

**SURNAMES IN CHILE**

**A study of the population of Chile through isonymy**

I. Barrai, A. Rodriguez-Larralde<sup>2</sup>, J. Dipierri<sup>1</sup>, E. Alfaro<sup>1</sup>, N. Acevedo<sup>3</sup>,

E. Mamolini, M. Sandri, A. Carrieri and C. Scapoli.

Dipartimento di Biologia ed Evoluzione, Università di Ferrara, 44121- Ferrara, Italy

<sup>1</sup>Instituto de Biología de la Altura, Universidad Nacional de Jujuy, 4600 – San Salvador De Jujuy, Argentina.

<sup>2</sup>Centro de Medicina Experimental, Laboratorio de Genética Humana, IVIC, 1020A -Caracas, Venezuela.

<sup>3</sup>Museo Nacional de Ciencias Naturales, Santiago, Chile

**Running title:** Surnames in Chile

Correspondence to:

Chiara Scapoli

Department of Biology and Evolution

University of Ferrara,

Via L. Borsari 46, - I-44121 Ferrara, Italy.

Telephone: +39 0532 455744; FAX: : +39 0532 249761

Email: scc@unife.it

Number of text pages: 15

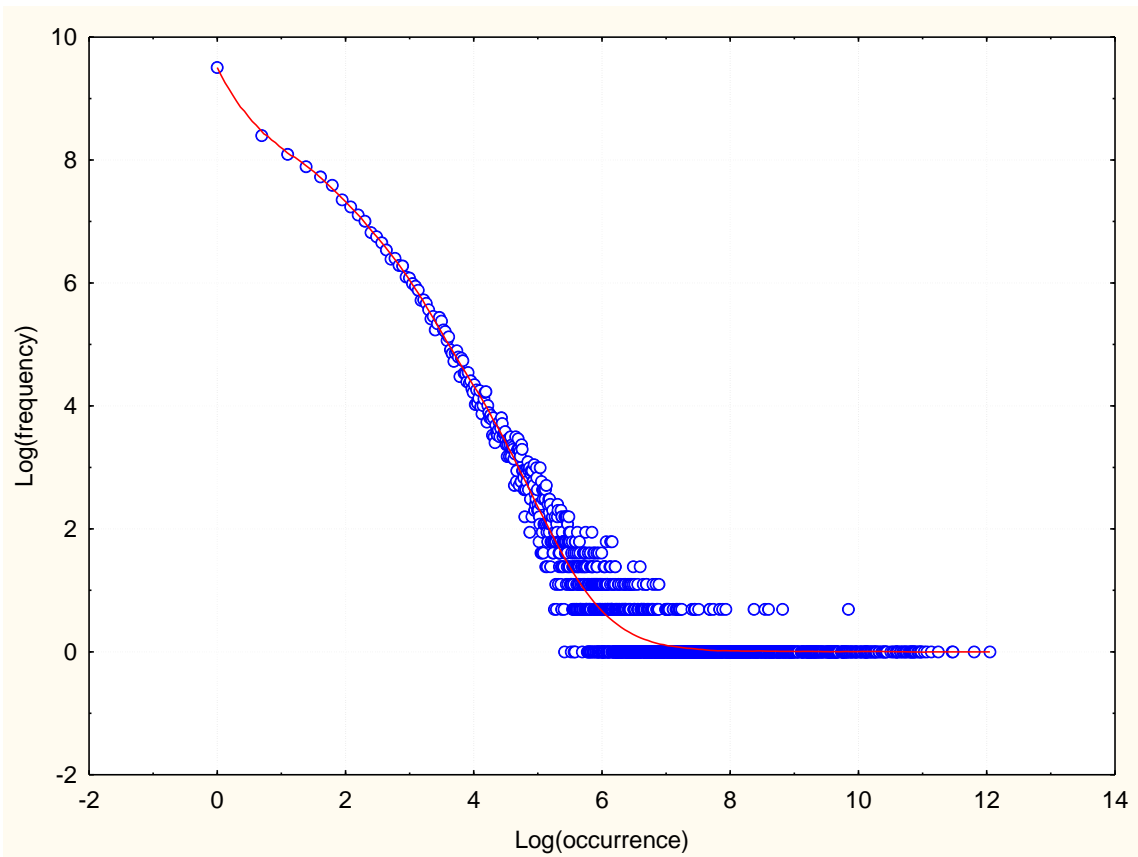
Literature pages: 4

Number of Tables : 2

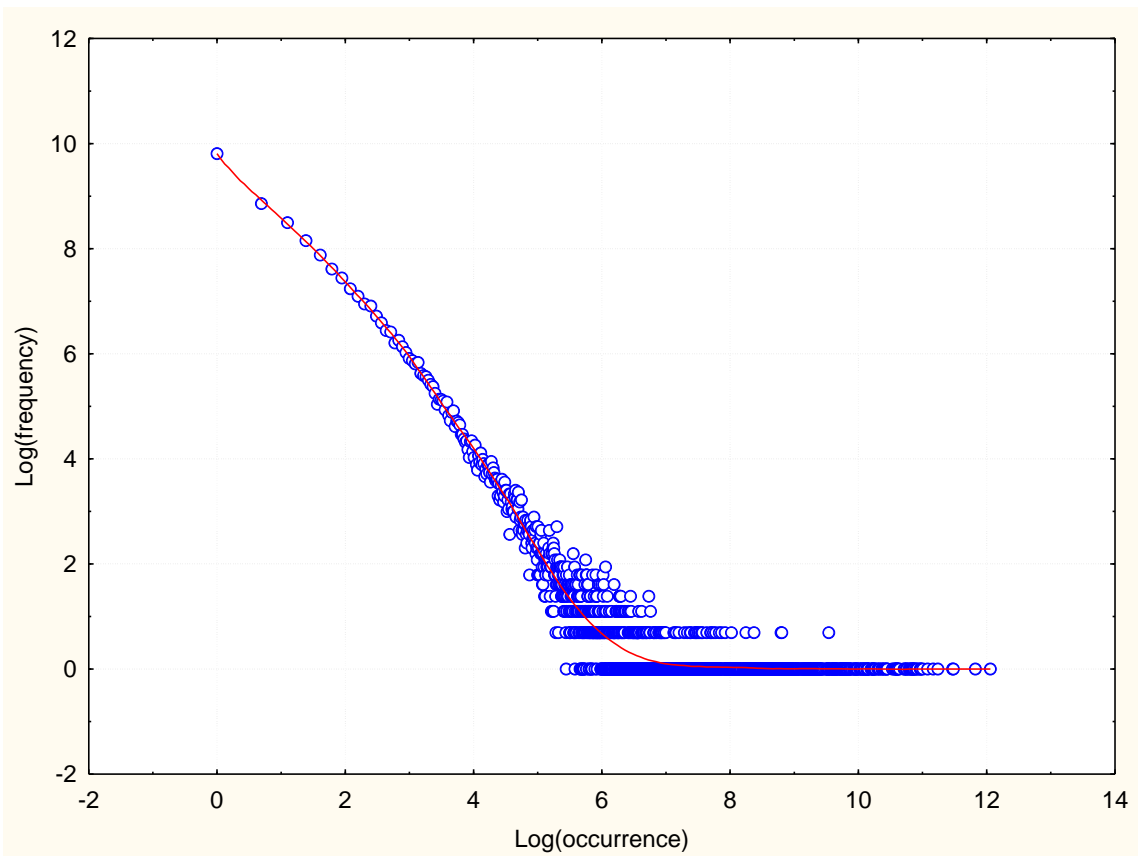
Number of Figures: 7

**KEYWORDS:** Chile, Population Structure, Isonymy, Inbreeding, Isolation by distance

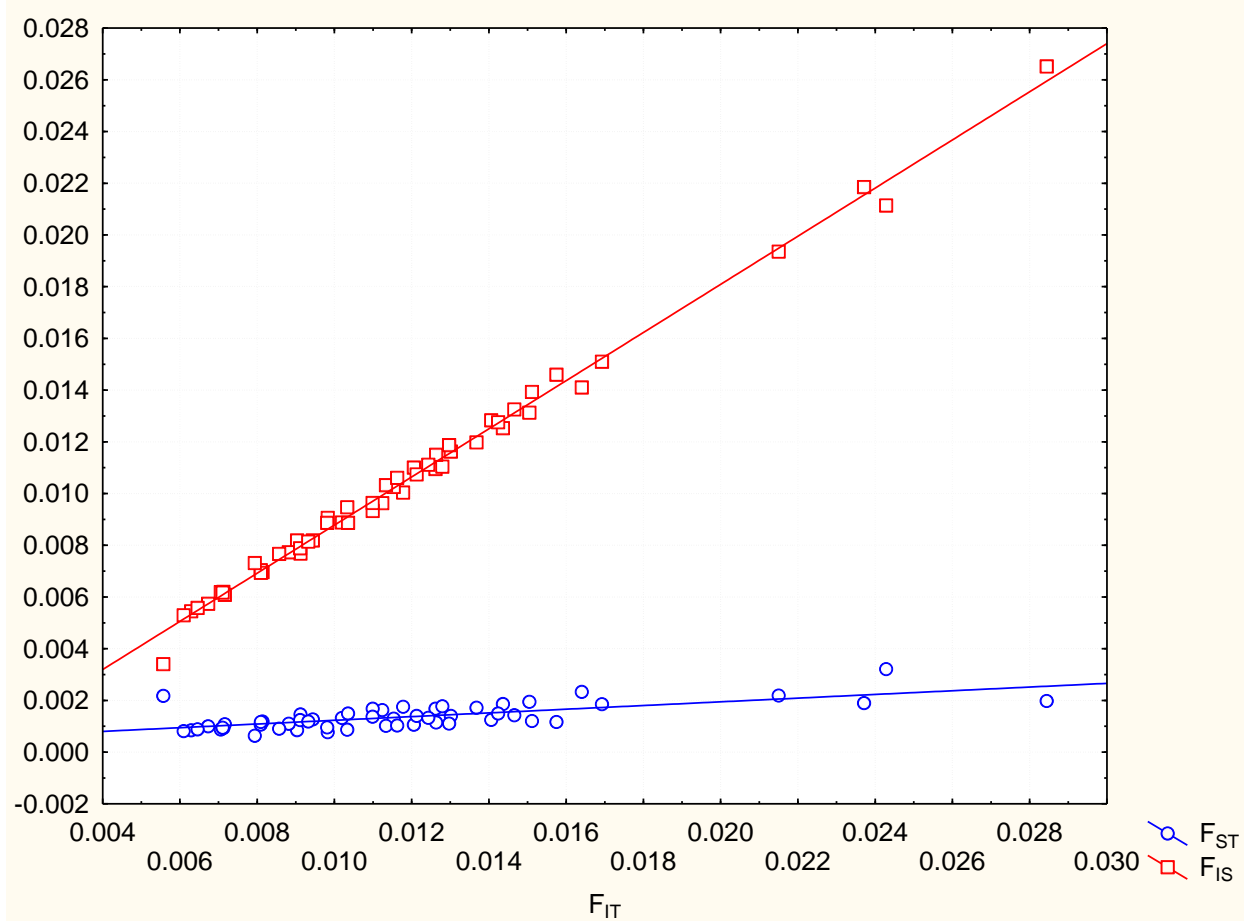
**ACKNOWLEDGMENTS:** The authors are grateful to the Director of the Servicio Electoral de la Republica de Chile Sr. Juan Ignacio Garcia Rodríguez, who made the data available, and to Sr. Dr. Ginés Mario Gonzalez Garcia, Embajador de la Republica Argentina en Chile. The work was supported by grants of the Italian Ministry of Universities and Research (MIUR) to Chiara Scapoli. The authors take pleasure in dedicating this work to Dr. Juan Pinto Cisternas of Valparaiso, a pioneer of isonymy studies.



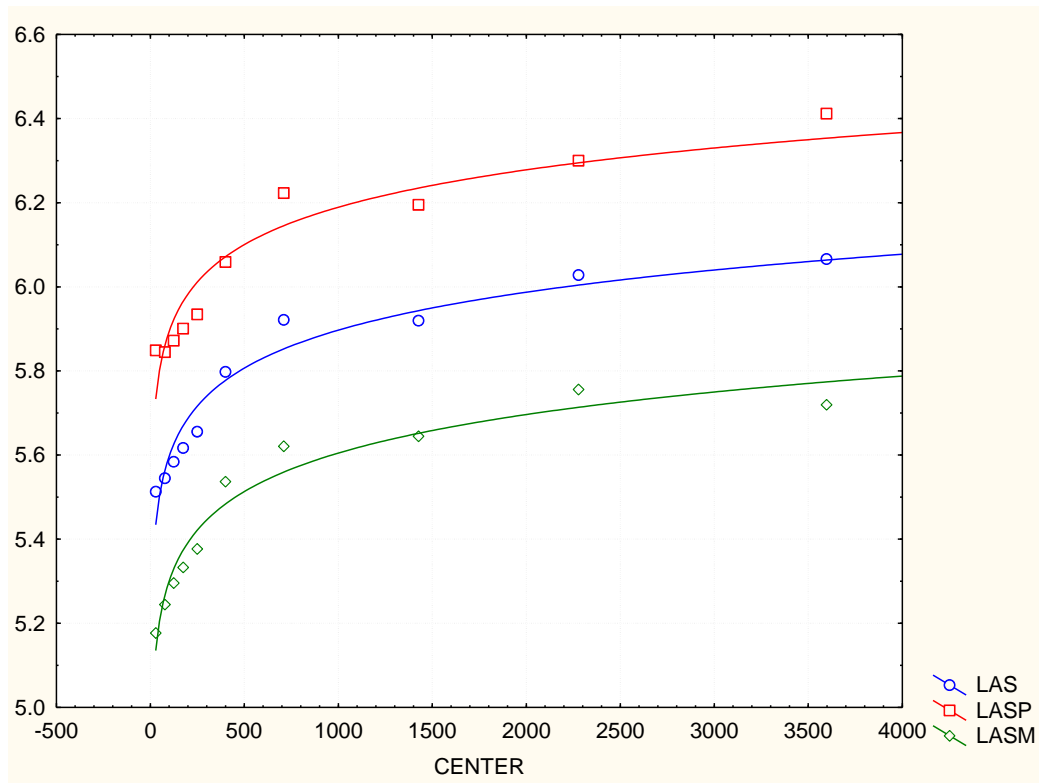
**Figure S1.** Frequency of a given occurrence of a surname as a function of its occurrence. 8.1 million paternal surnames, Chile 2006. Bilogarithmic scale.



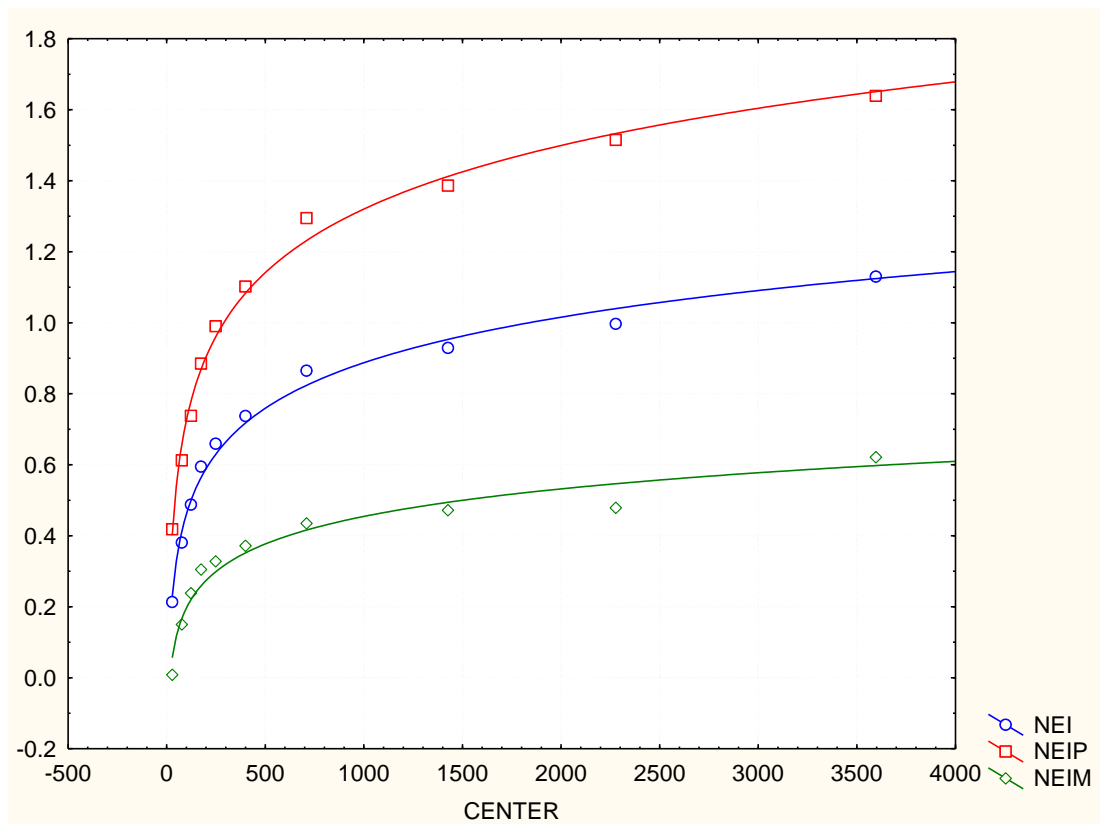
**Figure S2.** Frequency of a given occurrence of a surname as a function of its occurrence. 8.1 million maternal surnames, Chile 2006. Bilogarithmic scale.



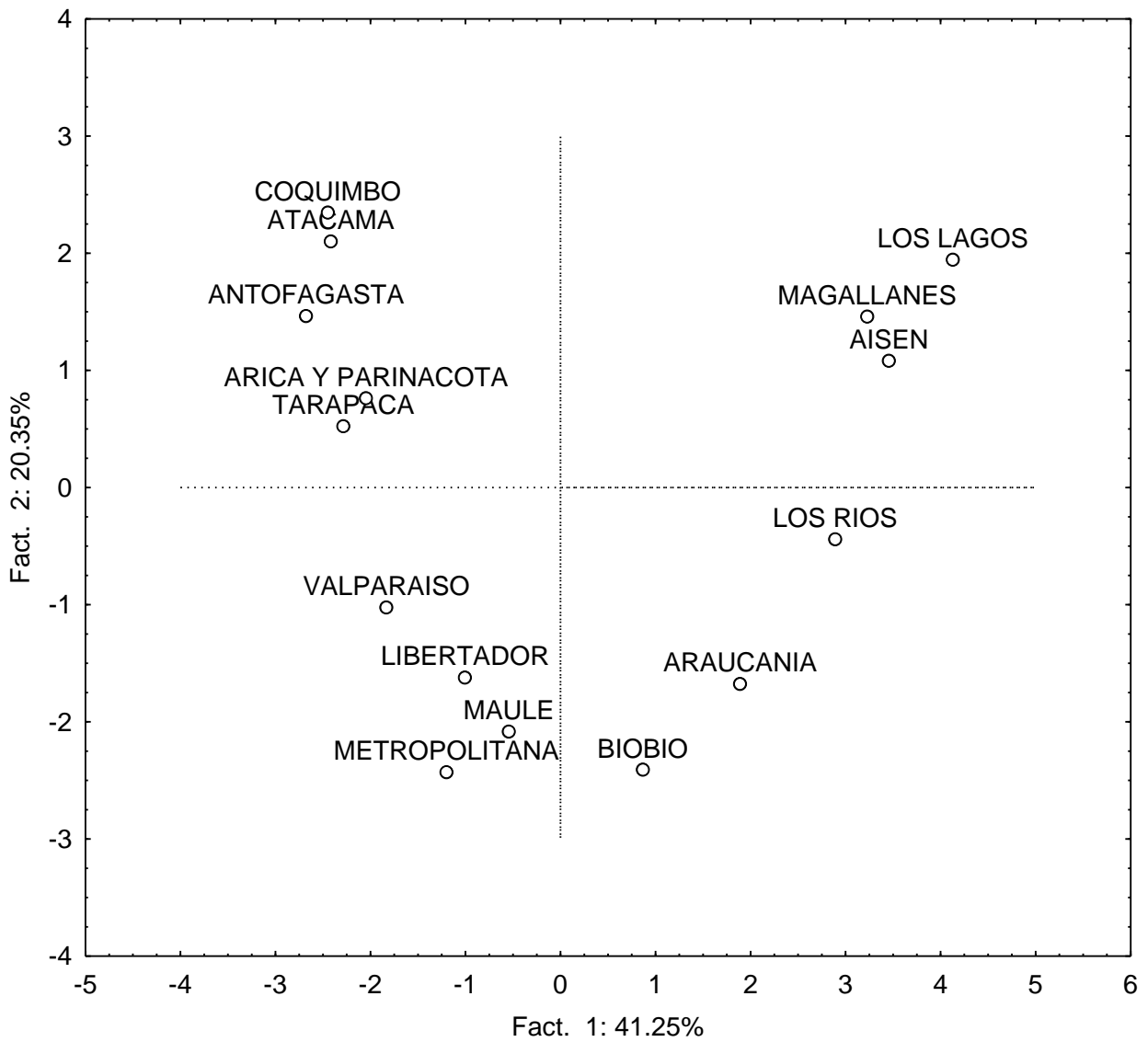
**Figure S3.** The components of inbreeding levels in 54 provinces of Chile. Local inbreeding seems to be the largest component, whereas random inbreeding tends to stay constant. Inbreeding from isonymy, 16 million surnames, Chile 2006.



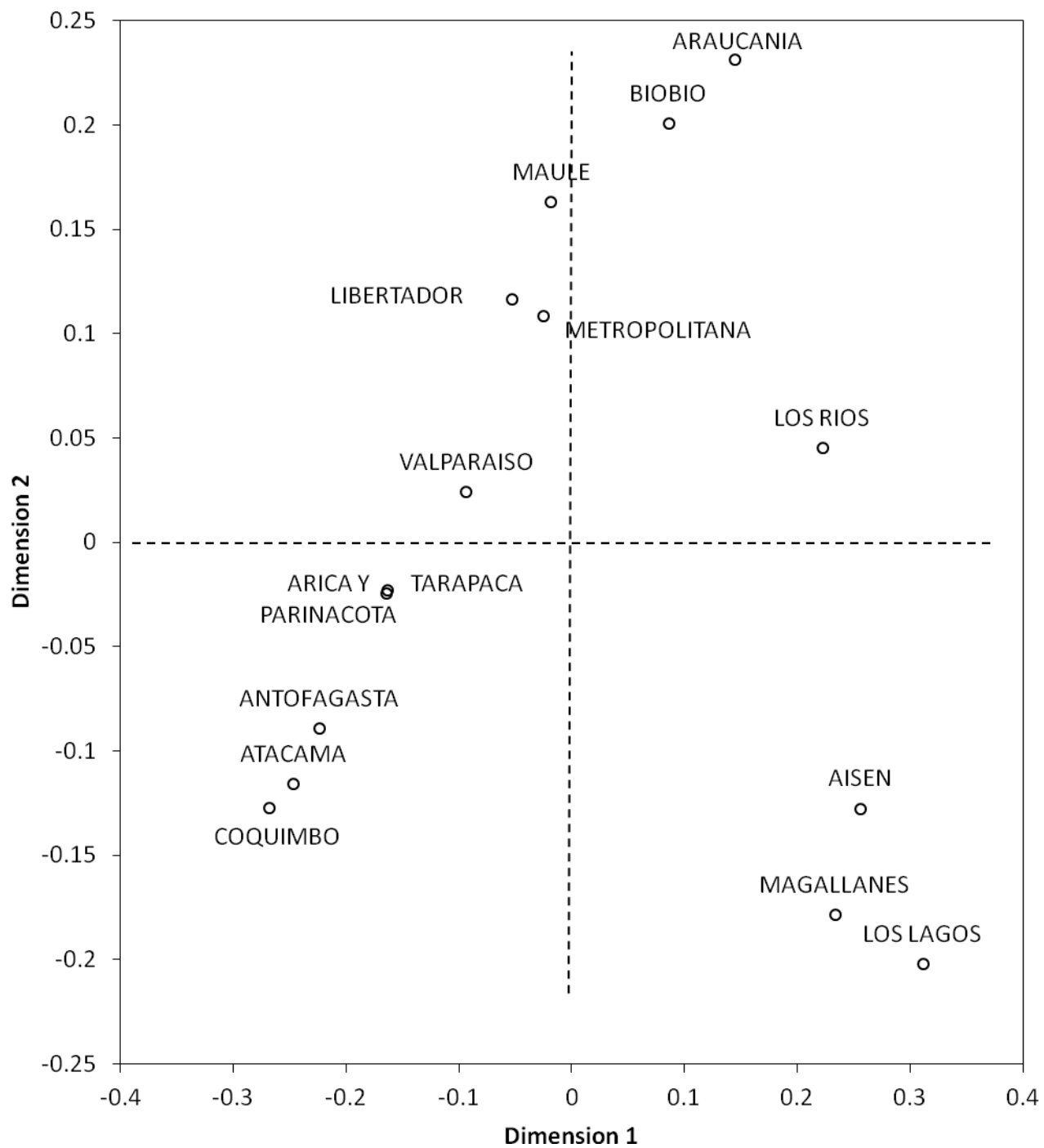
**Figure S4.** Variation of Lasker's distance on kilometers, Chile 2006. The belts are one standard deviation wide.



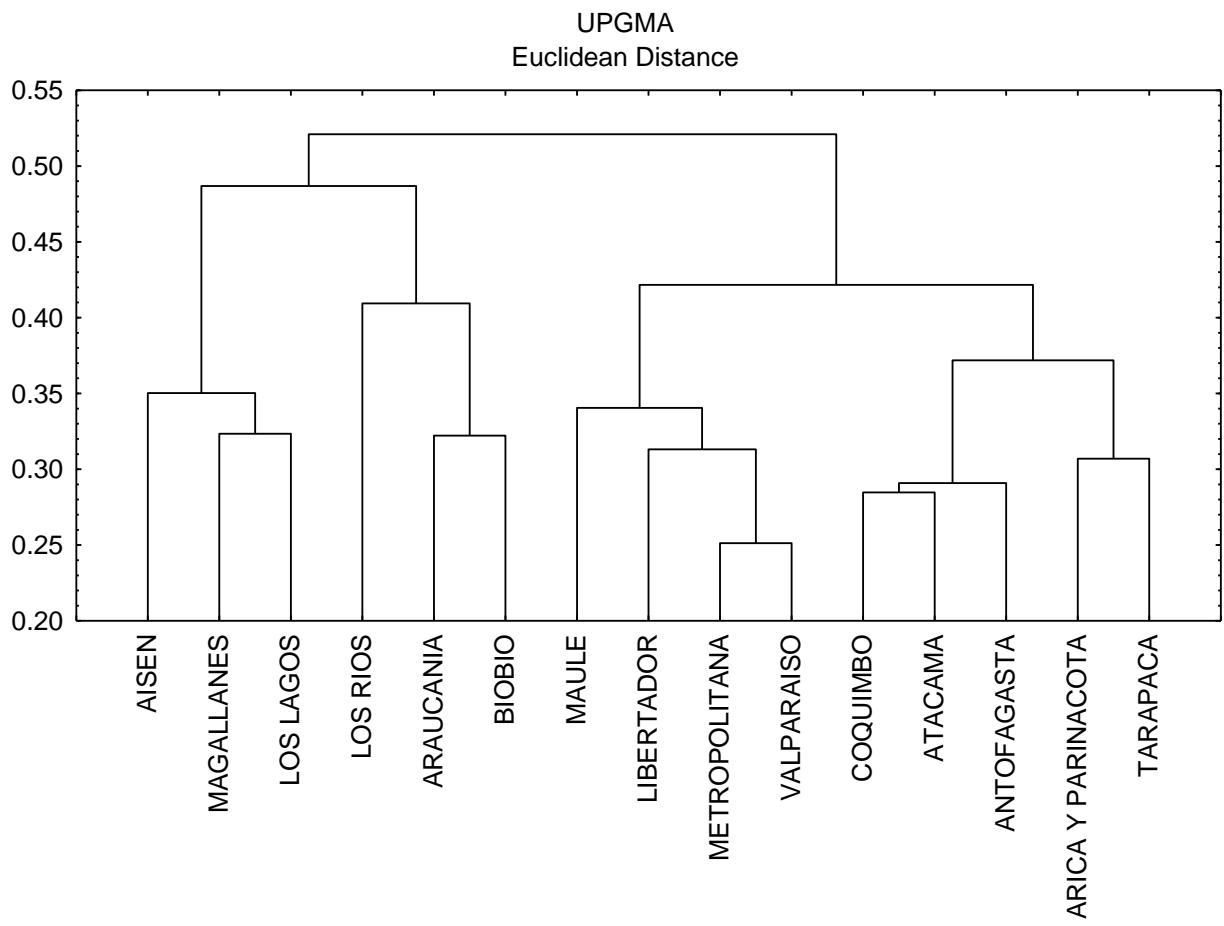
**Figure S5.** Variation of Nei's distance on kilometers, Chile 2006. The belts are one standard deviation wide.



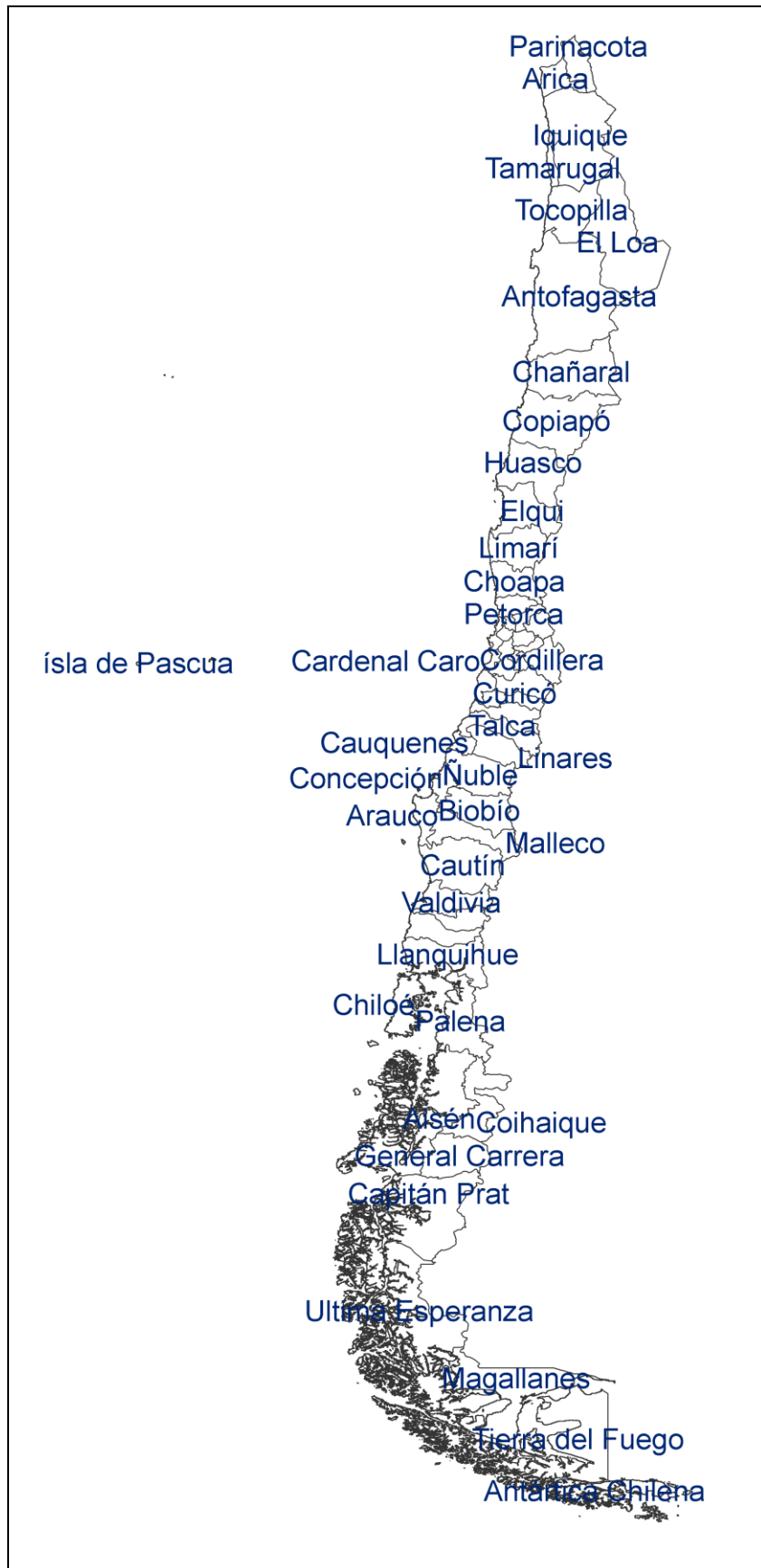
**Figure S6P.** Projection of the Euclidean distance matrix between regions on the first two Factors of PCA. Regions result geographically ordered in a concave arc from North to South.



**Figure S6M.** Projection of the Euclidean distance matrix between regions on the first two Dimensions of MDS. Regions result geographically ordered in a convex arc from North to South.

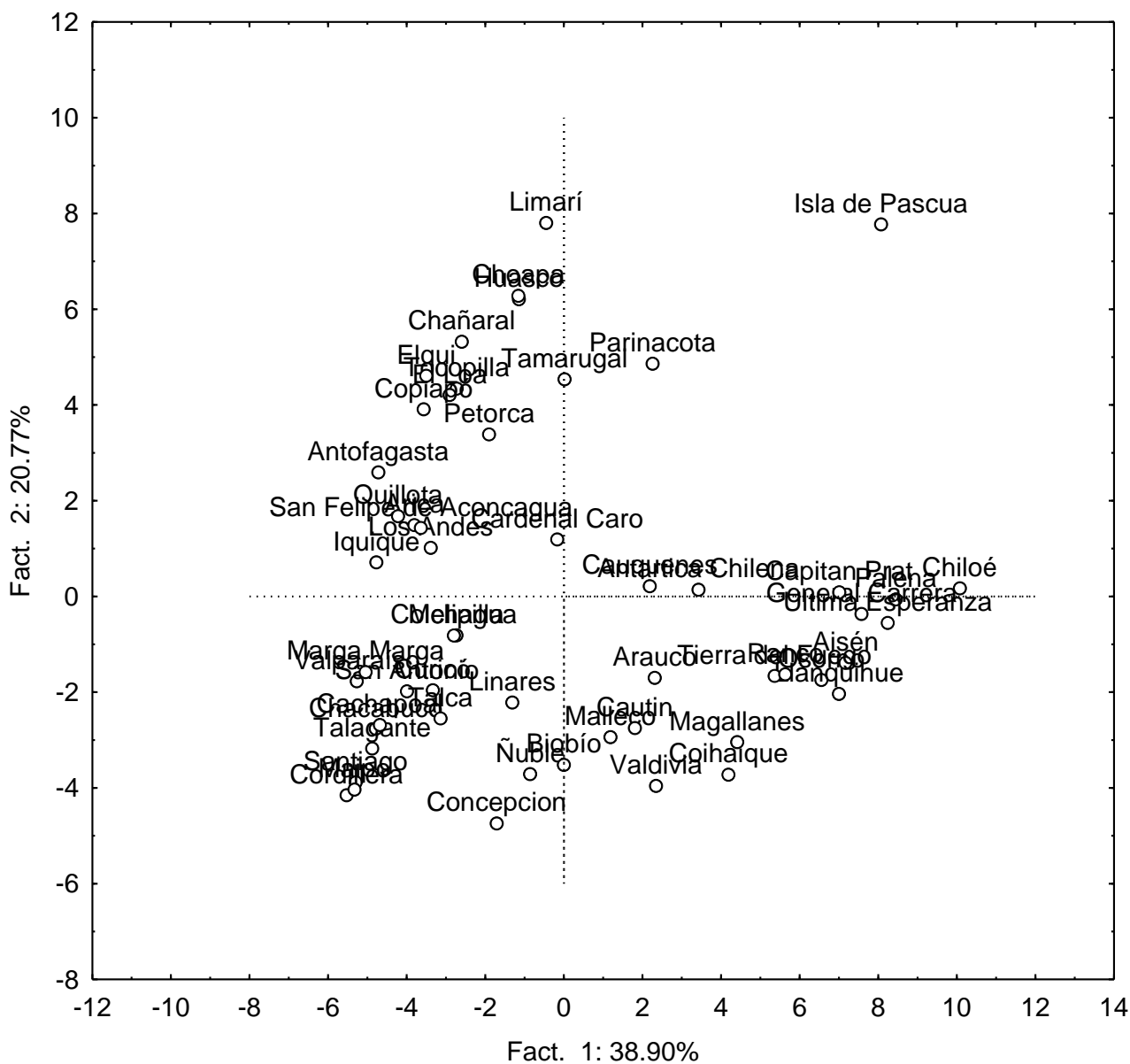


**Figure S7.** Dendrogram based on the matrix of Euclidean surname distances between regions. Note the strict North-South clustering from Tarapaca to Aisen.

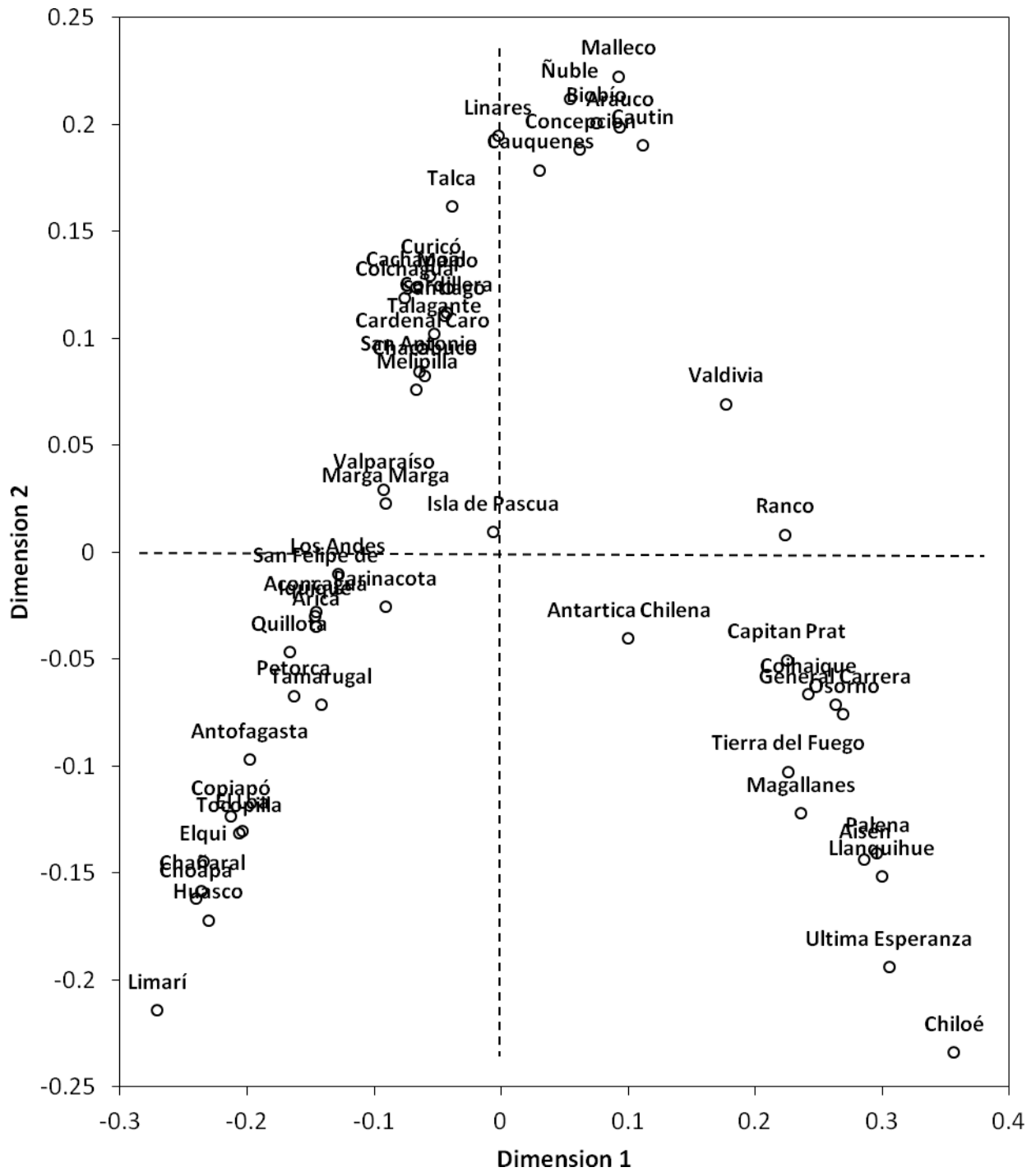


**Figure S8.** The provinces of Chile.

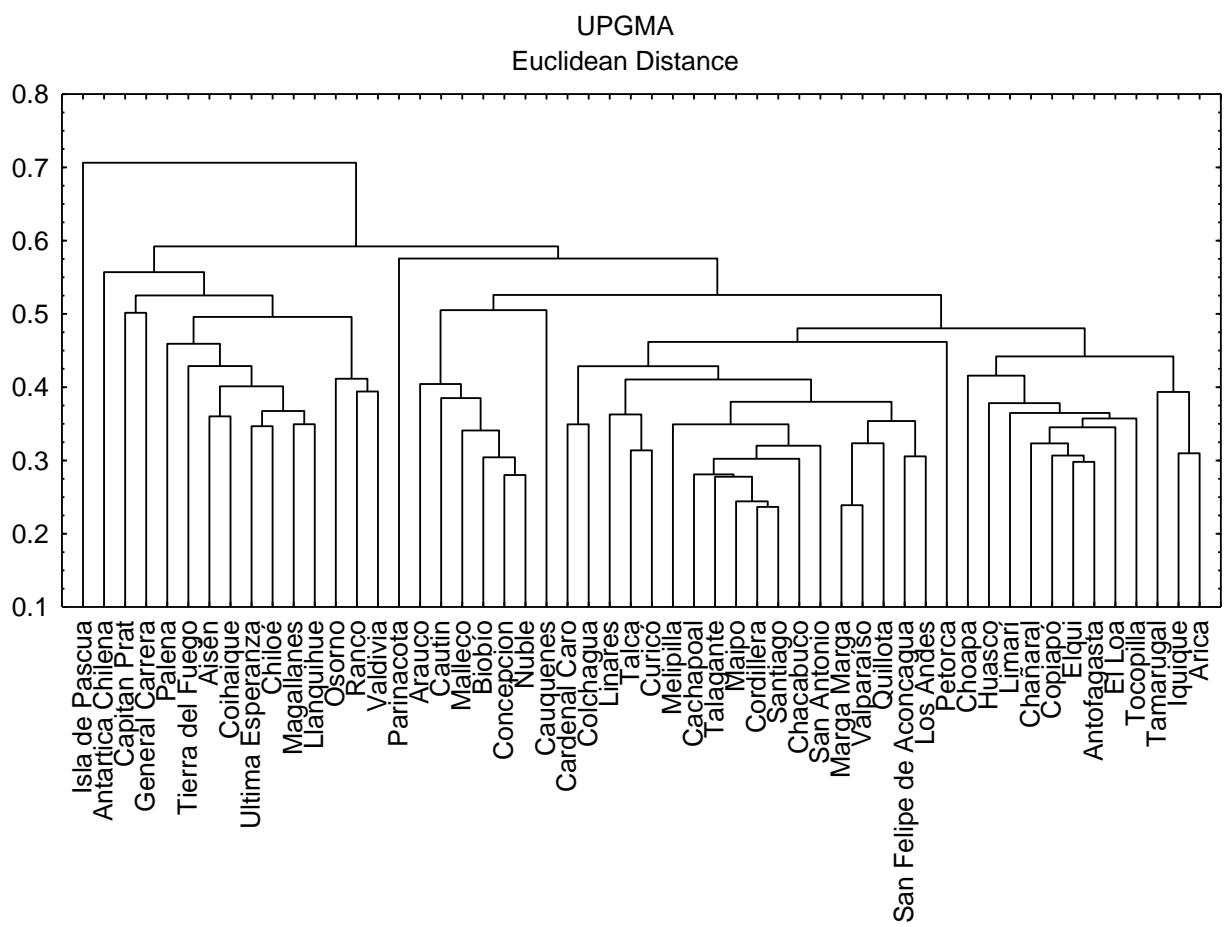




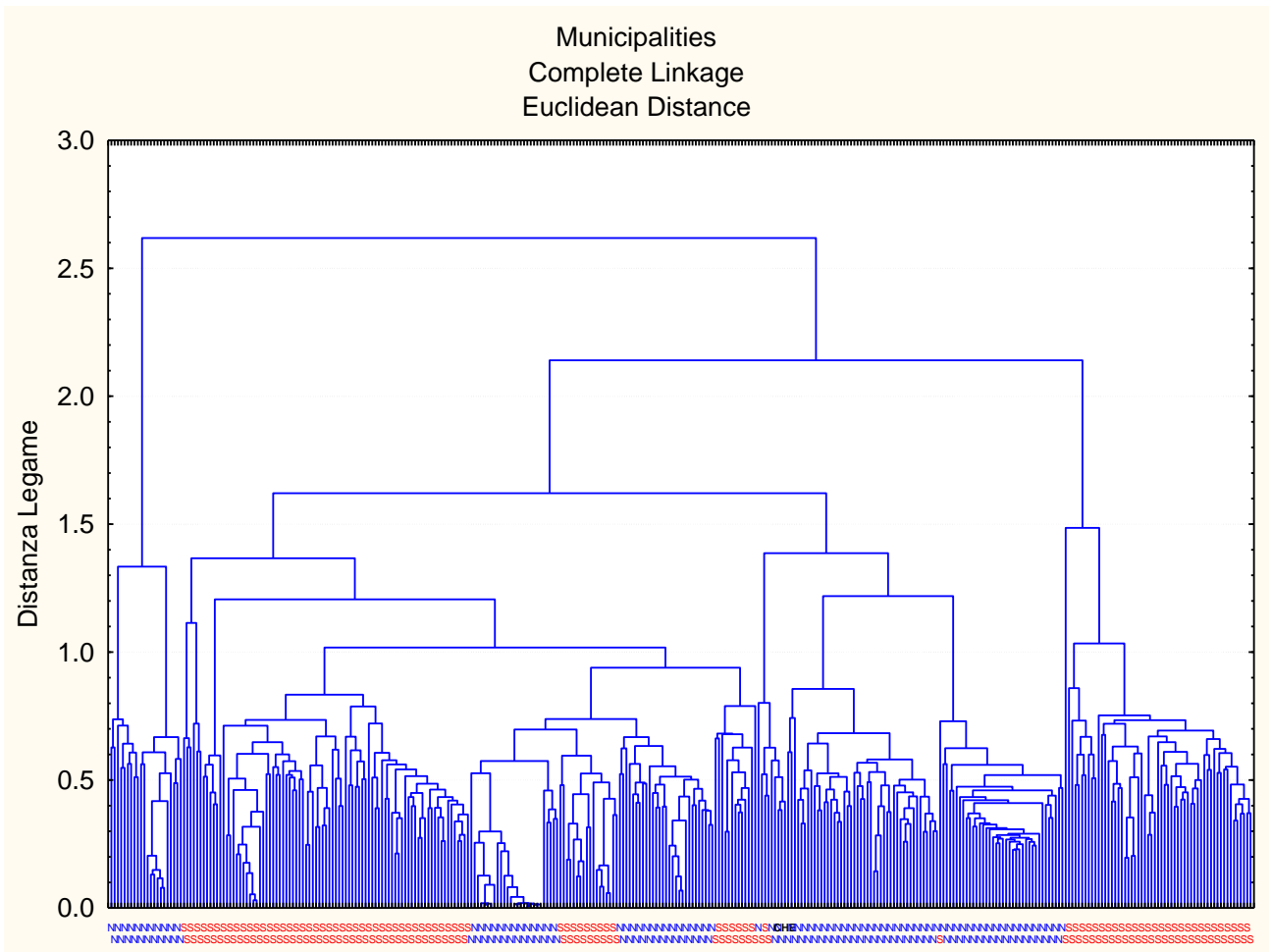
**Figure S9P.** Projection of the Euclidean distance matrix between provinces on the first two Factors of PCA. Note the counterclockwise arc ordering of the provinces in the second, third, fourth, and first quadrant. Note the outlier position of Isla de Pascua (Easter Island) in the first quadrant.



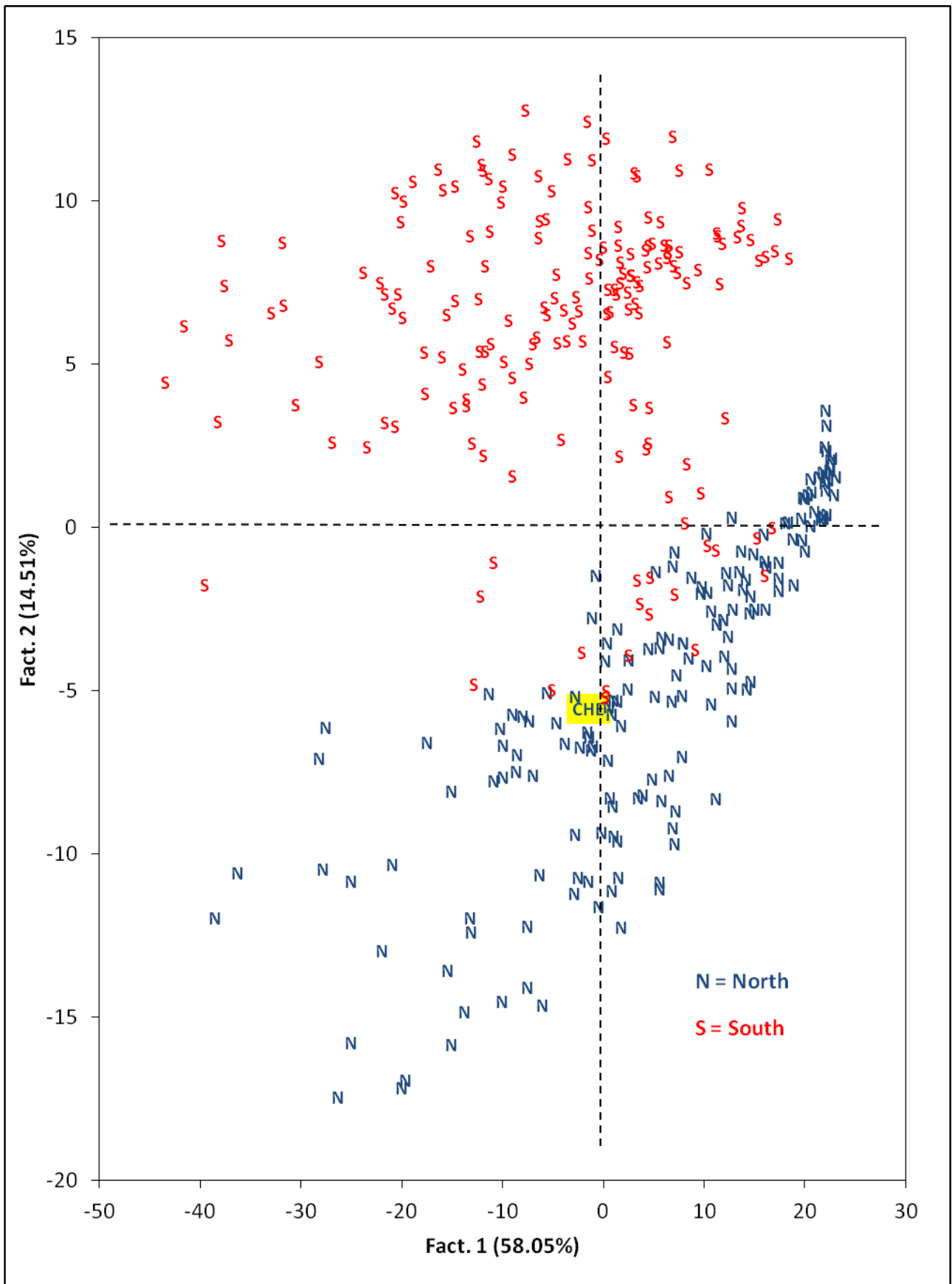
**Figure S9M.** Projection of the Euclidean distance matrix between provinces on the first two Dimensions of MDS. Note the counterclockwise arc ordering of the provinces in the third, second, first and fourth quadrant.



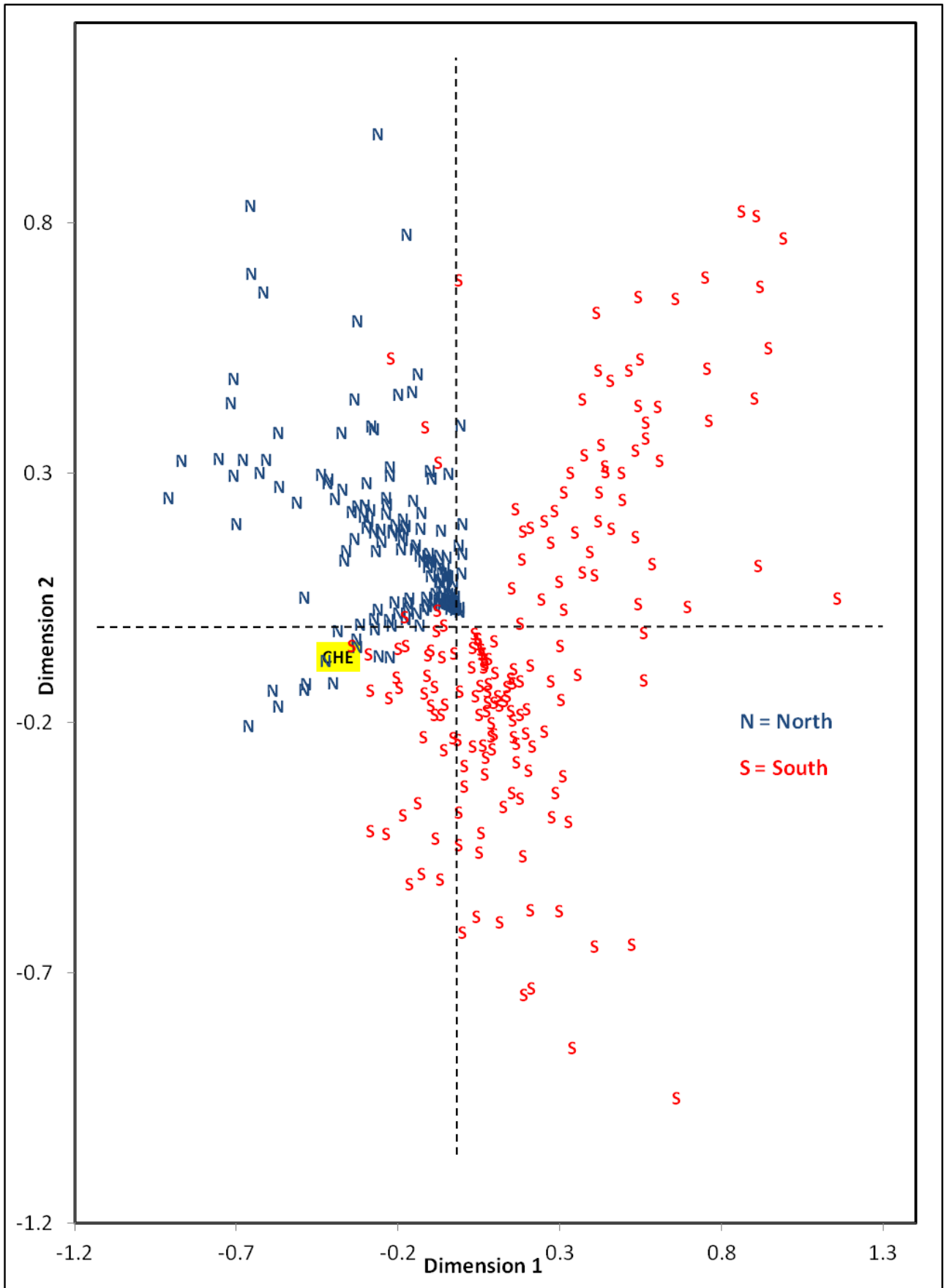
**Figure S10.** Dendrogram of the 54 provinces of Chile. Note the considerable North-South ordering of the provinces from Arica down to Tierra del Fuego and Antartica Cilena.



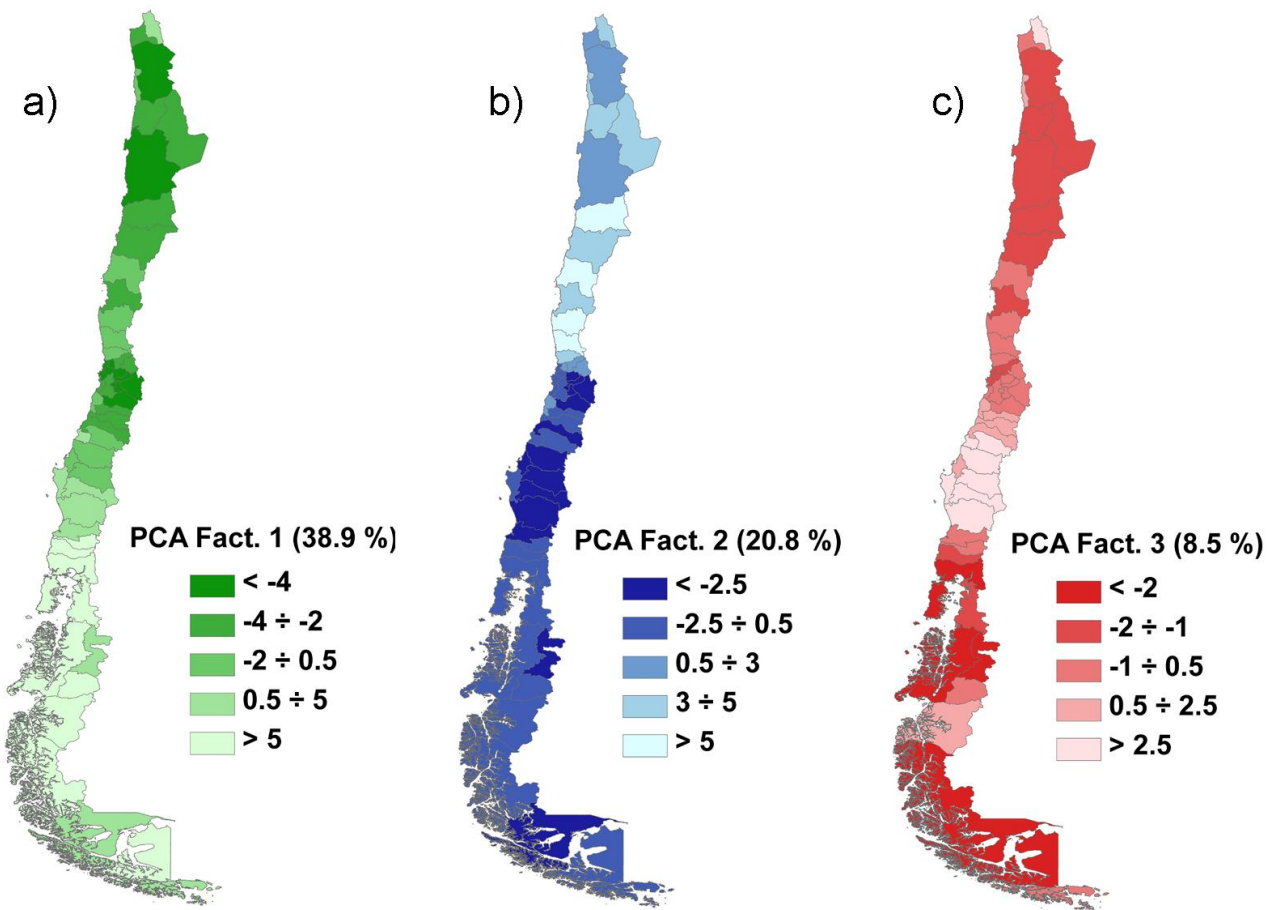
**Figure S11.** The dendrogram from the Euclidean distance matrix between the 346 communes of Chile. (Label CHE means Chepica municipality; N = North; S = South)



**Figure S12P.** Projection of the Euclidean distance matrix between the 346 communes of Chile on the first two Factors of PCA. Note that the North (N) and the South (S) clusters are inverted in the Projection. Label CHE means Chepica municipality.



**Figure S12M.** Projection of the Euclidean distance matrix between the 346 communes of Chile on the first two Dimensions of MDS. Label CHE means Chepica municipality.



**Figure S13.** Mapping of Euclidean's matrix of surname distances between departments in Chile on the first three components of Pca. The first component (a) green) represents 38.9% of the total dispersion, the second (b) blue) 20.8%, and the third component (c) red) 8.5%.











**Table S2.** The first 50 surnames in Chile, paternal and maternal series.

<i>Paternal</i>	<i>N</i>	<i>Maternal</i>	<i>N</i>
GONZALEZ	171266	GONZALEZ	172755
MUÑOZ	133998	MUÑOZ	136343
ROJAS	96090	ROJAS	96793
DIAZ	94912	DIAZ	95589
PEREZ	76643	PEREZ	76053
SOTO	68560	SOTO	70602
CONTRERAS	64057	CONTRERAS	64716
SILVA	61416	SILVA	60162
SEPULVEDA	58638	SEPULVEDA	58613
MARTINEZ	58330	MARTINEZ	58146
MORALES	57962	MORALES	57747
RODRIGUEZ	56353	LOPEZ	55210
LOPEZ	55418	RODRIGUEZ	55086
FUENTES	53657	FUENTES	53675
ARAYA	52596	TORRES	52383
TORRES	52393	ARAYA	52313
HERNANDEZ	52350	HERNANDEZ	52070
ESPINOZA	51132	FLORES	51562
VALENZUELA	50590	ESPINOZA	51042
FLORES	50462	VALENZUELA	50610
CASTILLO	49786	CASTILLO	49584
RAMIREZ	49316	RAMIREZ	49174
REYES	48353	REYES	48475
GUTIERREZ	46829	GUTIERREZ	46959
CASTRO	46104	CASTRO	46642
VARGAS	45640	VARGAS	46536
ALVAREZ	44468	VASQUEZ	45231
VASQUEZ	43960	ALVAREZ	44866
FERNANDEZ	42641	TAPIA	41821
TAPIA	42268	CARRASCO	40950
SANCHEZ	40822	SANCHEZ	40582
CORTES	40347	FERNANDEZ	40167
HERRERA	39925	CORTES	40100
CARRASCO	39811	GOMEZ	39858
GOMEZ	39795	HERRERA	39422
NUÑEZ	38325	NUÑEZ	39142
JARA	37672	JARA	38334
VERGARA	36579	VERGARA	37338
RIVERA	34321	FIGUEROA	34522
FIGUEROA	34072	RIVERA	34337
GARCIA	33385	RIQUELME	33967
RIQUELME	33114	VERA	33209
BRAVO	32903	MIRANDA	32747
VERA	32071	BRAVO	31986
MIRANDA	31696	GARCIA	31800
MOLINA	30464	MOLINA	30652
VEGA	30309	VEGA	30273
SANDOVAL	29194	CAMPOS	29762
CAMPOS	29116	SANDOVAL	28667
OLIVARES	28660	ORELLANA	28618