Received: 3 October 2016 | Revised: 12 December 2016 | Accepted: 1 February 2017

DOI 10.1002/ajhb.22982

SHORT REPORT

WILEY American Journal of Human Biology

Molecular polymorphisms of the ABO locus as informative markers of ancestry in Central Argentina

María Pía Tavella | Angelina García | Maia Pauro | Darío A. Demarchi | Rodrigo Nores

Instituto de Antropología de Córdoba (IDACOR), CONICET/Universidad Nacional de Córdoba, Córdoba 5000, Argentina

Correspondence
Dr. Rodrigo Nores, Museo de Antropología UNC, Av. Hipólito Yrigoyen 174, Córdoba 5000, Argentina.
Email: rodrigonores@hotmail.com

Funding information
Contract grant sponsors: PICT 2012 N° 711, PICT 2013 N° 1715

Abstract

Objectives: The aim of this study was to investigate the distribution of molecular polymorphisms of the ABO gene in four population samples from the province of Córdoba, in Central Argentina, and to compare them with other worldwide populations.

Methods: A total of 110 buccal swab samples from autochthonous individuals of Córdoba were typified. Molecular characterization of the allelic variants was performed by the analysis of exons 6 and 7 of the ABO gene using PCR-RFLP analysis. Additionally, the Native American AIM O1v542 was characterized by direct sequencing.

Results: The four Córdoba populations did not show significant geographic structure, although the frequency of the O1v542 haplotype, detected in all the populations studied, ranged from 0.019 to 0.222. The principal component analysis based on O allele distribution showed that the populations from Córdoba clustered close to the admixed populations of Santiago and Mexico City, and at intermediate distances between European and Native American populations, while being distant from the African population.

Conclusions: The results demonstrate that the analysis of the ABO system constitutes a useful tool for the study of the genetic structure and evolutionary history of human populations, reflecting accurately the relative contribution of parental continental contribution to the gene pool of admixed populations.

1 | INTRODUCTION

Since its discovery in 1900 by Karl Landsteiner, the ABO blood group system has been used as a marker of variation between populations in anthropological studies, mainly because it is highly polymorphic. The molecular basis of the gene encoding the glycosyltransferases that synthesize ABO antigens was established by Yamamoto, Clausen, White, Marken, & Hakomori (1990), and the main alleles were defined. Since then, a large number of ABO variants have been described, and new alleles are constantly being identified. Excluding the common alleles, about half of the remaining are because of new mutations and the other half can be better explained by intragenic recombination between common variants (Yip, 2002).

Although ABO polymorphisms are distributed worldwide, there is a remarkable variation in their frequency distribution among continental populations. Surprisingly, the O group, although not functional, is the most abundant phenotype worldwide, ranging from 0.61 to 0.98, and being almost fixed in Native Americans. In addition the highest frequencies of allele O02 (also called O1v) are found among these populations. A mutation derived from the O02 allele in the 542 position of exon 7 has been consistently described in both contemporary and precontact Native American populations in different frequencies over the continent (Villanea et al., 2013). These findings, taken together with the rarity of the O1v542 allele in non-American populations, support its use as an Ancestry Informative Marker (AIM) (Estrada-Mena et al., 2009). Present knowledge of the peopling of the

Americas suggests that this mutation would have occurred within an Asian migrant group, ancestral to all Native-Americans, during the isolation period in Beringia, prior to entering into the continent about 16,000 ybp (Villanea et al., 2013) in agreement with the Beringian Incubation Model (Tamm et al., 2007).

The population of Argentina, like those of other Latin American countries, has diverse ethnic origins. The immigration process can be divided in two main periods: (a) early settlements at the first half of the 16th century until the end of colonial times (circa 1810), characterized by the arrival of the Spaniards to the territory and by the traffic of African slaves; and (b) the second half of the 19th and early 20th centuries, marked by population expansion resulting from massive European immigration (mainly from Spain and Italy), favored by a process of economic internationalization (Pellegrino, 2002). All of these people merged in a complex process, as evident by the heterogeneous degrees of admixture observed in different regions of the country.

The ABO blood group has been studied to estimate admixture proportions and to investigate gene flow in different regions of Argentina, both from phenotypic (Avena et al., 2010; Morales, Dipierri, Alfaro, & Bejarano, 2000) and molecular data (Vaccaro et al., 2011), but not in the central region of the country. The aim of this study is to investigate the distribution of molecular polymorphisms of the ABO gene in four populations from the province of Córdoba and to compare them with other worldwide populations, as a contribution to the knowledge of the recent evolutionary history of central Argentina.

2 | MATERIALS AND METHODS

2.1 | Samples

A total of 117 buccal swab samples were obtained from healthy, unrelated individuals from four villages of Córdoba province. These include La Para (30.87S, 62.98W), located in the northeast near the Mar Chiquita lagoon; San Marcos Sierras (30.75S, 64.56W) and Villa de Soto (30.85S, 64.98W), settlements of the Sierras Pampeanas region; and Jovita (34.52S, 63.94W), a town located in the southwestern Pampas plains (Figure 1). Individuals included in the study had a family history of at least two generations in their residential location. Sampling was performed following the Helsinki protocol, with the informed consent of each donor. Our research program was approved by the Research Ethics Committee of CEMIC (Centro de Educación Médica en Investigaciones Clínicas "Norberto Quirno").

2.2 | DNA extraction and ABO genotyping

Genomic DNA was extracted from cheek swabs using the Accuprep Genomic DNA Kit (GenBiotech). The ABO genotypes were first characterized by PCR-RFLP, following a described method (Hummel, Schmidt, Kahle, & Herrmann, 2002). Two fragments of 103/104 bp and 64 bp from exons 6 and 7, respectively, were amplified in order to distinguish the A101, B101, O01, O02, and O03 alleles. Eight μl of the PCR products were digested sequentially with the corresponding restriction enzymes, that is, the exon 6's fragment in the first place and then exon 7, if necessary.

In a second stage, in order to trace the G542A mutation, a fragment corresponding to exon 7 from those samples that presented the allele O02, both in homozygous and heterozygous states, was amplified with primers ABO-3 and ABO-8 (Ogasawara et al., 1996). PCR products were purified and both strands were sequenced separately by Macrogen (Korea). Sequences were manually aligned using Allele A101 (GenBank AC000397) as reference sequence, using Sequencher 5.3 (Gene Codes Corporation). As there is no consensus terminology for ABO alleles, we followed the nomenclature proposed by the Blood Group Antigen Gene Mutation Database (http://www.ncbi.nlm.nih.gov/projects/mhc/xslcgi.fcgi?cmd = bgmut/home), except in the case of the allele O11, characterized by the mutation in position 542, where the name "O1v542" is more informative.

2.3 | Statistical analyses

Allele frequency estimates, Hardy-Weinberg equilibrium, and the Analysis of Molecular Variance (AMOVA) were carried out using the Arlequin v.3.5 software package (Excoffier & Lischer, 2010). To evaluate genetic affinities among the studied samples and other populations data obtained from the literature, a principal component analysis based on O alleles distribution was performed using the software Past 3.09 (Hammer, Harper, & Ryan, 2001).

3 | RESULTS AND DISCUSSION

Of the 117 samples, 110 were successfully genotyped for the ABO polymorphisms. A total of 62 sequences were obtained, including 44 from individuals that presented the O02 allele and 18 problematic samples that could not be genotyped by RFLP.

The distribution of ABO genotypes and allele frequencies in the four populations and in the total population, as well as the Hardy-Weinberg test results and observed and expected heterozygosis, are presented in Table 1. Deviation from Hardy-Weinberg's expectation was only observed in La Para, most likely because of a bias generated by the small sample size. The most frequent allele was O01, in line with a global tendency, except for La Para. Allele O02 was the second most frequent variant, a result in agreement with the high incidence of native mitochondrial lineages in Córdoba (76%) and the important Native American contribution in its

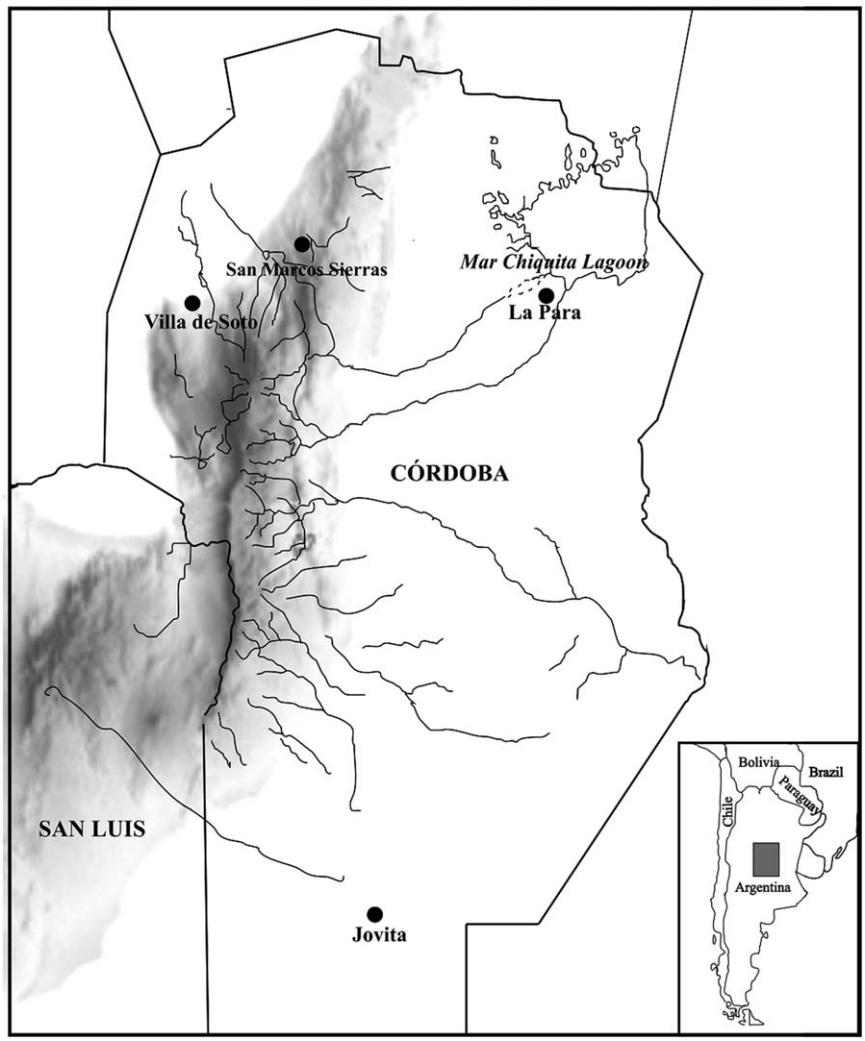

FIGURE 1 Map of Córdoba province, Argentina, in the southern cone of South America, showing the four locations sampled for this study

gene pool (40%) according to autosomal markers (García et al., 2015).

Three hypotheses have been proposed to explain the high frequency of O blood type in Native Americans: first, a founder effect during the initial peopling of the Americas; second, genetic drift related to European contact-associated population decline; and third, natural selection of the O blood group associated with some selective advantage in response to disease epidemics following European contact. In this regard, some pathogens mimic ABO antigens, which makes it more difficult for the immune system to detect and destroy them. Blood type A may be more susceptible to smallpox, for example, because a substance resembling the A antigen is present in the Variola virus. So, it could be assumed that the repeated and widespread smallpox epidemics of the 16th–19th centuries may have selected against Native Americans with A alleles and increased the frequency of O and B alleles in the surviving population (Halverson & Bolnick, 2008). However, further investigations show no changes in ABO allele frequencies between precontact and extant populations (Georges, Seidenberg, Hummel, & Fehren-Schmitz, 2012; Halverson & Bolnick, 2008). These results suggest that events following European contact had little impact on ABO frequencies in the Americas and support the founder effect hypothesis, which is more compatible with the evidence from other genetic loci. The fact that haplotypes O1 and O1v both constitute null alleles and, in principle, have identical selective values, supports the notion that chance could be responsible for both the predominance of O1v haplotypes and the O group fixation (Estrada-Mena et al., 2009).

In this regard, the AIM O1v542, introduced from Beringia by the population ancestral to most Native Americans (Villanea et al., 2013), was detected in all of the four populations studied, in variable frequencies. San Marcos Sierras shows the highest frequency (0.222), consistent with the demographic history of the Sierras Pampeanas, since this region used to have the highest concentration of native population settlements before the arrival of the Spaniards. The frequency of O1v542 in San Marcos Sierras is comparable to the average incidence of this polymorphism in the Native American

TABLE 1 Distribution of genotypic and allelic frequencies and Hardy-Weinberg's equilibrium (HWE) test in the studied populations

	Frequencies %										
ABO Genotypes	**La Para (*N*=29)**	**Villa de Soto (*N*=28)**	**San Marcos (*N*=27)**	**Jovita (*N*=26)**	**Total (*N*=110)**	**ABO Alleles**	**La Para**	**Villa de Soto**	**San Marcos**	**Jovita**	**Total**
A^1A^1	6.9	0	7.4	11.5	6.4	A101	0.190	0.179	0.167	0.269	0.200
A^1O^1	10.3	25	3.7	23	15.4	A201	0.017	0	0	0.019	0.009
A^1O^2	13.8	3.6	7.4	3.7	7.2	B101	0.052	0.018	0.019	0.038	0.032
A^1O^{1v542}	0	7.1	7.4	3.7	4.5	O01	0.310	0.429	0.278	0.308	0.332
A^2B^1	3.4	0	0	3.7	1.8	O02	0.362	0.304	0.259	0.288	0.305
B^1O^2	3.4	3.6	3.7	3.7	3.6	O1v542	0.035	0.071	0.222	0.019	0.086
B^1O^{1v542}	0	0	0	0	0	O03	0.017	0	0.019	0	0.009
B^1O^3	3.4	0	0	0	0.9	O03(649 + 689)	0.017	0	0	0	0.005
O^1O^1	17.2	17.9	11.1	7.7	13.6	O09	0	0	0.037	0	0.009
O^1O^2	17.2	17.9	14.8	19.2	17.3	O12	0	0	0	0.038	0.009
O^1O^{1v542}	0	3.6	11.1	0	3.6	O26	0	0	0	0.019	0.005
O^1O^3	0	0	3.7	0	0.9	**Obs. Het.**	0.586	0.679	0.63	0.692	
O^2O^2	17.2	14.3	7.4	11.3	10.8	**Exp. Het.**	0.745	0.699	0.791	0.760	
O^2O^{1v542}	3.4	7.1	11.1	0	5.4	**HWE p-value**	0.032[a]	0.287	0.291	0.173	
$O^{1v542}O^{1v542}$	0	0	7.4	0	1.8						
$O^{3v}O^{1v542}$	3.4	0	0	0	0.9						
O^9O^9	0	0	3.7	0	0.9						
$O^{12}O^1$	0	0	0	3.7	0.9						
$O^{12}O^2$	0	0	0	3.7	0.9						
$O^{26}O^2$	0	0	0	3.7	0.9						

[a]Statistically significant $P < 0.05$.

populations studied to date (Villanea et al., 2013). In contrast, the lowest frequency of this AIM was found in Jovita (0.019), a rural village of the southeastern plains of Córdoba province, where many European immigrants settled.

Some uncommon alleles of European origin were detected in low frequencies, such as O09, O12, O26, and O03. One particularly rare haplotype at a global scale, O03(649 + 689), was found in La Para. This variant was described by Hosseini-Maaf et al. (2005), when they screened samples presenting the allele O03 with an unclear serological blood group. They found that, although C649T polymorphism is a rare occurrence, G689A was present in 17% of the O03 screened chromosomes, and concluded that many of the rare novel alleles constitute a risk of error in all ABO genotyping methods used to date because of its wide range of serologic activity.

Taking into account that Native Americans nearly exclusively belong to group O and have low genetic variability at the ABO locus (Georges et al., 2012), alleles other than O01 and O02 (and its variant O1v542) could be considered to be of foreign origin. To that extent, the population of Jovita has the highest incidence of non-native polymorphisms (adding up to 38%).

Despite finding that the AMOVA results and pairwise FST distances did not show any significant difference between the studied populations ($p > 0.05$), they were considered discrete populations for the principal component analysis, attending to the above mentioned different tendencies in allelic frequencies distribution. Results are shown in Figure 2. The first component, accounting for 48.5% of the total variation, separates the African sample of Ivory Coast from the other populations. All of the

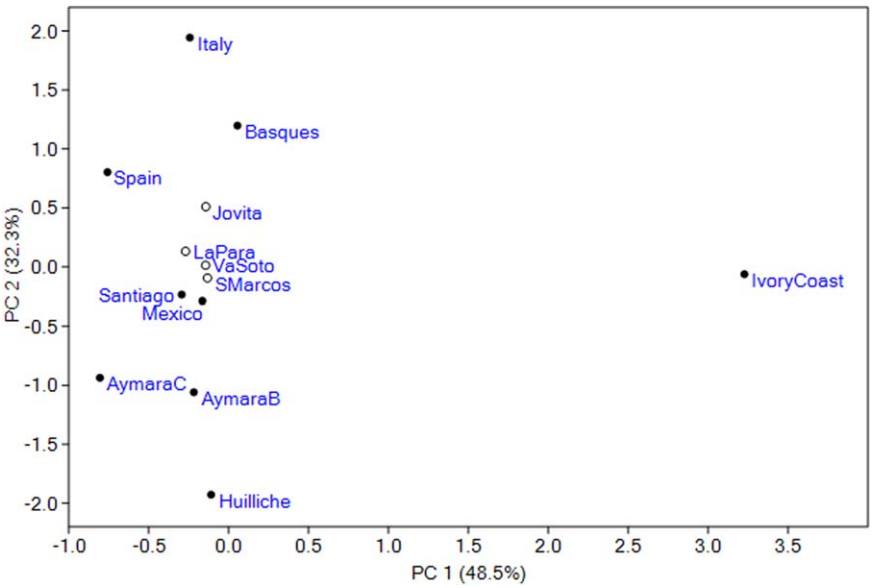

FIGURE 2 Principal component analysis based on the frequencies of O alleles in the studied populations (marked with empty circles) and population samples from Spain and Basques (Fregel, Maca-Meyer, Cabrera, González, & Larruga, 2005), Italy (Nishimukai et al., 2009), Akans of Ivory Coast and Aymara from Bolivia (Roubinet et al., 2001), Aymara and Huilliche from Chile (Llop, Henrıquez, Moraga, Castro, & Rothhammer, 2006), and admixed population from Santiago de Chile (Llop et al., 2006) and Mexico City (Estrada-Mena et al., 2009)

Córdoba samples cluster together at intermediate values of the second component, close to other admixed South American populations (Santiago de Chile and Mexico City). They also fall at intermediate distances between European and Native American populations, although Jovita appears closer to the European populations, as expected. It is interesting to note that this pattern, based on a single locus, is identical to that observed by García et al. (2015), who analyzed 10 autosomal AIMs. The results demonstrate that the analysis of the ABO system (more precisely the variants of the O blood group) constitutes a useful tool for the study of the genetic structure and evolutionary history of human populations, reflecting accurately the relative parental continental contributions to the gene pool of Latin American admixed populations.

ACKNOWLEDGMENTS

We thank the volunteers who donated samples. M.P.T. is a doctoral fellow, and A.G., D.A.D., and R.N. are Research Career Members of the Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET).

AUTHOR CONTRIBUTIONS

D.A.D. and R.N. supervised the study. M.P. and A.G. collected samples and extracted DNA. M.P.T. performed the genotyping. M.P.T. and D.A.D. analyzed the data. M.P.T., D.A.D., and R.N. wrote the manuscript with help from all co-authors.

REFERENCES

Avena, S. A., Parolin, M. L., Boquet, M., Dejean, C. B., Postillone, M. B., Alvarez, T. Y., … Carnese, F. R. (2010). Mezcla génica y linajes uniparentales en Esquel (Pcia. de Chubut). Su comparación con otras muestras poblacionales argentinas. *BAG—Journal of Basic and Applied Genetics*, [online] *21*(1), 1–14.

Estrada-Mena, B., Estrada, F. J., Ulloa-Arvizu, R., Guido, M., Méndez, R., Coral, R., … García-Carrancá, A. (2009). Blood Group O Alleles in Native Americans: Implications in the Peopling of the Americas. *American Journal of Physical Anthropology*, *142*, 85–94.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567.

Fregel, R., Maca-Meyer, N., Cabrera, V. M., González, A. M., & Larruga, J. M. (2005). Description of a simple multiplex PCR-SSCP method for AB0 genotyping and its application to the peopling of the Canary Islands. *Immunogenetics*, *57*(8), 572–578.

García, A., Demarchi, D. A., Tovo-Rodrigues, L., Pauro, M., Callegari-Jacques, S. M., Salzano, F. M., & Hutz, M. H. (2015). High interpopulation homogeneity in Central Argentina as assessed by Ancestry Informative Markers (AIMs). *Genetics and Molecular Biology*, *38*(3), 324–331.

Georges, L., Seidenberg, V., Hummel, S., & Fehren-Schmitz, L. (2012). Molecular characterization of ABO blood group frequencies in pre-Columbian Peruvian highlanders. *American Journal of Physical Anthropology*, *149*(2), 242–249.

Halverson, M. S., & Bolnick, D. A. (2008). An ancient DNA test of a founder effect in Native American ABO blood group frequencies. *American Journal of Physical Anthropology*, *137*, 342–347.

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*(1), 1–9.

Hosseini-Maaf, B., Irshaid, N. M., Hellberg, Å., Wagner, T., Levene, C., Hustinx, H., … Olsson, M. L. (2005). New and unusual O alleles at the ABO locus are implicated in unexpected blood group phenotypes. *Transfusion*, *45*(1), 70–81.

Hummel, S., Schmidt, D., Kahle, M., & Herrmann, B. (2002). ABO blood group genotyping of ancient DNA by PCR-RFLP. *International Journal of Legal Medicine*, *116*, 327–333.

Llop, E., Henrıquez, H., Moraga, M., Castro, M., & Rothhammer, F. (2006). Brief communication: molecular characterization of O alleles at the ABO locus in Chilean Aymara and Huilliche Indians. *American Journal of Physical Anthropology*, *131*, 535–538.

Morales, J. O., Dipierri, J. E., Alfaro, E., & Bejarano, I. F. (2000). Distribution of the ABO system in the argentine northwest: miscigenation and genetic diversity. *Interciencia*, *25*(9), 432–435.

Nishimukai, H., Fukumori, Y., Tsujimura, R., Okiura, T., Tanabe, R., Orimoto, C., & Ueda, N. (2009). Rare alleles of the ABO blood group system in two European populations. *Legal Medicine*, *11* (Suppl 1), S479–S481.

Ogasawara, K., Bannai, M., Saitou, N., Yabe, R., Nakata, K., Takenaka, M., … Tokunaga, K. (1996). Extensive polymorphism of ABO blood group gene: three major lineages of the alleles for the common ABO. *Human Genetics*, *97*(6), 777–783.

Pellegrino, A. (2002). *La migración internacional en América Latina. Tendencias y perfiles de los migrantes*. Presented at the Conferencia Hemisférica sobre Migración Internacional. CEPAL, Santiago, Chile.

Roubinet, F., Kermarrec, N., Despiau, S., Apoil, P. A., Dugoujon, J. M., & Blancher, A. (2001). Molecular polymorphism of O alleles in five populations of different ethnic origins. *Immunogenetics*, *53*(2), 95–104.

Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D. G., Mulligan, C. J., … Malhi, R. S. (2007). Beringian standstill and spread of Native American founders. *PLoS ONE*, *2*(9), e829. doi: 10.1371/journal.pone.0000829.

Vaccaro, M. S., Di Fabio Rocca, F., Russo, G., Parolin, M. L., Avena, S. A., Carnese, F. R., & Dejean, C. R. (2011). *Determinación molecular del sistema ABO en muestras cosmopolitas de la Argentina*. Presented at the X Jornadas Nacionales de Antropología Biológica, La Plata, Argentina.

Villanea, F. A., Bolnick, D. A., Monroe, C., Worl, R., Cambra, R., Leventhal, A., & Kemp, B. M. (2013). Brief communication: Evolution of a specific O allele (O1vG542A) supports unique ancestry of Native Americans. *American Journal of Physical Anthropology*, *151*(4), 649–657.

Yamamoto, F., Clausen, H., White, T., Marken, J., & Hakomori, S. (1990). Molecular genetic basis of the histo-blood group ABO system. *Nature*, *345*, 229–233.

Yip, S. P. (2002). Sequence variation at the human ABO locus. *Annals of Human Genetics*, *66*, 1–27.

How to cite this article: Tavella MP, García A, Pauro M, Demarchi DA, Nores R. Molecular polymorphisms of the ABO locus as informative markers of ancestry in Central Argentina. *Am J Hum Biol.* 2017;00:e22982. https://doi.org/10.1002/ajhb.22982