




ORIGINAL ARTICLE

Open Access



# Analysis of 27 Y-chromosomal STR loci of the Mestizo Peruvian population

Carlos David Neyra-Rivera<sup>1\*</sup> , Edgardo Delgado Ramos<sup>2,3</sup> , Ericson Gutierrez Ingunza<sup>4</sup>  and Bruce Budowle<sup>5</sup> 

## Abstract

**Background** Y-chromosome-specific short tandem repeat markers reside on the non-recombinant portion of the Y chromosome and are paternally inherited. These properties make Y-STRs a useful tool in investigations of sexual assault and other violent crimes, paternity cases, genealogical testing, and evolutionary studies. Native populations reside in the Peruvian coast, mountains, and jungle; however, the Mestizos are the largest population group in Peru.

**Results** Of the 290 samples, 283 distinct haplotypes were observed. Gene diversities for each Y-STR marker ranged from 0.4271 to 0.8597. The haplotype diversity for the total sample was 0.9964.

**Conclusions** Genetic affinity was identified between the Peruvian Mestizo population and other Peruvian populations (except Jivaro, Awajun, and Shipibo-Conibo populations) and Bolivian Mestizo populations using multidimensional scaling and phylogenetic tree analysis

**Keywords** Y-STRs, Haplotypes, Mestizo Peruvian population

## Background

Y-STR markers are inherited via the paternal lineage, present haploid, and except for the pseudoautosomal regions do not experience recombination. These properties make Y-STRs a useful tool in sexual assault investigations, paternity testing, genealogical testing, and evolutionary studies (Roewer 2009).

The populations of Peru have undergone different migratory processes, contributing to Peruvian population variation and native populations in the coast, mountains, and jungle as well as populations with

substantial genetic admixture. Because of this diversity, studies have been carried out on autosomal STRs markers from different Peruvian populations (Delgado & Neyra 2018; Neyra et al. 2019, 2020; Neyra-Rivera et al. 2020; Neyra-Rivera et al. 2021a, b; Bermejo et al. 2022; Neyra-Rivera et al., 2022) and also on Y-STRs from native populations (Tineo et al. 2015; Neyra-Rivera et al., 2022; Neyra-Rivera et al. 2023) analysis in the case of Y-STRs the genetic diversity (GD, probability that two randomly chosen individuals can be differentiated by a particular genetic marker(s) (Nei 1987)), haplotype diversity (HD, probability that two randomly chosen individuals can be differentiated by a particular set of markers comprising a haplotype(s) (Nei 1987)), probability of a random match (MP, probability that two randomly selected individuals in a population share identical genotypes (Jones 1972)), discrimination capacity (DC, probability that two randomly selected individuals will be genetically differentiated at a particular marker(s) (Smalldon and Moffat 1973)), haploid match probability (HMP, probability that two randomly selected individuals will share the same haplotype), and

\*Correspondence:

Carlos David Neyra-Rivera  
carlosdavidmp@outlook.es

<sup>1</sup> Universidad Privada Peruano Alemana, Lima, Perú

<sup>2</sup> Escuela de Pos Grado, Universidad Privada Norbert Wiener, Lima, Perú

<sup>3</sup> Unidad de Biología Molecular y de Genética (UNBIMOG), Instituto de Medicina Legal y Ciencias Forenses del Ministerio Público, Lima, Perú

<sup>4</sup> Unidad de Investigación Para La Generación y Síntesis de Evidencias en Salud, Universidad San Ignacio de Loyola, Lima, Perú

<sup>5</sup> Department of Forensic Medicine, University of Helsinki, Helsinki, Finland



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

pairwise genetic distance (RST, the correlation of allele sizes (rather than allelic states) between genetic markers sampled within populations (Excoffier 2001)). However, there have not been studies with large population samples (i.e., greater than 200 individuals) and with at most 23 Y-STR markers. Only one study with 83 individuals and 23 Y-STR markers (accession number YA005599) is represented in the YHRD database (Yhrd.org). The present investigation seeks to expand on available genetic data by studying a larger sample set of Mestizos for 27 Y-STR markers and compare these genetic data to assess relationships with populations associated either geographically or historically.

## Methods

### Samples and experiment

DNA from 290 unrelated individuals were collected in Nucleid Cards (Copan). The individuals reside in different departments of Peru (Amazonas  $n=2$ , Ancash  $n=15$ , Apurimac  $n=4$ , Arequipa  $n=15$ , Ayacucho

$n=15$ , Cajamarca  $n=2$ , Cusco  $n=16$ , Huancavelica  $n=4$ , Huanuco  $n=22$ , Ica  $n=4$ , Junín  $n=9$ , La Libertad  $n=2$ , Lambayeque  $n=1$ , Lima  $n=123$ , Loreto  $n=4$ , Madre de Dios  $n=7$ , Moquegua  $n=1$ , Piura  $n=5$ , Puno  $n=31$ , Tacna  $n=2$ , Tumbes  $n=2$ , and Ucayali  $n=4$ ) (Fig. 1). All participants freely and voluntarily signed informed consent to participate in this study.

A 1.2 mm punch was taken from each card and amplified by direct PCR using the Yfiler™ Plus PCR Amplification Kit (Life Technologies) following the manufacturer's instructions. Amplified products were detected in the Applied Biosystem™ 3500XL Genetic Analyzer (Life Technologies) following the manufacturer's recommended protocol. After capillary electrophoresis, the data were imported into the GeneMapper® ID-X v1.5 software (Thermo Fisher Scientific 2015) to generate the genetic profiles which were analyzed according to the manufacturer's recommendations. The population data were deposited in the YHRD database (accession number: YA006020).



**Fig. 1** Departments of Peru from which the 290 samples were collected. (a) Amazonas  $n=2$ , (b) Ancash  $n=15$ , (c) Apurimac  $n=4$ , (d) Arequipa  $n=15$ , (e) Ayacucho  $n=15$ , (f) Cajamarca  $n=2$ , (g) Cusco  $n=16$ , (h) Huancavelica  $n=4$ , (i) Huanuco  $n=22$ , (j) Ica  $n=4$ , (k) Junín  $n=9$ , (l) La Libertad  $n=2$ , (m) Lambayeque  $n=1$ , (n) Lima  $n=123$ , (o) Loreto  $n=4$ , (p) Madre de Dios  $n=7$ , (q) Moquegua  $n=1$ , (r) Piura  $n=5$ , (s) Puno  $n=31$ , (t) Tacna  $n=2$ , (u) Tumbes  $n=2$ , and (v) Ucayali  $n=4$

**Statistical analysis**

For the calculation of the allelic frequencies for 23 of the 27 STRs, the Power Stats V1.2 software was used (Ter-eba 1919); for the DYS385a/b and DYF387S1 markers, the “type” frequencies were calculated by the counting method. GD and HD were calculated with the following formulas  $GD=(n/(n-1)) (1-\Sigma Pi^2)$  and  $HD=(n/(n-1)) (1-\Sigma Xi^2)$ , where  $Pi$  is the frequency of the  $i$ th allele,  $n$  is the number of samples analyzed, and  $Xi$  represents the calculated frequency of the  $i$ th haplotype (Sandoval et al., 2013). The MP was calculated as  $MP=\Sigma Xi^2$  and DC was determined by dividing the number of different observed haplotypes with the sum total number of identified haplotypes. HMP was calculated with the formula  $HMP=1-HD$ . AMOVA was used to RST values between Y-STR datasets. The RST values were then used for MDS analysis. RST and p values between different populations were estimated by AMOVA and visualized on a MDS plot (YHRD) (Willuweit & Roewer 2007).

**Results**

**Allelic frequencies and forensic statistical parameter analyses**

This study analyzed 27 Y STRs in 290 males collected across different departments in Peru. Results were obtained for all 27 Y-STR loci in all samples, and no null alleles were observed. Allele (or type) frequencies for each locus of the combined data set are shown in Table 1. The DYS437 marker displayed the fewest number of alleles (3 alleles), while the DYS385a/b marker had the highest number of types (45 types). GD values ranged between 0.4271 (DYS437) and 0.9415 (DYS385a/b), and only two Y-STR loci had values less than 0.5 (DYS437

and DYS391). The two highest diversity markers, based on GD, were DYS385a/b (0.9415) and DYF387S1 (0.9363) while the lowest diversity was observed at the DYS437 locus (0.4271).

**Haplotype diversity**

Since there is no recombination between the Y STRs, haplotype data better display the diversity within the population. There were 283 distinct haplotypes in the total data set ( $n=290$ ) with HD and DC values of 0.9964 and 0.9759, respectively (Table 2). The largest number of the same haplotype was two and seven pairs of two haplotypes were observed (H78 –Puno and Cajamarca-, H105 –Huancavelica and Cusco-, H124 –Lima-, H201 –Lima-, H210 –Lima-, H253 –Madre de Dios and Lima- and H255 –Lima and Cusco-). Ten microvariants were identified and were observed in the markers DYS627 (alleles 18.2, 20.2, and 21.2), DYS448 (alleles 16.2, 17.2, and 18.2), and DYS448 (allele 18.2).

**Population structure for forensic analyzes**

Using an online tool NevGen Y DNA Predictor (Cetkovic Gentula & Nevski, 2015) provisional haplogroups (most probable and major clade only) for the distinct haplotypes of the Mestizos from Peru were determined. In the Mestizo population, the majority (56.18%) belonged to haplogroup Q, with the rest distributed among haplogroups R (22.61%), J (7.07%), E (6.36%), G (2.83%), T (2.47%), I (2.12%), and L (0.35%) (Table 2).

**Genetic distances**

The RST values for genetic distances and the MDS plot are shown in Table 3 and Fig. 2 for the

**Table 1** Allele frequencies and gene diversities of 27 Y-chromosome STR loci for the Mestizo population from Peru ( $n=290$ )

Allele	DYS375	DYS381	DYS448	DYS385b	DYS449	DYS385a	DYS391	YCAT444	DYS448	DYS391	DYS448	DYS449	DYS392	DYS381	DYS375	DYS447	DYS449	DYS381	DYS449	Allele	No.	DYS385a/b	Allele	No.	DYF387S1					
7																				10.13	1	0.0014	33.36	1	0.0024					
8										0.0483			0.0724							10.15	1	0.0014	33.36	2	0.0069					
10										0.7000			0.1552							11.11	1	0.0010	33.40	2	0.0069					
11										0.2041			0.2100							11.12	4	0.0136	34.28	1	0.0054					
12	0.0034	0.0739								0.6069			0.2414	0.0101						11.13	5	0.0172	34.35	2	0.0069					
13										0.0034			0.5739	0.0724						11.13	1	0.0014	34.36	5	0.0172					
14	0.0034	0.3345				0.0034				0.0172			0.2691							11.14	45	0.1552	34.37	3	0.0103					
15	0.0034	0.0034								0.1148			0.6066							11.15	5	0.0172	34.38	7	0.0241					
16	0.0000									0.3172			0.0172							11.16	2	0.0069	34.39	10	0.0161					
16.2										0.0069			0.0172							12.12	2	0.0069	34.41	1	0.0034					
17	0.1759									0.2931			0.0103							12.13	1	0.0014	34.42	1	0.0034					
17.2										0.0103										12.14	1	0.0014	34.39	7	0.0241					
18	0.3207									0.1517										12.15	2	0.0069	34.36	30	0.1310					
18.2										0.0609			0.0034							12.16	2	0.0069	34.37	11	0.0379					
19	0.2552									0.1481			0.3176							12.17	1	0.0014	34.38	25	0.0862					
20	0.0724									0.2307			0.4793							12.18	2	0.0069	34.39	43	0.1483					
20.2										0.0034										12.19	1	0.0014	34.40	10	0.0345					
21	0.0310									0.1448			0.2172							13.14	7	0.0241	34.41	4	0.0138					
21.2										0.0034										13.15	7	0.0241	34.42	2	0.0069					
22	0.0034									0.5069			0.1276							13.16	9	0.0310	34.36	11	0.0379					
23										0.2655										13.17	2	0.0069	34.37	19	0.0655					
24	0.0414									0.0241										13.18	3	0.0103	34.28	12	0.0414					
25										0.0138										13.19	2	0.0069	34.39	9	0.0310					
26										0.0103										13.20	1	0.0014	34.40	3	0.0103					
27										0.0103										14.14	12	0.0414	34.41	2	0.0069					
28										0.0552										14.15	10	0.0345	34.37	4	0.0138					
29										0.2100										14.16	8	0.0276	34.38	11	0.0379					
30										0.3448										14.17	31	0.1069	34.39	10	0.0345					
31										0.2828										14.18	21	0.0724	34.40	2	0.007					
32										0.0793										14.19	5	0.017	34.38	13	0.0463					
33										0.0069										14.20	3	0.010	34.39	7	0.024					
34										0.0207										15.15	4	0.014	34.40	5	0.017					
35										0.0966										15.16	14	0.046	34.41	1	0.003					
36										0.0069										15.17	12	0.041	34.39	3	0.010					
37										0.1586										15.18	22	0.076	34.40	2	0.007					
38										0.1586										15.19	14	0.046	34.41	1	0.003					
39										0.1862										16.20	7	0.024								
40										0.1069										16.16	4	0.014								
41										0.0483										16.17	5	0.017								
42										0.0103										16.18	4	0.014								
43										0.0138										16.20	1	0.003								
44																				17.17	5	0.017								
45																				17.18	1	0.003								
																				17.19	2	0.007								
IK.H																					1	0.003								
Na	11	5	6	6	14	5	11	6	5	7	4	6	8	5	4	6	8	5	4	10	11	3	13	5	7	11	5			
GD	0.7864	0.9466	0.6512	0.7969	0.8201	0.5412	0.7721	0.5787	0.5632	0.6422	0.4554	0.5915	0.6673	0.6430	0.7462	0.7919	0.7213	0.4271	0.8597	0.5160	0.7083	0.8291	0.5614							
																					45	36								
																					0.9415									

NA Numbers of allele, GD Genetic diversity



**Table 3** Pairwise Rst value estimates (below the diagonal) and corresponding *p* value (above the diagonal) for the Mestizo population from Peru and 14 reference population datasets

Population	Peru [Mestizo] (n=283)	Bolivia [Boliviano] (n=218) <sup>a</sup>	Bogotá, Colombia [Mestizo] (n=184) <sup>b</sup>	Oxapampa, Peru [Ashaninka] (n=57) <sup>c</sup>	Amazonas, Peru [Awajun] (n=18) <sup>d</sup>	Puno, Peru [Aymara] (n=90) <sup>e</sup>	Amazonas, Peru [Chachapoya] (n=106) <sup>f</sup>	Cuzco, Peru [Chumbivilca] (n=10) <sup>g</sup>	Arequipa, Peru [Chuquiabamba] (n=16) <sup>h</sup>	Junin, Peru [Huanca] (n=13) <sup>g</sup>	Loreto, Peru [Iquito] (n=85) <sup>d</sup>	Amazonas, Peru [Jivaro] (n=21) <sup>f</sup>	Lima, Peru [Quechua] (n=11) <sup>f</sup>	Ucayali, Peru [Shipibo-Conibo] (n=21) <sup>g</sup>	Spain [Spanish] (n=2406) <sup>f</sup>
Peru [Mestizo] (n=283)	-	0.1771	0.0000	0.0017	0.0046	0.0000	0.0285	0.0000	0.0005	0.0014	0.0046	0.0551	0.2577	0.0013	0.0000
Bolivia [Boliviano] (n=218) <sup>a</sup>	0.0017	-	0.0000	0.0017	0.0016	0.0000	0.0101	0.0009	0.0006	0.0017	0.0010	0.0257	0.1251	0.0023	0.0000
Bogotá, Colombia [Mestizo] (n=184) <sup>b</sup>	0.0889	0.0891	-	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0355	0.0000	0.0000
Oxapampa, Peru [Ashaninka] (n=57) <sup>c</sup>	0.0359	0.0374	0.2072	-	0.0004	0.0005	0.0000	0.0002	0.0010	0.0003	0.0020	0.0068	0.0013	0.0000	0.0000
Amazonas, Peru [Awajun] (n=18) <sup>d</sup>	0.0776	0.1004	0.2535	0.1055	-	0.0007	0.0000	0.0020	0.0026	0.0078	0.0036	0.0971	0.0039	0.0002	0.0000
Puno, Peru [Aymara] (n=90) <sup>e</sup>	0.1006	0.0982	0.2932	0.0423	0.0981	-	0.0000	0.0747	0.3081	0.1221	0.0000	0.0008	0.0004	0.0000	0.0000
Amazonas, Peru [Chachapoya] (n=106) <sup>f</sup>	0.0098	0.0163	0.0587	0.1015	0.1542	0.1606	-	0.0000	0.0000	0.0000	0.0000	0.0014	0.3505	0.0000	0.0000
Cuzco, Peru [Chumbivilca] (n=10) <sup>g</sup>	0.1976	0.1952	0.3963	0.1917	0.2377	0.0451	0.2871	-	0.8148	0.2576	0.0001	0.0004	0.0020	0.0000	0.0000
Arequipa, Peru [Chuquiabamba] (n=16) <sup>h</sup>	0.1425	0.1424	0.3464	0.1114	0.1644	0.0050	0.2189	-0.0415	-	0.7902	0.0001	0.0007	0.0019	0.0000	0.0000
Junin, Peru [Huanca] (n=13) <sup>g</sup>	0.1273	0.1342	0.3330	0.1282	0.1404	0.0257	0.1975	0.0225	-0.0343	-	0.0005	0.0016	0.0043	0.0000	0.0000
Loreto, Peru [Iquito] (n=85) <sup>d</sup>	0.0190	0.0306	0.1769	0.0294	0.0683	0.0727	0.0562	0.2143	0.1325	0.1297	-	0.0852	0.0099	0.0000	0.0000
Amazonas, Peru [Jivaro] (n=21) <sup>f</sup>	0.0274	0.0416	0.1795	0.0537	0.0550	0.0895	0.0863	0.2601	0.1652	0.1725	0.0212	-	0.0323	0.0010	0.0000

**Table 3** (continued)

Population	Peru [Mestizo] (n=283)	Bolivia [Boliviano] (n=218) <sup>a</sup>	Bogotá, Colombia [Mestizo] (n=184) <sup>b</sup>	Oxapampa, Peru [Ashaninka] (n=57) <sup>c</sup>	Amazonas, Peru [Awajun] (n=18) <sup>d</sup>	Puno, Peru [Aymara] (n=90) <sup>e</sup>	Amazonas, Peru [Chachapoya] (n=106) <sup>f</sup>	Cuzco, Peru [Chumbivilca] (n=10) <sup>g</sup>	Arequipa, Peru [Chuquiabamba] (n=16) <sup>h</sup>	Junin, Peru [Huancá] (n=13) <sup>g</sup>	Loreto, Peru [Iquito] (n=85) <sup>d</sup>	Amazonas, Peru [Jiwaro] (n=21) <sup>f</sup>	Lima, Peru [Quechua] (n=11) <sup>f</sup>	Ucayali, Peru [Shipibo-Conibo] (n=21) <sup>g</sup>	Spain [Spanish] (n=2406) <sup>f</sup>
Lima, Peru [Quechua] (n=11) <sup>f</sup>	0.0104	0.0316	0.0680	0.1548	0.1870	0.1942	0.0033	0.3275	0.2450	0.2043	0.0841	0.0905	-	0.0001	0.0000
Ucayali, Peru [Shipibo-Conibo] (n=21) <sup>g</sup>	0.0830	0.0837	0.1873	0.1621	0.1900	0.2080	0.1347	0.3928	0.3082	0.3215	0.1011	0.1260	0.2086	-	0.0000
Spain [Spanish] (n=2406) <sup>f</sup>	0.2377	0.2300	0.0486	0.3685	0.4239	0.4559	0.2005	0.5589	0.5178	0.5058	0.3423	0.3547	0.2351	0.3350	-

<sup>a</sup> María Brion, Antonio Salas, Francesca Brisighelli, Jorge Mario Cárdenas, Alejandro Blanco Vereá, Sandra Belezá, Manuel Fondévila, Juan Carlos Jaime, María Victoria Lareu, José Lovo-Gómez, Angel Carracedo contributed 223 Haplotypes on October 4, 2012 Filed under accession number YA003761

<sup>b</sup> Yunis JJ, Acevedo LE, Campo DS, and Yunis EJ. (2005), "Population data of Y-STR minimal haplotypes in a sample of Caucasian-Mestizo and African descent individuals of Colombia," *Forensic Sci Int* 151(2-3):307-13

<sup>c</sup> Tineo DH, Lolola S, Paredes FV, Amaya YC, Simão F, Carvalho EF, Gusmão L. Genetic characterization of 27 Y-STR loci in the native population of Ashaninka from Peru. *Forensic Sci. Int. Genet. Suppl. Ser.* 2015. 5: e220-e222. DOI: doi.org/10.1016/j.fsigss.2015.09.088

<sup>d</sup> Neyra-Rivera CD, Robles CS, Delgado E., Velásquez MRE., Budowle B. Analysis of 27 STR markers of the chromosome Y in the population of the jungle from Peru. *Revista Española de Medicina Legal.* 2023., <https://doi.org/10.1016/j.reml.2022.10.001>

<sup>e</sup> Neyra-Rivera CD., Arenas T., Ramos D., Ramos D., Reinoso V., Rey C. and Budowle B. (2021), "Population data of 27 Y-chromosome STRs in Aymara population from Peru," *Australian Journal of Forensic Sciences* 54(5):596-610

<sup>f</sup> Evelyn Guevara, Antti Sajantila contributed 24 Haplotypes on June 19, 2015 Filed under accession number YA004109

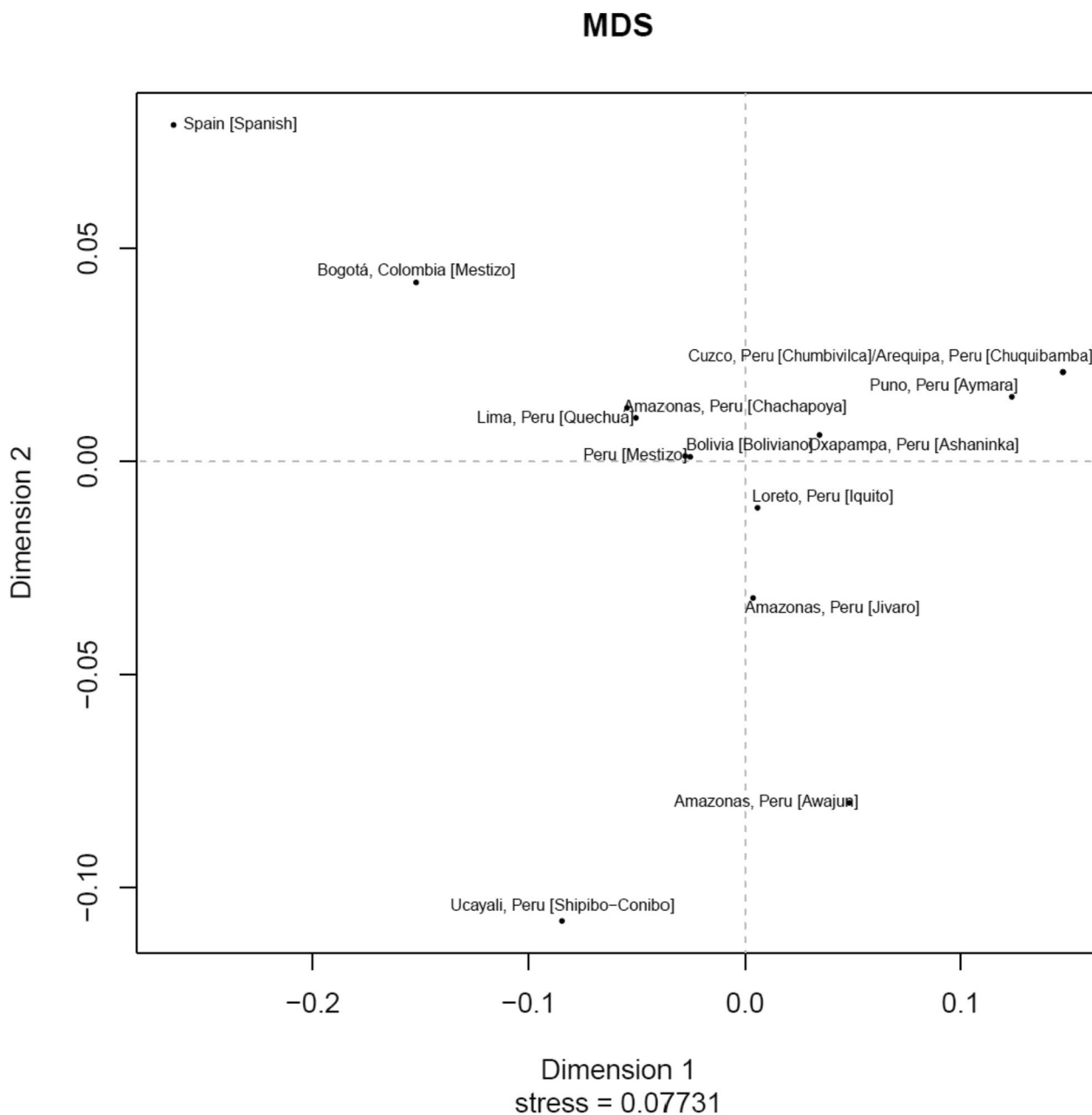
<sup>g</sup> Roewer Lutz et al. Continent-wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans. *PLoS genetics* vol. 9.4 (2013): e1003460. <https://doi.org/10.1371/journal.pgen.1003460>

<sup>h</sup> Claudia Fiorella Barletta Carrillo contributed 16 Haplotypes on March 13, 2013 (for reference see Roewer L., Nothmager M., Gomes V., González M., Corach D., Sala A., Alechine E., Palha T., Santos N., Ribeiro-Dos-Santos A., Geppert M., Willuweit S., Nagy M., Zweynert S., Baeta M., Núñez C., Martínez-Jarreta B., González-Andrade F., De F., Silva D., Builes J.J., Turbón D., Parra L., Arroyo-Pardo E., Toscanini U., Borjas L., Barletta C., Ewart E., Santos S. and Krawczak M. (2013), "Continent-wide decoupling of Y-chromosomal genetic variation from language and geography in native South Americans," *PLoS Genet* 9(4):e1003460

<sup>i</sup> Rukesh Shrestha, Chuan-Chao Wang contributed 11 Haplotypes on May 22, 2014 Filed under accession number YA003949

<sup>j</sup> Martínez-Cadenas C., Blanco-Verea A., Hernando B., George BJ., Brion M., Carracedo A., Salas A. and Capelli C. (2016), "The relationship between surname frequency and Y chromosome variation in Spain," *European Journal of Human Genetics* 24:120-8

<sup>k</sup> Martínez-Cruz B., Mendizabal I., Harmant C., Pablo D., Ioana M., Angelicheva D., Kouvatzi A., Makukh H., Netea MG., Pamjav H., Zalián A., Tournev I., Marushakova E., Popov V., Bertranpetit J., Kalaydjieva L., Quintana-Murci L., Comas D. and Genographic-Consortium (2016), "Origins, admixture and founder lineages in European Roma," *Eur J Hum Genet* 24(6):937-43



**Fig. 2** Multidimensional scaling (MDS) analysis for Mestizo population ( $n = 290$ ) and 14 other sample populations. Bolivia [Boliviano] ( $n = 218$ )<sup>a</sup>; Bogotá, Colombia [Mestizo] ( $n = 184$ )<sup>b</sup>; Oxapampa, Peru [Ashaninka] ( $n = 57$ )<sup>c</sup>; Amazonas, Peru [Awajun] ( $n = 18$ )<sup>d</sup>; Puno, Peru [Aymara] ( $n = 90$ )<sup>e</sup>; Amazonas, Peru [Chachapoya] ( $n = 106$ )<sup>f</sup>; Cuzco, Peru [Chumbivilca] ( $n = 10$ )<sup>g</sup>; Arequipa, Peru [Chuquibamba] ( $n = 16$ )<sup>h</sup>; Junín, Peru [Huanca] ( $n = 13$ )<sup>i</sup>; Loreto, Peru [Iquito] ( $n = 85$ )<sup>j</sup>; Amazonas, Peru [Jivaro] ( $n = 21$ )<sup>k</sup>; Lima, Peru [Quechua] ( $n = 11$ )<sup>l</sup>; Ucayali, Peru [Shipibo-Conibo] ( $n = 21$ )<sup>m</sup>; and Spain [Spanish] ( $n = 2406$ )<sup>n</sup> deposited in the YHRD. Filed under accession numbers: a: YA005867, b: YP000047, c: YA005583, d: YA005891, e: YP000861, f: YA004108, g: YA003820, h: YA003821, i: YA003822, j: YP000860, k: YA005058, l: YA003949, m: YA003825, n: YP001017

Mestizo population and other datasets from Boliviano (Bolivia), Mestizo (Bogotá, Colombia), Aymara (Puno, Peru), Ashaninka (Oxapampa, Peru), Awajun (Amazonas, Peru), Chachapoya (Amazonas, Peru), Chumbivilca (Cuzco, Peru), Chuquibamba (Arequipa, Peru),

Huanca (Junín, Peru), Iquito (Loreto, Peru), Jivaro (Amazonas, Peru), Quechua (Lima, Peru), Shipibo-Conibo (Ucayali, Peru), and Spanish (Spain). There were 17 Y-STRs in common among the data sets (DYS456, DYS389I, DYS390, DYS389II, DYS458,



DYS19, *DYS385*, *DYS393*, *DYS391*, *DYS439*, *DYS635*, *DYS392*, *YGATAH4*, *DYS437*, *DYS438*, *DYS448*).

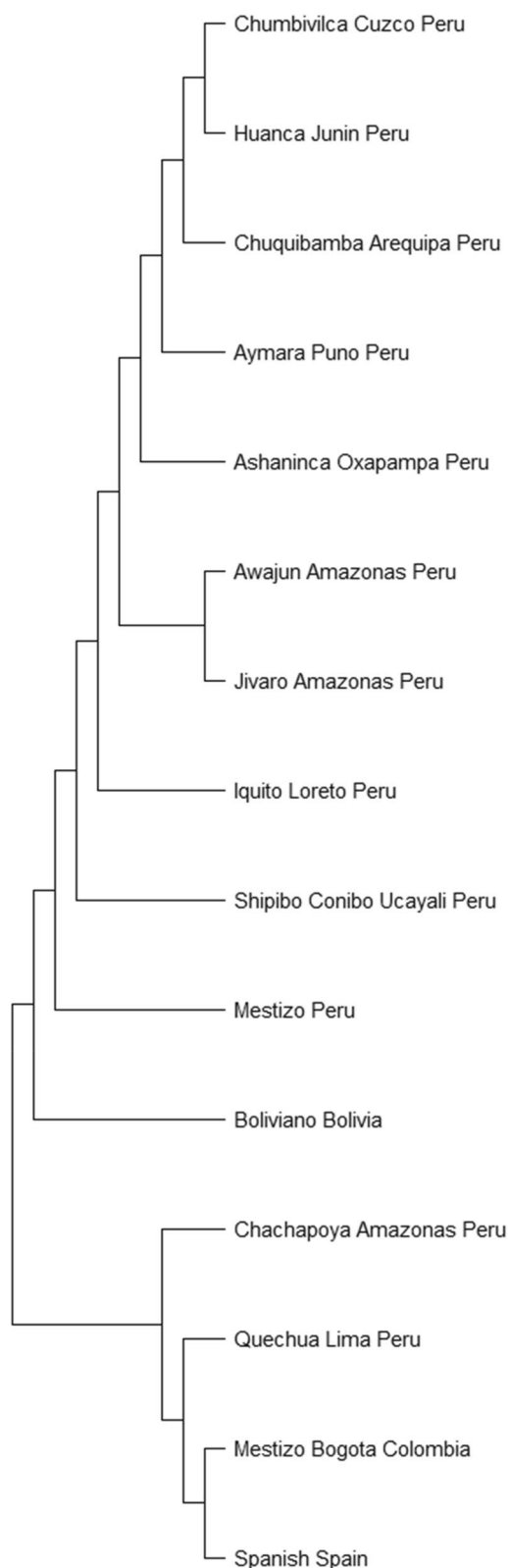
**Discussion**

Molecular biology technologies offer enhanced capabilities for human identification. Forensic DNA typing has an important role in some criminal investigations, which in turn benefits society, by providing critical data for developing investigative leads to identify donors of biological evidence who may be perpetrators of crime as well as to exonerate those that may be wrongly associated with biological evidence. In order to contribute to assessing the significance of an association by DNA between a reference sample and an evidence sample, a population study on a sample population of Peruvians was carried out by analyzing 27 Y-STRs contained within a widely used commercial kit (i.e., Yfiler™ Plus PCR Amplification Kit).

The per locus GD of the studied population ranged between 0.4271 (*DYS437*) and 0.9415 (*DYS385a/b*). In contrast, the GD values for the Peruvian Aymara population were between 0.2424 (*DYS437*) and 0.9298 (*DYS385a/b*), Peruvian Ashaninca populations were between 0.1337 (*DYS437*) and 0.6582 (*DYS438*) and the population of the Peruvian Jungle were between 0.000 (*DYS635*, *DYS448* and *DYS437*) and 0.7094 (*DYS449*) (Tineo et al. 2015; Neyra-Rivera et al., 2022; Neyra-Rivera et al. 2023). The values have a similar distribution with those of the Mestizo population of Spain, *DYS393* (0.4625) had the lowest GD (Martin et al., 2004).

Because there is no recombination on the Y chromosome where the Y-STRs reside, haplotype diversity is more meaningful than GD to compare diversity among populations. The Y-STR HD was 0.9964 with 283 distinct haplotypes out of 290 individuals analyzed in the Peruvian population studied herein. This level of diversity is slightly lower than that of the Peruvian Aymara (0.9998, *n*=90) and Peruvian Asháninca (1.000) populations and higher than that of the Peruvian Jungle population (0.9910, *n*=141) (Tineo et al. 2015; Neyra-Rivera et al., 2022; Neyra-Rivera et al. 2023). In addition, the diversity value was lower than that of the Mestizo populations of Spain (0.9995, *n*=144), Colombia (0.9971, *n*=137), and Bolivia (1.000, *n*=100) (Martin et al., 2004; Vullo et al. 2015; Yunis et al. 2005). The difference compared with the Mestizo populations may be due in part to fewer samples being typed, a fewer number of markers tested (17 vs 27 markers), and/or stochastic variation.

The specific haplogroups carried by the males analyzed in the Peruvian Mestizo population provide insight into the male lineages of the population. The most prevalent was haplogroup Q (56.18%), which also was the most common haplogroup in the Ashaninca (between 99 and 100%), Aymara (86.52%), and La Selva (91.56%)



**Fig. 3** UPGMA tree based on 17 Y-STR markers in common among 15 populations and Rst estimates



populations. The lower proportion of haplogroup Q in the Mestizo population may be due to admixture that has occurred and continues to occur in Peru; for various reasons, native populations have migrated from their communities to the cities in Peru (Tineo et al. 2015; Neyra-Rivera et al., 2022; Neyra-Rivera et al. 2023).

The RST values were used to generate a UPGMA tree using the Molecular Evolutionary Genetics Analysis (MEGA) 6.06 software (Tamura et al. 2013). The Mestizo population was closest to Bolivian and other Peruvian populations (except Chachapoya and Quechua populations) (Fig. 3). The Mestizo population was closest to the Bolivian population (Bolivia, Boliviano) and farthest from the Jivaro and Awajun (Amazonas, Peru) and Shipibo-Conibo (Ucayali) (Fig. 2). These results suggest that there is a high genetic affinity between the Mestizo population and the other Peruvian populations (except Jivaro, Awajun and Shipibo-Conibo populations). Additionally, the Jivaro, Awajun, and Shipibo-Conibo populations were distant from all groups analyzed which may be due to genetic differences and/or small population sample sizes ( $n \leq 21$ ) (Fig. 3).

## Conclusions

In conclusion, 27 Y-STR loci in 290 unrelated Mestizo males were typed, providing a larger dataset for this population group. The data obtained have been deposited in the YHRD database (accession number: YA006020). These data are larger than any other available data for the Mestizo population from Peru. For example, YHRD (Yhrd.org) contains 394 haplotypes (394 with 12 Y-STR haplotypes, 229 with 17 Y-STR haplotypes, 83 with 23 Y-STR haplotypes, and until the study herein 0 with 27 Y-STR haplotypes) (filed under accession numbers YA003399 and YA005599). A specific limitation of this study is the small sample sizes for some groups which can lead to errors in the allele frequency estimates and subsequent population genetic calculations. Regardless of this limitation, the analyses carried out herein show that the Y-STR haplotypes are highly polymorphic and thus can be useful in human identification forensic cases, with proper statistical methods that account for uncertainty, as well as may have value for studies on the history of the human populations in Latin America.

## Abbreviations

AMOVA	Analysis of molecular variance
MDS	Multidimensional scaling plot
PCR	Polymerase chain reaction
YHRD	Y Haplotype Reference Database
Y-STR	Y-chromosome-specific short tandem repeat

## Acknowledgements

The authors thank the volunteers for donating their samples for the research. Also, to the Universidad Privada Peruano Alemana for financing this research.

## Authors' contributions

Carlos Neyra, Edgardo Delgado, and Ericson Gutierrez participated in the design of the study. Carlos Neyra and Edgardo Delgado processed the samples and obtained the genetic profiles. Carlos Neyra, Edgardo Delgado, Ericson Gutierrez, and Bruce Budowle carried out the statistical analysis. All participated in the writing and approval of the final version of the article.

## Funding

We thank the Universidad Privada Peruano Alemana for financing the development of the research (RECTORAL RESOLUTION No. 013–2022-REC).

## Availability of data and materials

The datasets analyzed during the current study are available from the corresponding author.

## Declarations

### Ethics approval and consent to participate

The study was approved by the ethics committee of Instituto de Medicina Tropical "Daniel Alcides Carrion" from San Marcos National University (Certificate of Approval CIEI-2018–015).

### Consent for publication

Consent has been given by all authors for publication of this study.

### Competing interests

The authors declare that they have no competing interests.

Received: 21 April 2023 Accepted: 24 September 2023

Published online: 16 October 2023

## References

- Bermejo ME, Ge J, Budowle B, Infante C, Neyra-Rivera CD (2022) Genetic study with 21 autosomal STRs in five Peruvian macro regions for human identification purposes. *Legal Med* 57:1344–6223. <https://doi.org/10.1016/j.legalmed.2022.102073>
- Cetkovic M, Nevski A (2015). NevGen Y-DNA haplogroup predictor. In: <http://www.nevgen.org/>. Accessed 20 Feb 2023
- Delgado E, Neyra CD (2018) Allele frequencies of 21 autosomal STR markers in a mixed race Peruvian population applied to forensic practice. *Rev Espanola De Medicina Leg* 45(3):92–97. <https://doi.org/10.1016/j.reml.2018.09.002>
- Excoffier L (2001) Analysis of population subdivision. In: Balding DJ, Bishop M, Cannings C (eds) *Statistical Genetics*. John Wiley & Sons, Chichester, pp 271–324
- Jones DA (1972) Blood lood samples- probability of discrimination. *J Forensic Sci Soc* 12(2):355–359. [https://doi.org/10.1016/s0015-7368\(72\)70695-7](https://doi.org/10.1016/s0015-7368(72)70695-7)
- Martín P, García-Hirschfeld J, García O, Gusmão L, García P, Albarrán C, Sancho M, Alonso A (2004) A Spanish population study of 17 Y-chromosome STR loci. *Forensic Sci Int* 139(2–3):231–235. <https://doi.org/10.1016/j.forsciint.2003.11.012>
- Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York, pp 39–46
- Neyra CD, Delgado E, Robles C, Velasquez M, Caceres O, Budowle B (2020) Population genetic study of a Peruvian population using human identification STRs. *Int J Legal Med* 134(6):2071–2073. <https://doi.org/10.1007/s00414-020-02418-6>
- Neyra CD, Suarez M, Cueva E, Bailón H, Gutierrez E (2019) Identificación genética de recién nacidos en Perú: un estudio piloto. *Rev Chil Pediatr* 90(1):26–35. <https://doi.org/10.32641/rchped.v90i1.730>
- Neyra-Rivera CD, Budowle B (2022) Genetic variation of different Peruvian populations using 23 autosomal STR markers. *Forensic Sci Int Genet Suppl Ser* 8:259–262. <https://doi.org/10.1016/j.fsigss.2022.10.055>
- Neyra-Rivera CD, Delgado E, Diaz F, Quispe JS, Ge J, Budowle B (2021a) Genetic study with autosomal STR markers in people of the Peruvian jungle for human identification purposes. *Can Soc Forensic Sci* 54(3):117–138. <https://doi.org/10.1080/00085030.2021.1933811>
- Neyra-Rivera CD, Robles CS, Delgado E, Velasquez MRE, Budowle B (2023) Analysis of 27 STR markers of the chromosome Y in the population of the jungle from Peru. *Rev. Espanola de Medicina Leg* ISSN 0377–4732. <https://doi.org/10.1016/j.reml.2022.10.001>

- Neyra-Rivera CD, Ticona A, Delgado E, Velasquez MRE, Budowle B (2020) Allelic frequencies with 23 autosomic STRS in the Aymara population of Peru. *Int J Legal Med* 135(3):779–781. <https://doi.org/10.1007/s00414-020-02448-0>
- Neyra-Rivera CD, Ticona A, Delgado E, Velasquez MRE, Caceres O, Budowle B (2021) Population data of 27 Y-chromosome STRS in Aymara population from Peru. *Aust J Forensic Sci* 1–15. <https://doi.org/10.1080/00450618.2021.1882571>
- Roewer L (2009) Y chromosome STR typing in crime casework. *Forensic Sci Med Pathol* 5(2):77–84. <https://doi.org/10.1007/s12024-009-9089-5>
- Sandoval JR, Lacerda DR, Jota MS, Salazar-Granara A, Vieira PP, Acosta O, Cuellar C, Revollo S, Fujita R, Santos FR, Genographic Project Consortium (2013) The genetic history of indigenous populations of the Peruvian and Bolivian Altiplano: the legacy of the Uros. *PLoS ONE* 8(9):e73006. <https://doi.org/10.1371/journal.pone.0073006>
- Smalldon KW, Moffat AC (1973) The calculation of discriminating power for a series of correlated attributes. *J Forensic Sci* 13(4):291–295. [https://doi.org/10.1016/S0015-7368\(73\)70828-8](https://doi.org/10.1016/S0015-7368(73)70828-8)
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 30 (12): 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Tereba A (1919) Tools for analysis of population statistics. *Profiles DNA* 2:14–16
- Thermo Fisher Scientific (2015). GeneMapper® iii ID-X Software v1.5 Reference Guide. In: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/100031707\\_GeneMapIDX\\_ver1\\_5\\_ReferenceGuide.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/100031707_GeneMapIDX_ver1_5_ReferenceGuide.pdf). Accessed 7 February 2023
- Tineo DH, Loiola S, Paredes FV, Noli LR, Amaya YC, Simão F, Carvalho EF, Gusmão L (2015) Genetic characterization of 27 Y-STR loci in the native population of Ashaninka from Peru. *Forensic Sci Int Genet Suppl Ser* 5:e220–e222. <https://doi.org/10.1016/j.fsigss.2015.09.088>
- Vullo C, Gomes V, Romanini C, Oliveira AM, Rocabado O, Aquino J, Amorim A, Gusmão L (2015) Association between Y haplogroups and autosomal AIMs reveals intra-population substructure in Bolivian populations. *Int J Legal Med* 129(4):673–680. <https://doi.org/10.1007/s00414-014-1025-x>
- Willuweit S, Roewer L (2007) Y chromosome haplotype reference database (YHRD): update. *Forensic Sci Int Genet* 1:83–87. <https://doi.org/10.1016/j.fsigen.2007.01.017>
- Yunis JJ, Acevedo LE, Campo DS, Yunis EJ (2005) Population data of Y-STR minimal haplotypes in a sample of Caucasian-Mestizo and African descent individuals of Colombia. *Forensic Sci Int* 151(2–3):307–313. <https://doi.org/10.1016/j.forsciint.2005.02.005>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---