

1 Mendelian Randomization Analysis of the Relationship Between 2 Native American Ancestry and Gallbladder Cancer Risk

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

27 **Abstract**

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

35
36 **Methods:** Markers of Mapuche ancestry were selected based on the informativeness
37 for assignment measure and then used as instrumental variables in two-sample
38 mendelian randomization (MR) analyses and complementary sensitivity analyses.

39
40 **Results:** We found evidence of a causal effect of Mapuche ancestry on GBC risk
41 (inverse variance-weighted (IVW) risk increase of 0.8% for every 1% increase in
42 Mapuche ancestry proportion, 95% CI 0.4% to 1.2%, $p = 6.6 \times 10^{-5}$). Mapuche
43 ancestry was also causally linked to gallstone disease (IVW risk increase of 3.6% per
44 1% increase in Mapuche proportion, 95% CI 3.1% to 4.0%, $p = 1.0 \times 10^{-59}$),
45 suggesting a mediating effect of gallstones in the relationship between Mapuche
46 ancestry and GBC. In contrast, the proportion of Mapuche ancestry showed a
47 negative causal effect on BMI (IVW estimate -0.006 kg/m^2 per 1% increase in
48 Mapuche proportion, 95% CI -0.009 to -0.003 , $p = 4.4 \times 10^{-5}$).

49
50 **Conclusions:** The results presented here may have significant implications for GBC
51 prevention and are important for future admixture mapping studies. Given that the
52 association between Mapuche ancestry and GBC risk previously noted in
53 observational studies appears to be causal, primary and secondary prevention
54 strategies that take into account the individual proportion of Mapuche ancestry could
55 be particularly efficient.

56

57 **Introduction**

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

63
64 The development of GBC is probably driven by a combination of environmental
65 exposures and genetic predisposition (Hundal & Shaffer, 2014). Symptoms are often
66 absent or unspecific until the disease has progressed to a non-curative stage, leaving
67 patients with few treatment options (Hundal & Shaffer, 2014; Roa & de Aretxabala,
68 2015; Stinton & Shaffer, 2012). As GBC is mostly diagnosed at an advanced stage,
69 the 5-year survival rate is low, reported at between 5% and 30% depending on the
70 country of origin of the study population (Bertran et al., 2010; Hundal & Shaffer,
71 2014; Kanthan et al., 2015; Stinton & Shaffer, 2012; Zhu et al., 2020). Gallbladder
72 removal (cholecystectomy) is a valuable tool for primary and secondary prevention of
73 GBC, but little progress has been made in individualized risk prediction and early
74 diagnosis (Roa & de Aretxabala, 2015).

75
76 The incidence of GBC shows wide geographical and ethnic variation (Mehrotra et al.,
77 2018). High-income regions such as Western Europe, the United States and
78 Australia have 1–2 cases per 100,000 person-years. In contrast, the largest Native
79 American people in Chile – the Mapuche – shows the highest incidence of GBC in
80 the world, with more than 20 cases per 100,000 person-years (Campbell et al.,
81 2017). Observational studies have found a strong association between the individual
82 proportion of Native American Mapuche ancestry and GBC risk: each 1% increase in
83 the proportion of Mapuche ancestry translates into a 3.7% increase in GBC mortality
84 (Lorenzo Bermejo et al., 2017). However, the observed association may arise from
85 other established GBC risk factors, especially gallstones and elevated body mass
86 index (BMI) (Hundal & Shaffer, 2014; Roa & de Aretxabala, 2015; Stinton & Shaffer,
87 2012). The Mapuche have a high prevalence of gallstone disease, a relative GBC
88 risk of 4.9 has been observed in individuals with a history of gallstones, and recent
89 studies have found evidence of a causal effect of gallstones on GBC risk for
90 genetically admixed Chileans (odds ratio [OR] = 1.97) (Barahona Ponce et al., 2021;

91 Mehrotra et al., 2018; Stinton & Shaffer, 2012). Native American ancestry has also
92 been associated with increased BMI: the World Cancer Research Fund considers
93 that higher body fatness marked by BMI probably causes GBC, and relative GBC
94 risks of 1.59 for women and 1.09 for men per five-point increase in BMI have been
95 reported (Campbell et al., 2017; Hundal & Shaffer, 2014; Roa & de Aretxabala, 2015;
96 World Cancer Research Fund/American Institute for Cancer Research, 2018).
97 Furthermore, multiparity in women, low socio-economic status, chronic inflammation
98 and lifestyle choices such as cigarette smoking and alcohol consumption could also
99 confound the association between ethnicity and GBC, justifying the use of MR to
100 assess the causal relationship between Mapuche ancestry and GBC risk (Lugo et al.,
101 2020; Roa & de Aretxabala, 2015; Stinton & Shaffer, 2012; Wistuba & Gazdar, 2004).

102
103 In response to the high GBC mortality rates and lack of treatment options, but based
104 on weak scientific evidence, the Chilean government is implementing a GBC
105 prevention program that provides financial support to gallstone patients for
106 prophylactic cholecystectomy (Roa & de Aretxabala, 2015). Prioritizing prophylactic
107 cholecystectomy for those at high risk of GBC, taking into account causal risk factors
108 rather than observational associations potentially attributable to confounding, would
109 optimize the efficiency of current GBC prevention measures. Mapuche ancestry per
110 se neither causes nor prevents disease, but statistical inference on the causal effect
111 of the individual's proportion of Mapuche ancestry on GBC risk allows for a distinction
112 between correlation and confounding-free causation, allowing for prioritization of
113 causal risk factors for prevention and a better understanding of the underlying
114 etiological mechanisms leading to GBC development.

115
116 With this goal in mind, we preselected markers of Mapuche ancestry in a panel of
117 Native American, European and African individuals, relying on the informativeness
118 for assignment measure (Rosenberg et al., 2003). We then used a subset of ancestry
119 markers as instrumental variables for the individual proportion of Mapuche ancestry
120 in two-sample MR to assess the causal effect of Mapuche ancestry on GBC risk. It
121 should be noted that population stratification, typically a limitation in MR studies, was
122 the exposure of interest in the present study.

123

124 Results

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

143
144 **Figure 2** depicts the genetic principal components, and the estimated proportions of
145 Mapuche, Aymara and European ancestry in the two samples used for the main MR
146 analysis. Individuals with a proportion of Mapuche ancestry above the 95th percentile
147 in the sample are represented in orange; blue and green highlight individuals with the
148 largest proportions of Aymara and European ancestry, respectively. The first principal
149 component in admixed Chileans distinguished between European and Native
150 American ancestry, and the second principal component separated the two main
151 types of Native American ancestry in Chile: Mapuche and Aymara. A strong
152 correlation between the second principal component and the proportion of Mapuche
153 ancestry was found in the two samples used for MR analysis (panels C and F in
154 **Figure 2**). In the three plots depicting sample II (panels D, E and F in **Figure 2**),
155 crosses represent GBC patients and circles represent population-based controls. The
156 concentration of GBC patients among individuals with the highest proportion of
157 Native American ancestry, especially Mapuche ancestry, was striking.

158 MR analysis of the association Mapuche ancestry → GBC risk revealed no
 159 heterogeneity among instrumental variables as a proxy for pleiotropy (inverse
 160 variance-weighted (IVW) Cochran’s Q statistic $p = 0.99$) and no horizontal pleiotropy
 161 (MR-Egger intercept $p = 0.87$; **Table 1**). Neither outliers nor weak instrument biases
 162 were apparent in the scatter and funnel plots (panels A and B in **Figure 3**). We found
 163 evidence of a causal effect of Mapuche ancestry on GBC risk (IVW OR = 1.008,
 164 which translates into a GBC risk increase of 0.8% for every 1% increase in the
 165 proportion of Mapuche ancestry, 95% confidence interval [CI] 0.4% to 1.2%, $p =$
 166 6.6×10^{-5}). MR-Egger (OR = 1.009, 95% CI 1.004 to 1.015, $p = 6.7 \times 10^{-4}$) and
 167 weighted median estimates (OR = 1.009, 95% CI 0.991 to 1.028, $p = 0.33$) were
 168 consistent with IVW estimates. As described above, the second principal component
 169 of genetic variability in admixed Chileans reflects the individual proportion of
 170 Mapuche ancestry, and, as expected, the causal effect vanished after including the
 171 second principal component in the calculation of summary statistics ($p = 0.62$). In
 172 contrast, consideration of the first and third to tenth principal components translated
 173 into stronger estimates of the causal effect (1.4% risk increase, 95% CI 0.8% to
 174 1.9%, $p = 5.9 \times 10^{-7}$).

175
 176 **Table 1** MR results on the causal relationship between Native American Mapuche ancestry
 177 as exposure and GBC, gallstone disease and BMI as outcomes of interest. Cochran’s Q
 178 statistic p-values higher than 0.05 are suggestive of no instrument heterogeneity as a proxy
 179 for pleiotropy, and MR-Egger intercept p-values higher than 0.05 are consistent with no
 180 horizontal pleiotropy. Results from sensitivity analyses on the association Mapuche ancestry
 181 → GBC risk after including the first to the tenth principal component, both with and without
 182 the second principal component, which reflects the individual proportion of Mapuche ancestry
 183 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

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

188

189 Further sensitivity analysis considered two-step MR and multivariable MR (MVMR) to
190 investigate the potential mediating effects of gallstone disease and BMI on the link
191 between Mapuche ancestry and GBC. For two-step MR mediation analysis, we first
192 assessed the causal effect of Mapuche ancestry on gallstone disease. Neither
193 heterogeneity among instrumental variables (Q statistic $p = 0.95$) nor horizontal
194 pleiotropy (MR-Egger intercept $p = 0.63$) was noted. We found evidence of a causal
195 effect of Mapuche ancestry on the risk of gallstone disease (IVW risk increase of
196 3.6% for every 1% increase in Mapuche proportion, $p = 1.0 \times 10^{-59}$; MR-Egger OR =
197 1.032, 95% CI 1.026 to 1.039, $p = 5.2 \times 10^{-24}$; weighted median OR = 1.031, 95% CI
198 1.010 to 1.051, $p = 0.003$). These results are visually represented in **Figure 3** (panels
199 C and D).

200
201 Since we found a causal effect of Mapuche ancestry on gallstones in the present
202 study, and a causal effect of gallstones on GBC risk has already been demonstrated
203 for Chileans, the two-step MR analysis suggests that gallstones mediate the
204 association between Mapuche ancestry and GBC risk. To validate this hypothesis,
205 we also performed MVMR. The instrumental variables used for MVMR were weak
206 (Mapuche ancestry F statistics = 3.49; gallstones F statistics = 1.37), but horizontal
207 pleiotropy was not found (Q statistic of 425 on 432 degrees of freedom, $p = 0.59$). In
208 agreement with the results of two-step MR, MVMR revealed a causal effect of
209 gallstones on GBC risk (OR = 1.263, $p = 2.0 \times 10^{-9}$) but no concurrent causal effect of
210 Mapuche ancestry on GBC risk (OR=1.0001, $p = 0.95$).

211
212 Neither heterogeneity among instrumental variables (Q statistic $p = 0.96$) nor
213 horizontal pleiotropy (MR-Egger intercept $p = 0.90$) was detected in MR analysis of
214 the association Mapuche ancestry \rightarrow BMI. We found evidence of a negative causal
215 effect of Mapuche ancestry on BMI (IVW $\beta = -0.006$ kg/m² per 1% increase in
216 Mapuche proportion, 95% CI -0.009 to -0.003, $p = 4.4 \times 10^{-5}$). The corresponding
217 scatter and funnel plots are shown in panels E and F of **Figure 3**. MR-Egger ($\beta = -$
218 0.007 kg/m² per 1% Mapuche ancestry proportion, 95% CI -0.020 to 0.006, $p = 0.31$)
219 and weighted median estimates ($\beta = -0.006$ kg/m² per 1% Mapuche ancestry
220 proportion, 95% CI -0.011 to -0.002, $p = 3.8 \times 10^{-3}$) showed good agreement with the
221 negative causal effect estimated by IVW.

222 Discussion

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

239
240 Gallstones are highly prevalent in Chileans and one of the most important risk factors
241 for GBC: observational studies report a relative GBC risk of 9.2–10.1 for individuals
242 with gallstones larger than 3 cm (Miquel et al., 1998; Stinton & Shaffer, 2012). We
243 therefore examined the possible mediating effect of gallstone disease on the
244 association between Mapuche ancestry and GBC risk. The evidence for a causal
245 effect of Mapuche ancestry on gallstone disease was quite robust: the estimated
246 causal OR was 1.036 per 1% Mapuche ancestry proportion. This finding, together
247 with the previously reported causal effect of gallstones on GBC risk in Chileans,
248 support a mediating effect of gallstones based on two-step MR (Barahona Ponce et
249 al., 2021). The MVMR results, based on a relatively small sample size and weak
250 instrumental variables, showed good agreement with two-step MR and also
251 supported a mediating effect of gallstones between Mapuche ancestry and GBC risk;
252 however, large prospective Chilean datasets with complete information on GBC,
253 gallstones and individual genotype data are urgently needed to conduct formal
254 mediation analyses and accurately quantify the magnitude of the direct and indirect
255 effects of gallstones, which are highly relevant for more precise GBC prevention.

256

257 The International Agency for Research on Cancer, the American Institute for Cancer
258 Research and the World Cancer Research Fund all consider obesity to be a likely
259 cause of GBC, and high BMI is another important GBC risk factor, with a direct effect
260 on GBC risk reported for Chileans and an indirect effect mediated by gallstones for
261 Europeans (Barahona Ponce et al., 2021; Lauby-Secretan et al., 2016; World Cancer
262 Research Fund/American Institute for Cancer Research, 2018). However, studies of
263 the association between Native American ancestry and BMI have yielded
264 contradictory results. BMI decreased by 0.13 m/kg² for each 10% increase in the
265 proportion of Native American ancestry in Mexican–American women (Hu et al.,
266 2015), and the association was also negative in a Hispanic–Mexican study (Tang et
267 al., 2006). In contrast, BMI increased by 0.56 m/kg² per 10% increase in the
268 proportion of Native American ancestry in Native Americans (Norden-Krichmar et al.,
269 2014). Native American ancestry was positively correlated with obesity in Mexico and
270 Peru, whereas no association was found in Brazil, Chile, and Colombia, and higher
271 Native American ancestry was associated with overweight and obesity, but only
272 among foreign-born Latina women (Ziv et al., 2006). In this study we found clear
273 evidence of a negative causal effect of Mapuche ancestry on BMI, probably ruling out
274 a positive mediating effect of BMI on the association between Mapuche ancestry and
275 GBC risk.

276

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

289

290

291 Previous studies have demonstrated the importance of subdividing Native American
292 ancestry into its main subcomponents –in the case of admixed Chileans, the two
293 major Native American subcomponents are Mapuche and Aymara ancestry (Lorenzo
294 Bermejo et al., 2017). While combined Native American ancestry showed no
295 association with GBC mortality and the proportion of Aymara ancestry showed a
296 negative association with GBC mortality, each 1% increase in the Mapuche
297 proportion translated into a 3.7% increase in the risk of death due to GBC (95% CI
298 3.1 to 4.3%). To assess the contribution of health system access to the identified
299 association between Mapuche ancestry and GBC risk, hospitalization rates due to
300 gallbladder removal (cholecystectomy) were included as an additional explanatory
301 variable in the fitted regression models; this showed a minor role of health system
302 access in the estimated increase in GBC mortality risk. The MR-based OR estimated
303 in the present study confirmed a causal effect of Mapuche ancestry on GBC risk,
304 while reducing the effect size to 0.8% (95% CI 0.4% to 1.2%). This result suggests
305 that GBC prevention strategies that consider the individual proportion of Mapuche
306 ancestry may be particularly efficient, and that the difference between the
307 observational (3.7%) and causal (0.8%) increase in GBC risk is due to confounders
308 of the association between Mapuche ancestry and GBC risk, which may include
309 modifiable risk factors. Further research is warranted.

310

311 Beyond the practical relevance for GBC prevention, the results presented here may
312 also be useful for designing future studies. On the one hand, the evidence of a
313 causal effect of Mapuche ancestry on GBC risk underpins the potential of admixture
314 mapping to identify novel GBC susceptibility variants, possibly in combination with
315 subsequent association testing (note that the genome-wide significance level is much
316 higher for admixture mapping than for association mapping) (Horimoto et al., 2021).
317 For example, assuming that the average proportion of Mapuche ancestry in Chileans
318 is about 40%, approximately 1100 GBC patients are needed to detect a Mapuche
319 haplotype with a risk ratio of 1.8, consistent with the estimated causal OR of 1.008
320 per 1% proportion of Mapuche ancestry in the whole Chilean genome (McKeigue,
321 2005). On the other hand, the 3.7% increased GBC mortality risk per 1% Mapuche
322 ancestry proportion previously reported in observational studies probably
323 overestimates the contribution of Mapuche ancestry to GBC risk and would lead to

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

327

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

346

347

348 **Materials and Methods**

349 **Ethics approval**

350 The study protocol conformed to the ethical guidelines of the 1975 Declaration of
351 Helsinki and was approved by the ethics committees of Servicio de Salud
352 Metropolitano Oriente, Santiago de Chile (#06.10.2015, #08.03.2016 and
353 #12.11.2019), Servicio de Salud Metropolitano Sur Oriente, Santiago de Chile
354 (#15.10.2015 and #05.04.2018), Servicio de Salud Metropolitano Central, Santiago
355 de Chile (#1188-2015), Servicio de Salud Coquimbo, Coquimbo, Chile
356 (#01.04.2016), Servicio de Salud Maule, Talca, Chile (#05.11.2015), Universidad
357 Católica del Maule, Talca, Chile (#102-2020), Servicio de Salud Concepción,
358 Concepción, Chile (ID: 16-11-97 and ID:19-12-111), Servicio de Salud Araucanía
359 Sur, Temuco, Chile (#10.02.2020), Servicio de Salud Valdivia, Valdivia, Chile
360 (ID:438), Centro de Bioética, Universidad del Desarrollo, Clínica Alemana de
361 Santiago, Santiago de Chile (#2018-97, ID 678) and Unidad de Investigación
362 Hospital San Juan de Dios, Santiago de Chile (#6182), the Medical Faculties of
363 Universidad de Chile (approval #123-2012 and #11.10.2012) and Pontificia
364 Universidad Católica de Chile (#11-159), and Universidad de Tarapacá and
365 University College London as described in (Ruiz-Linares et al., 2014). All participants
366 provided written informed consent before enrolment.

367

368 **Study participants**

369 The Chilean genome is a mixture of chromosomal segments from two major Native
370 American peoples, the Aymara in the north and the Mapuche in the south of the
371 country; Europeans; and, to a lesser extent, Africans (Eyheramendy et al., 2015;
372 Hundal & Shaffer, 2014; Ruiz-Linares et al., 2014). The reference panel for
373 preselection of ancestry-informative markers and estimation of individual ancestry
374 proportions in genetically admixed Chileans therefore included 63 Aymara (Reich et
375 al., 2012) and 28 Mapuche (Lindo et al., 2018; Reich et al., 2012), as well as 206
376 Europeans (99 Utah residents with Northern and Western European ancestry [CEU]
377 and 107 Iberians from Spain [IBS]), and 108 African Yorubans from Ibadan, Nigeria
378 (YRI) from the 1000 Genome Project (The 1000 Genomes Project Consortium,
379 2015).

380

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

402

403 **Preselection of markers of Mapuche ancestry**

404 To preselect genetic variants robustly associated with our exposure of interest, the
405 individual proportion of Mapuche ancestry, we first chose markers of Mapuche
406 ancestry based on the informativeness for assignment measure I_n (Rosenberg et al.,
407 2003). For each genetic variant with $j = 1, \dots, N$ possible alleles in $i = 1, \dots, K$
408 subpopulations we calculated:

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

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

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

420

421 **Selection of instrumental variables**

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

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

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

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

442

443

444

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

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

463
464 Population stratification is particularly relevant in MR studies of genetically admixed
465 individuals. To deal with potential stratification, association statistics are typically
466 adjusted for the main principal components of genetic variability. However, the
467 individual proportion of Mapuche ancestry was the exposure of interest in this study
468 of Chileans, who show continuous gradients of ancestry. To check the sensitivity of
469 MR results to population stratification, we estimated the genetic principal components
470 in sample I and sample II using the *eigenstrat* function (Patterson et al., 2006);
471 examined the correlation between the estimated principal components and the
472 proportion of Mapuche ancestry; adjusted the association statistics for (1) all of the
473 first ten principal components or (2) the first ten principal components with the
474 exception of the principal component correlated with the proportion of Mapuche
475 ancestry; and repeated the MR analyses.

476
477 Since a causal effect of gallstones and BMI on GBC risk has recently been reported
478 (Barahona Ponce et al., 2021), we investigated the mediating effects of gallstone

479 disease and BMI as a surrogate marker for obesity, as potential mechanisms
480 explaining the relationship between Mapuche ancestry and GBC risk. The
481 unavailability of large prospective Chilean datasets with complete information on
482 GBC, gallstone, BMI and individual genotypes precluded the implementation of
483 formal mediation analyses, and we decided first to apply two-step MR to assess
484 mediation (Zheng et al., 2017). In the first MR step, IV for the proportion of Mapuche
485 ancestry were used to assess the causal effect of the Mapuche ancestry proportion
486 on the potential mediator (gallstone disease or BMI) by MR as described above. The
487 second MR step was based on published findings, also based on genetically
488 admixed Chileans, on the causal effect of the mediators on GBC risk. Evidence of
489 association in the two MR steps (Mapuche ancestry \rightarrow gallstone disease and
490 gallstone disease \rightarrow GBC) would imply some degree of mediation between Mapuche
491 ancestry and GBC risk on the part of the intermediate trait (gallstones). We also
492 performed MVMR to simultaneously assess the causal effects of Mapuche ancestry
493 and gallstones on GBC risk. MVMR allows the estimation of mediation effects when
494 considering potentially correlated risk factors utilizing genetic instruments within the
495 MR framework. We used the R package 'MVMR', following the workflow suggested
496 by the software developers (Sanderson et al., 2021). We calculated F statistics to
497 monitor instrument strength, Q statistics to assess instrument validity including
498 horizontal pleiotropy, and finally estimated the direct effects of the considered
499 exposures (Mapuche ancestry and gallstone disease) simultaneously on the outcome
500 of interest (GBC risk). The IV used included the previously selected IV for Mapuche
501 ancestry along with five genetic variants (one palindromic variant was removed)
502 robustly associated with gallstone disease in Chileans (Joshi et al., 2016). The
503 sample I and sample II used to assess the association between IV and Mapuche
504 ancestry, and between IV and gallstone disease, did not overlap. They consisted of
505 1,703 Chileans (1,861 minus those used as controls in sample II) and 351 gallstone
506 patients and matched controls, respectively. The sample used to assess the direct
507 effect between IV and gallstone disease on GBC risk was the same as the sample II
508 utilized for the MR analysis of Mapuche ancestry \rightarrow GBC risk (412 Chilean GBC
509 cases and 412 matched controls).

510
511 MR analyses were conducted using the R version of MR-Base, which provides
512 convenient tools for the harmonization of the association statistics, including

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

520

521 References

- 522 Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of
523 ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.
524 <https://doi.org/10.1101/gr.094052.109>
- 525 Barahona Ponce, C., Scherer, D., Brinster, R., Boekstegers, F., Marcelain, K.,
526 Gárate-Calderón, V., Müller, B., de Toro, G., Retamales, J., Barajas, O.,
527 Ahumada, M., Morales, E., Rojas, A., Sanhueza, V., Loader, D., Rivera, M. T.,
528 Gutiérrez, L., Bernal, G., Ortega, A., ... Lorenzo Bermejo, J. (2021). Gallstones,
529 Body Mass Index, C-Reactive Protein, and Gallbladder Cancer: Mendelian
530 Randomization Analysis of Chilean and European Genotype Data. *Hepatology*,
531 *73*(5), 1783–1796. <https://doi.org/10.1002/hep.31537>
- 532 Bertran, E., Heise, K., Andia, M. E., & Ferreccio, C. (2010). Gallbladder cancer:
533 Incidence and survival in a high-risk area of Chile. *International Journal of*
534 *Cancer*, *127*(10), 2446–2454. <https://doi.org/10.1002/ijc.25421>
- 535 Blandino, A., Scherer, D., Rounge, T. B., Umu, S. U., Boekstegers, F., Barahona
536 Ponce, C., Marcelain, K., Gárate-Calderón, V., Waldenberger, M., Morales, E.,
537 Rojas, A., Munoz, C., Retamales, J., de Toro, G., Barajas, O., Rivera, M. T.,
538 Cortés, A., Loader, D., Saavedra, J., ... Bermejo, J. L. (2022). Identification of
539 Circulating lncRNAs Associated with Gallbladder Cancer Risk by Tissue-Based
540 Preselection, Cis-eQTL Validation, and Analysis of Association with Genotype-
541 Based Expression. *Cancers*, *14*(3). <https://doi.org/10.3390/cancers14030634>
- 542 Bowden, J., Spiller, W., Del Greco, F. M., Sheehan, N., Thompson, J., Minelli, C., &
543 Smith, G. D. (2018). Improving the visualization, interpretation and analysis of
544 two-sample summary data Mendelian randomization via the Radial plot and
545 Radial regression. *International Journal of Epidemiology*, *47*(4), 1264–1278.
546 <https://doi.org/10.1093/ije/dyy101>
- 547 Campbell, P. T., Newton, C. C., Kitahara, C. M., Patel, A. V., Hartge, P., Koshiol, J.,
548 McGlynn, K. A., Adami, H.-O., Berrington De Gonzalez, A., Beane Freeman, L.
549 E., Bernstein, L., Buring, J. E., Freedman, N. D., Gao, Y. T., Giles, G. G.,
550 Gunter, M. J., Jenab, M., Liao, L. M., Milne, R. L., ... Gapstur, S. M. (2017).
551 Body Size Indicators and Risk of Gallbladder Cancer: Pooled Analysis of
552 Individual-Level Data from 19 Prospective Cohort Studies. *Cancer Epidemiology*
553 *Biomarkers and Prevention*, *26*(4), 597–606. <https://doi.org/10.1158/1055-9965.EPI-16-0796>
- 554 Eyheramendy, S., Martinez, F. I., Manevy, F., Vial, C., & Repetto, G. M. (2015).
555 Genetic structure characterization of Chileans reflects historical immigration
556 patterns. *Nature Communications*, *6*, 6472. <https://doi.org/10.1038/ncomms7472>
- 557 Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., Laurin,
558 C., Burgess, S., Bowden, J., Langdon, R., Tan, V. Y., Yarmolinsky, J., Shihab,
559 H. A., Timpson, N. J., Evans, D. M., Relton, C., Martin, R. M., Davey Smith, G.,
560 Gaunt, T. R., & Haycock, P. C. (2018). The MR-Base platform supports
561 systematic causal inference across the human phenome. *ELife*, *7*, e34408.
562 <https://doi.org/10.7554/eLife.34408>
- 563 Horimoto, A. R. V. R., Xue, D., Thornton, T. A., & Blue, E. E. (2021). Admixture
564 mapping reveals the association between Native American ancestry at 3q13.11
565 and reduced risk of Alzheimer's disease in Caribbean Hispanics. *Alzheimer's*
566 *Research and Therapy*, *13*. <https://doi.org/10.1186/s13195-021-00866-9>
- 567 Hu, H., Huff, C. D., Yamamura, Y., Wu, X., & Strom, S. S. (2015). The Relationship
568 between Native American Ancestry, Body Mass Index and Diabetes Risk among
569 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>

- 622 Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and
623 eigenanalysis. *PLoS Genetics*, 2(12), 2074–2093.
624 <https://doi.org/10.1371/journal.pgen.0020190>
- 625 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D.,
626 Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007).
627 PLINK: A Tool Set for Whole-Genome Association and Population-Based
628 Linkage Analyses. *American Journal of Human Genetics*, 81(3), 559–575.
629 <https://doi.org/10.1086/519795>
- 630 Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.
631 V., Rojas, W., Duque, C., Mesa, N., García, L. F., Triana, O., Blair, S., Maestre,
632 A., Dib, J. C., Bravi, C. M., Bailliet, G., Corach, D., Hünemeier, T., ... Ruiz-
633 Linares, A. (2012). Reconstructing Native American population history. *Nature*,
634 488(7411), 370–374. <https://doi.org/10.1038/nature11258>
- 635 Roa, I., & de Aretxabala, X. (2015). Gallbladder cancer in Chile: what have we
636 learned? *Current Opinion in Gastroenterology*, 31(3), 269–275.
637 <https://doi.org/10.1097/MOG.000000000000164>
- 638 Rosenberg, N. A., Li, L. M., Ward, R., & Pritchard, J. K. (2003). Informativeness of
639 Genetic Markers for Inference of Ancestry. *American Journal of Human*
640 *Genetics*, 73(6), 1402–1422. <https://doi.org/10.1086/380416>
- 641 Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C.,
642 Arias, W., Fuentes, M., Pizarro, M., Everardo, P., de Avila, F., Gómez-Valdés, J.,
643 León-Mimila, P., Hunemeier, T., Ramallo, V., Silva de Cerqueira, C. C., Burley,
644 M. W., Konca, E., de Oliveira, M. Z., Veronez, M. R., ... Gonzalez-José, R.
645 (2014). Admixture in Latin America: Geographic Structure, Phenotypic Diversity
646 and Self-Perception of Ancestry Based on 7,342 Individuals. *PLoS Genetics*,
647 10(9). <https://doi.org/10.1371/journal.pgen.1004572>
- 648 Sanderson, E., Spiller, W., & Bowden, J. (2021). Testing and correcting for weak and
649 pleiotropic instruments in two-sample multivariable Mendelian randomization.
650 *Statistics in Medicine*, 40(25), 5434–5452. <https://doi.org/10.1002/sim.9133>
- 651 Stinton, L. M., & Shaffer, E. A. (2012). Epidemiology of Gallbladder Disease:
652 Cholelithiasis and Cancer. *Gut and Liver*, 6(2), 172–187.
653 <https://doi.org/10.5009/gnl.2012.6.2.172>
- 654 Tang, H., Jorgenson, E., Gadde, M., Kardia, S. L. R., Rao, D. C., Zhu, X., Schork, N.
655 J., Hanis, C. L., & Risch, N. (2006). Racial admixture and its impact on BMI and
656 blood pressure in African and Mexican Americans. *Human Genetics*, 119(6),
657 624–633. <https://doi.org/10.1007/s00439-006-0175-4>
- 658 The 1000 Genomes Project Consortium. (2015). A global reference for human
659 genetic variation. *Nature*, 526(7571), 68–74. <https://doi.org/10.1038/nature15393>
- 660 The Global Cancer Observatory. (2020). *Global Cancer Observatory: Cancer Today -*
661 *Gallbladder cancer fact sheet*. Gco.iarc.fr.
662 <https://gco.iarc.fr/today/data/factsheets/cancers/12-Gallbladder-fact-sheet.pdf>
- 663 Wistuba, I. I., & Gazdar, A. F. (2004). Gallbladder cancer: lessons from a rare
664 tumour. *Nature Reviews Cancer*, 4(9), 695–706. <https://doi.org/10.1038/nrc1429>
- 665 World Cancer Research Fund/American Institute for Cancer Research. (2018). *Diet,*
666 *Nutrition, Physical Activity and Cancer: A Global Perspective. Continuous*
667 *Update Project Expert Report 2018*. Dietandcancerreport.Org.
668 <http://dietandcancerreport.org>
- 669 Zheng, J., Baird, D., Borges, M.-C., Bowden, J., Hemani, G., Haycock, P., Evans, D.
670 M., & Davey Smith, G. (2017). Recent Developments in Mendelian
671 Randomization Studies. *Current Epidemiology Reports*, 4(4), 330–345.
672 <https://doi.org/10.1007/s40471-017-0128-6>

- 673 Zhu, X., Zhang, X., Hu, X., Ren, H., Wu, S., Wu, J., Wu, G., Si, X., & Wang, B.
674 (2020). Survival analysis of patients with primary gallbladder cancer from 2010
675 to 2015: A retrospective study based on SEER data. *Medicine*, 99(40), e22292.
676 <https://doi.org/10.1097/MD.00000000000022292>
- 677 Ziv, E., John, E. M., Choudhry, S., Kho, J., Lorizio, W., Perez-Stable, E. J., &
678 Burchard, E. G. (2006). Genetic ancestry and risk factors for breast cancer
679 among Latinas in the San Francisco Bay Area. *Cancer Epidemiology Biomarkers*
680 *and Prevention*, 15(10), 1878–1885. [https://doi.org/10.1158/1055-9965.EPI-06-](https://doi.org/10.1158/1055-9965.EPI-06-0092)
681 [0092](https://doi.org/10.1158/1055-9965.EPI-06-0092)
- 682

683 **Figure legends**

684

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

689

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

696

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

701

702 **Funding**

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

718

719 **Conflict of Interest Statement**

720 The authors declare that they have no conflicts of interest.

721

722 **Acknowledgements**

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

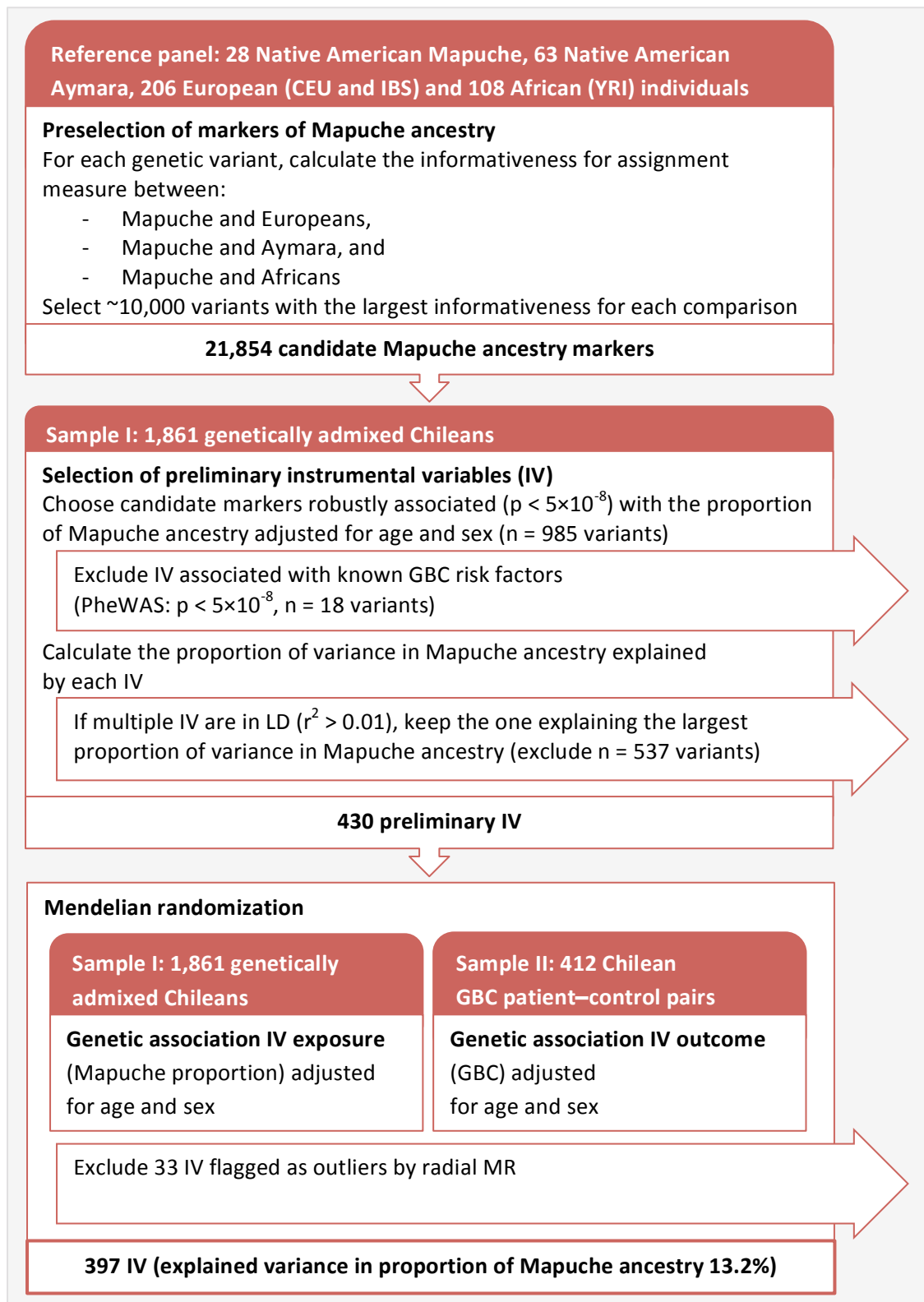


Figure 1 Flowchart describing the main analyses, from the selection of markers of Mapuche ancestry in a reference panel composed of Mapuche, Aymara, European and African individuals, to the two-sample MR based on 397 selected instrumental variables after exclusion of outlying instruments based on radial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

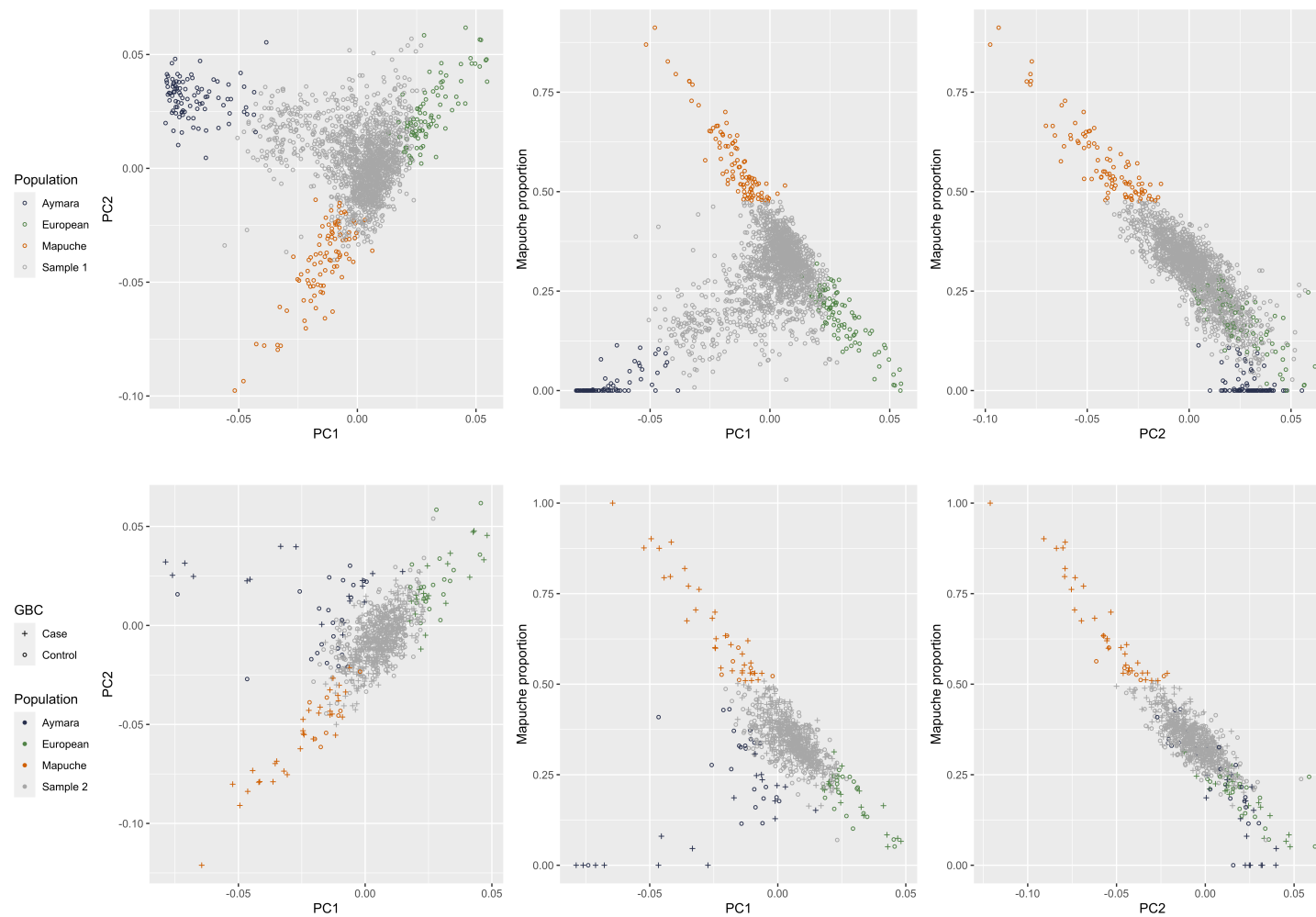


Figure 3 Scatter and funnel plots from MR analyses after exclusion of outlying instruments according to radial MR on the association between Mapuche ancestry and GBC risk (panels A and B), gallstone disease (panels C and D) and BMI (panels E and F)

