

Y-chromosome diversity in central Portugal reveals genetic signatures regarding Phoenician colonization

Rui Martiniano, Augusto Abade, Licínio Manco

Research Centre for Anthropology and Health (CIAS) – Department of Life Sciences,
University of Coimbra, 3000 Coimbra, Portugal

Email: lmanco@antrop.uc.pt

Number of illustrations: 1

Number of tables: 2

Abstract

The genetic diversity of human populations in Portugal results from several different demographic events that occurred in distinct historical periods. Phoenician migrations reached the Portuguese coast in the 1st Millennium BCE which may have influenced the genetic background in coastal areas. The main objective of this study was to examine if different patterns of Y-chromosome diversity explained by putative Phoenician signatures can be observed between littoral versus most inland regions in central Portugal. A total of 143 male DNA samples from Coimbra district, in central-west region of Portugal, were typed for 16 Y-SNPs and eight Y-STRs using standard molecular methodologies. Thirteen different haplogroups were identified, being the typical Western European haplogroup R1b1b2-M269 the most common (0.601), followed by J-M304 (0.147) and E1b1b1-M35 (0.112). Taking haplogroup J-M304 individually, whose origin maps to the Middle East, a higher frequency was observed in littoral (0.166) compared with inland region (0.108). Moreover, a significant difference ($p=0.022$) was observed when the littoral region was compared with the innermost region of Beira Interior, in the same geographic area of the country. Y-STR haplotypes previously associated with ancient Phoenician expansions were found within haplogroups J-M304 and E1b1b1a-M78. In conclusion, this study supports the hypothesis a maritime entry of Middle Eastern haplogroups in the coastal region of Central Portugal, most probable through the Phoenician migrations in the first millennium BCE.

Key words: Y-SNPs; Y-STRs; Central Portugal; Middle East; Phoenicians.

Introduction

The genetic diversity of human populations in Portugal results from several different migrations that occurred in distinct historical periods. The largest part included the Palaeolithic colonization about 40 000 years ago, followed by re-expansions from refugia after the Last Glacial Maximum in the Mesolithic period (Rootsi et al. 2004). The spread of Neolithic farmers from the Near East into Europe, around 10 000 years ago, related to the development of agricultural practices, is considered the second most important demographic event (Semino et al. 2000). Other subsequent settlements of the territory included various Indo-European peoples among which Celts, Roman, Germanic or Greeks (Saraiva, 1992). In the last two millennia the long residence of two important communities with different origins, the Sephardic-Jews, most arriving during the Roman period, and the North African Arab-Berbers, invading the territory in 711 CE, also had a considerable demographic impact in the territory (Adams et al. 2008).

It is also well known that ancient Mediterranean contacts have been established across the Portuguese coast, including Phoenician settlements in the first millennium BCE, as attested by archaeological ruins at the coastal area in central Portugal implanted on the right bank of the ancient estuary of the Mondego river (ancient ruins of Santa Olaia – municipality of Figueira da Foz) (Tavares, 2004; Arruda, 1999). A recent work by Zalloua et al. (2008), using Y-chromosome diversity, comparing historically documented Phoenician-influenced sites throughout the Mediterranean shore with neighbouring non-Phoenician sites, found weak but systematic signatures shared by these Phoenician sites.

In the last years, several studies of the Y-chromosome polymorphisms contributed to the knowledge of the Portuguese demographic history on its male component (Gonçalves et al. 2005; Beleza et al. 2006; Adams et al. 2008). The central region of Portugal, namely Coimbra district, was also accomplished with Y-chromosome studies; however comparisons between littoral and inland areas were not considered to explain higher than expected frequencies of specific lineages, including Y-haplogroups with probable Middle Eastern origin: J-M304 lineage with a 25% frequency was drawn from a small population sample in Coimbra district (Beleza et al. (2006).

The main purposes of this study were i) to address a more accurate distribution for Y-chromosome lineages in Coimbra district; ii) to examine if different patterns of Y-chromosome diversity explained by putative Phoenician genetic signatures can be observed between littoral and most inland regions in central Portugal .

Material and Methods

Samples

Buccal swabs were collected from a total of 143 non-related male volunteers of Coimbra district (central Portugal; total area, ~4000 km²): 97 from the coastal area, at municipalities of Montemor-o-Velho, Figueira da Foz, Cantanhede and Mira, and 46 from the innermost region of the district (distributed by eight different municipalities). Individuals were chosen based on his male geographical origin considering the last three generations. Only individuals with different surnames were considered to ensure non-affiliation. Written informed consent was obtained from all participants.

Genotyping

Genomic DNA was extracted using the QIAamp DNA Micro kit (Qiagen GmbH, Hilden, Germany). For the identification of the most common haplogroups described in the Portuguese population 16 biallelic markers (M213, M9, M269, M304, M35, M78, M81, M123, M410, M12, M70, M172) were studied by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) or automatic sequencing by Sanger's dideoxy chain termination reaction (M267, M170, M201, M22), using *primers* previously described (Karafet et al. 2008). Haplogroups were defined according to the nomenclature proposed by Karafet et al. (2008).

Haplotypes were obtained from the analysis of eight microsatellites (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 and DYS388), selected on basis of Zalloua et al. (2008) work to differentiate Phoenician genetic signals, or Adams et al. (2008) study on Sephardic Jews samples. PCR fragments were analysed on an automate ALFexpressTM II sequencer (Amersham Biosciences, Uppsala, Sweden) using Cy5-labeled *primer* sequences obtained online at http://www.cstl.nist.gov/div831/strbase/ystr_fact.htm.

Statistical analysis

Haplogroup frequencies were estimated by simple counting. The software Arlequin v3.11 (Excoffier and Schneider, 2005; <http://cmpg.unibe.ch/software/arlequin3/>) was used to estimate diversity values and pairwise F_{ST} genetic distances. Statistical differences between population groups for individual SNPs were assessed by Fisher's exact probability tests using contingency tables on SISA (www.quantitativeskills.com). Results were compared with several reference mainland regions of Portugal reported in Beleza et al. 2006. Comparisons for individual haplogroups were made only between regions of the same geographic area (central Portugal) to avoid allele stratifications easily found in Y-chromosome variation along mainland Portugal. A median-joining network of STR haplotypes in haplogroup J1-M267 was constructed and drawn with the Network 4.610 program (Bandelt et al. 1999; <http://www.fluxus-engineering.com>).

Results and Discussion

Observed haplogroups and respective frequencies are shown in Table 1 for whole sampled population as for sub-populations of littoral and inland regions of Coimbra district. The 16 typed SNPs allowed identifying 13 different haplogroups. The typical Western European haplogroup R1b1b2-M269 was the most common, with an overall frequency of 0.601, followed by J-M304 and E1b1b1-M35, with frequencies of 0.147 and 0.112 respectively. The overall genetic diversity was 0.625. These proportions were similar to that observed in other studies for Portuguese or Iberian populations (Beleza et al. 2006; Adams et al., 2008). When pairwise differentiation tests based on haplogroup frequencies were computed between Coimbra district and the eight reference Portuguese populations previously reported by Beleza et al. (2006) no significant distances were observed (F_{ST} p values $\gg 0.05$), except for the region of Alentejo with a marginally significant distance (F_{ST} p value = 0.046). In the same way, no significant differences were found between littoral and inland regions of Coimbra district (F_{ST} p value = 0.777).

Taking haplogroup J-M304 individually, whose origin maps to the Middle East, a higher frequency of 0.147 was found in Coimbra district tending to corroborate results obtained by Beleza et al. (2006) (0.25 frequency). In littoral, a higher frequency (0.166) was

found when compared to the inland region (0.108), with no significant distances ($p=0.14$) between the two regions. However, a significant difference ($p=0.022$) was found when the littoral region was compared with the innermost region of Beira Interior (data from Beleza et al. 2006) in the same geographic area of the country. This gradient frequency on J-M304 lineages from the littoral to the innermost regions in central Portugal favours the hypothesis that the entry of J chromosomes in the littoral region may have occurred mainly by sea, similar to what has been suggested for Mediterranean coastal regions (Di Giacomo et al. 2004).

The study of Y-STR haplotypes within J1-M267 lineages reveals one chromosome in the littoral area (sample ID 6; Table 2) carrying the PCS1+ core haplotype 14-13-16-24-10-11-12, described by Zalloua et al. (2008) as “Phoenician Colonization Signal” (PCS) in Mediterranean coastal regions. In a network constructed using J1-M267 STR haplotypes (Figure 1), no matches were observed between our study population and Portuguese Jews (Nogueiro et al. 2010) or Sephardic Jews from Iberian Peninsula (Adams et al. 2008). Moreover, no similarity was found with the North African Arab modal haplotype, DYS19-14 / DYS388-17 / DYS390-23 / DYS391-11 / DYS392-11 (Nebel et al. 2002; Bosch et al. 2001).

Regarding J2-M172 branch, eight J2a and three J2b chromosomes were found in whole population of Coimbra district, reaching an overall frequency of 0.077 (0.083 in littoral region). Also a chromosome from the inland Coimbra region (sample ID 12; Table 2) classified as J2a-M410 carry the core PCS1+ signal. One single J2a chromosome, found in littoral region, presents a Jewish Y-chromosome haplotype (sample ID 11; Table 2) when compared to data reported by Nogueiro et al. (2010) and Adams et al. (2008) (Figure 1).

Thus, despite influence of the Neolithic demic diffusion or other demographic contributions cannot be discarded, the Phoenician settlement hypothesis must be considered as a main justification for the considerably high frequency of J-M304 haplogroup in the littoral region of Coimbra district, reporting Phoenician archaeological evidences (Arruda, 1999).

Regarding the proportions of other individual lineages with origin attributed to the Middle East, haplogroup E1b1b1a-M78 reaches a moderate frequency of 0.041 in littoral region (overall frequency of 0.028 in Coimbra district) (Table 1). The presence of this lineage in south coast of Europe has been suggested as a possible indicator of Greek expansions (Semino et al. 2004). In concordance, the STR analysis within this haplogroup E1b1b1a-M78 allowed us to find one haplotype 13-13-17-24-10-11-13 (sample ID 28; Table 2) described by Zalloua et al. (2008) as a “Greek Colonization Signal” (GCS).

Haplogroup G-M201 was observed with a 0.059 frequency in whole population sample, in range with data from other studies in Portugal (Beleza et al., 2006; Adams et al. 2008). Curiously, a higher frequency in inland Coimbra district (0.108) was observed when compared with littoral (0.021) (Table 1). With its origin placed in Middle East, it is presumed that people carrying G-M201 haplogroup took part in the Neolithic demic expansions. Nevertheless, Y-STR haplotypes in G-M201 chromosomes found in Coimbra district showed no similarities with Y-haplogroup G2a reported by Lacan et al. (2011a,b) in ancient DNA from two Neolithic populations from Southern France and Northern Spain.

In conclusion, the study of Y chromosome diversity in Coimbra district (central Portugal) allowed to identify the major haplogroups described for European populations in general and for the Portuguese population in particular. The high frequency of J-M304 haplogroup in littoral regions, showing a decreasing gradient to inland regions, in association

with previously described “Phoenician Colonization Signals”, suggest that the entry of Middle Eastern haplogroups, in particular the J-M304 chromosomes, in coastal region of Central Portugal, may have occurred mainly by sea, most probable through the Phoenician migrations in the first millennium BCE.

Acknowledgements:

This work was performed with the support of *Fundação para a Ciência e Tecnologia* (FCT) (PEst-OE/SADG/UI0283/2011). The authors thank all anonymous volunteer donors.

Conflict of interest:

The authors declare no conflict of interest.

References

Adams SM, Bosch E, Balaesque PL, Ballereau SJ, Lee AC, Arroyo E, López-Parra AM, Aler M, Grifo MS, Brion M, Carracedo A, Lavinha J, Martínez-Jarreta B, Quintana-Murci L, Picornell A, Ramon M, Skorecki K, Behar DM, Calafell F, Jobling MA (2008) The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet* 83:725-736.

Arruda AM (2000) Los fenícios en Portugal: Fenícios y mundo indígena en el centro y sur de Portugal. Barcelona: Cuadernos de Estudios Mediterráneos. Barcelona.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37-48

Beleza S, Gusmão L, Lopes A, Alves C, Gomes I, Giouzeli M, Calafell F, Carracedo A, Amorim A. (2006) Micro-phylogeographic and demographic history of Portuguese male lineages. *Ann Hum Genet* 70:181-194.

Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J (2001) High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am J Hum Genet* 68:1019–1029.

Cupido M (2011) Toponímia gandraesa-Apontamentos. Mira: Jornal Voz de Mira Press. 285 p.

Di Giacomo F, Luca F, Popa LO, Akar N, Anagnou N, Banyko J, Brdicka R, Barbujani G, Papola F, Ciavarella G, Cucci F, Di Stasi L, Gavrila L, Kerimova MG, Kovatchev D, Kozlov AI, Loutradis A, Mandarino V, Mammi' C, Michalodimitrakis EN, Paoli G, Pappa KI, Pedicini G, Terrenato L, Tofanelli S, Malaspina P, Novelletto A. (2004) Y chromosomal

haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum Genet* 115:357-371.

Excoffier LG, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47-50.

Gonçalves R, Freitas A, Branco M, Rosa A, Fernandes AT, Zhivotovsky LA, Underhill PA, Kivisild T, Brehm A. (2005) Y-chromosome lineages from Portugal, Madeira and Açores record elements of Sephardim and Berber ancestry. *Ann Hum Genet* 69:443-454.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18:830-838.

Lacan M, Keyser C, Ricaut FX, Brucato N, Tarrús J, Bosch A, Guilaine J, Crubézy E, Ludes B (2011a) Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proc Natl Acad Sci USA* 108:18255-18259.

Lacan M, Keyser C, Ricaut FX, Brucato N, Duranthon F, Guilaine J, Crubézy E, Ludes B (2011b) Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *Proc Natl Acad Sci USA* 108:9788-9791.

Nebel A, Landau-Tasserion E, Filon D (2002) Genetic evidence for the expansion of Arabian tribes into the Southern Levant and North Africa. *Am J Hum Genet* 70:1594–1596.

Nogueiro I, Manco L, Gomes V, Amorim A, Gusmão L (2010) Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities. *Am J Phys Anthropol* 141:373-381.

Rootsi, S., Magri, C., Kivisild, T., Benuzzi, G., Help, H., Bermisheva, M., Kutuev, I., Barac, L., Pericic, M., Balanovsky, O., et al. (2004). Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.* 75, 128–137.

Saraiva JH (1992) *História concisa de Portugal*. Sintra: Publicações Europa-América Press. 375 p.

Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., et al. (2000). The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290, 1155–1159.

Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ, Zhivotovsky LA, King R, Torroni A, Cavalli-Sforza LL, Underhill PA, Santachiara-Benerecetti AS. (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74:1023–1034.

Stieglitz RR (1990) The geopolitics of the Phoenician Littoral in the Early Iron Age. *Bull Am Schools Orient Res* 279:9–12.

Tavares AA (2004) A costa portuguesa há dois mil anos. In Tavares AA, Tavares MJF, Cardoso JL, *Evolução, Geohistórica do Litoral Português e Fenómenos Correlativos: Geologia, História, Arqueologia e Climatologia*, pp. 427-440.

Zalloua PA, Platt DE, El Sibai M, Khalife J, Makhoul N, Haber M, Xue Y, Izaabel H, Bosch E, Adams SM, Arroyo E, López-Parra AM, Aler M, Picornell A, Ramon M, Jobling MA, Comas D, Bertranpetit J, Wells RS, Tyler-Smith C; Genographic Consortium. (2008) Identifying genetic traces of historical expansions: phoenician footprints in the Mediterranean. *Am J Hum Genet* 83:633–642.

Full address of the author/s:

Research Centre for Anthropology and Health (CIAS)
Department of Life Sciences
University of Coimbra
Rua Arco da Traição
3030-056 Coimbra (Portugal)

Table 1 – Most common Y-SNP haplogroups and frequencies found in Coimbra district.

Population	n	H	M35	M78	M81	M123	M213	M201	M170	M267	M172	M9	M22	M269	M70
			E1b1b1*	E1b1b1a	E1b1b1b	E1b1b1c	F*	G	I	J1	J2	K*	L	R1b1b2	T
Littoral Coimbra	97	0.630	2.1	4.1	3.1	1.0	1.0	2.1	5.2	8.3	8.3	4.1	0.0	59.8	1.0
Inland Coimbra	46	0.615	0.0	0.0	6.5	6.5	0.0	10.8	4.3	4.3	6.5	0.0	0.0	60.9	0.0
Total Coimbra	143	0.625	1.4	2.8	4.2	2.8	0.7	4.9	4.9	7.0	7.7	2.8	0.0	60.1	0.7

Table 2 – Y chromosome STR haplotypes in 32 unrelated individuals from Coimbra district within haplogroups with Middle East origin.

Haplogroup	Sample ID	Population	DYS alleles						
			19	389I	389b	390	391	392	393
J1	1	iCoimbra	14	13	16	25	10	11	12
J1	2	iCoimbra	14	13	18	23	10	11	12
J1	3	Littoral	14	13	17	23	10	11	13
J1	4	Littoral	14	13	16	25	10	11	12
J1	5	Littoral	14	13	16	25	10	11	12
J1	6	Littoral	14	13	16	24	10	11	12
J1	7	Littoral	14	13	16	25	10	11	12
J1	8	Littoral	14	13	16	25	10	11	12
J1	9	Littoral	14	13	16	25	10	11	12
J1	10	Littoral	14	13	16	25	10	11	12
J2a	11	iCoimbra	15	13	16	23	10	11	12
J2a	12	iCoimbra	14	13	16	24	10	11	12
J2a	13	iCoimbra	14	12	18	23	10	11	12
J2a	14	Littoral	14	13	17	24	10	11	12
J2a	15	Littoral	14	13	16	23	11	11	12
J2a	16	Littoral	14	12	18	23	10	11	12
J2a	17	Littoral	14	13	18	22	10	11	12
J2a	18	Littoral	15	13	16	23	9	11	12
J2b	19	Littoral	14	12	16	24	11	11	12
J2b	20	Littoral	14	12	16	24	11	11	12
J2b	21	Littoral	15	12	16	24	10	11	13
E1b1b1a	22	Littoral	13	13	18	24	10	11	12
E1b1b1a	23	Littoral	13	13	17	25	10	11	12
E1b1b1a	24	Littoral	13	13	18	23	10	11	13
E1b1b1a	25	Littoral	13	13	17	24	10	11	13
G	26	iCoimbra	16	12	16	21	11	11	14
G	27	iCoimbra	15	12	17	21	10	11	16
G	28	iCoimbra	15	12	18	21	10	11	15
G	29	iCoimbra	16	12	17	23	10	11	13
G	30	iCoimbra	15	12	17	22	10	11	14
G	31	Littoral	15	12	16	23	10	11	13
G	32	Littoral	15	12	17	22	10	11	13

Population codes are as follows:

Littoral: littoral region of Coimbra district.

iCoimbra: inland region of Coimbra district.

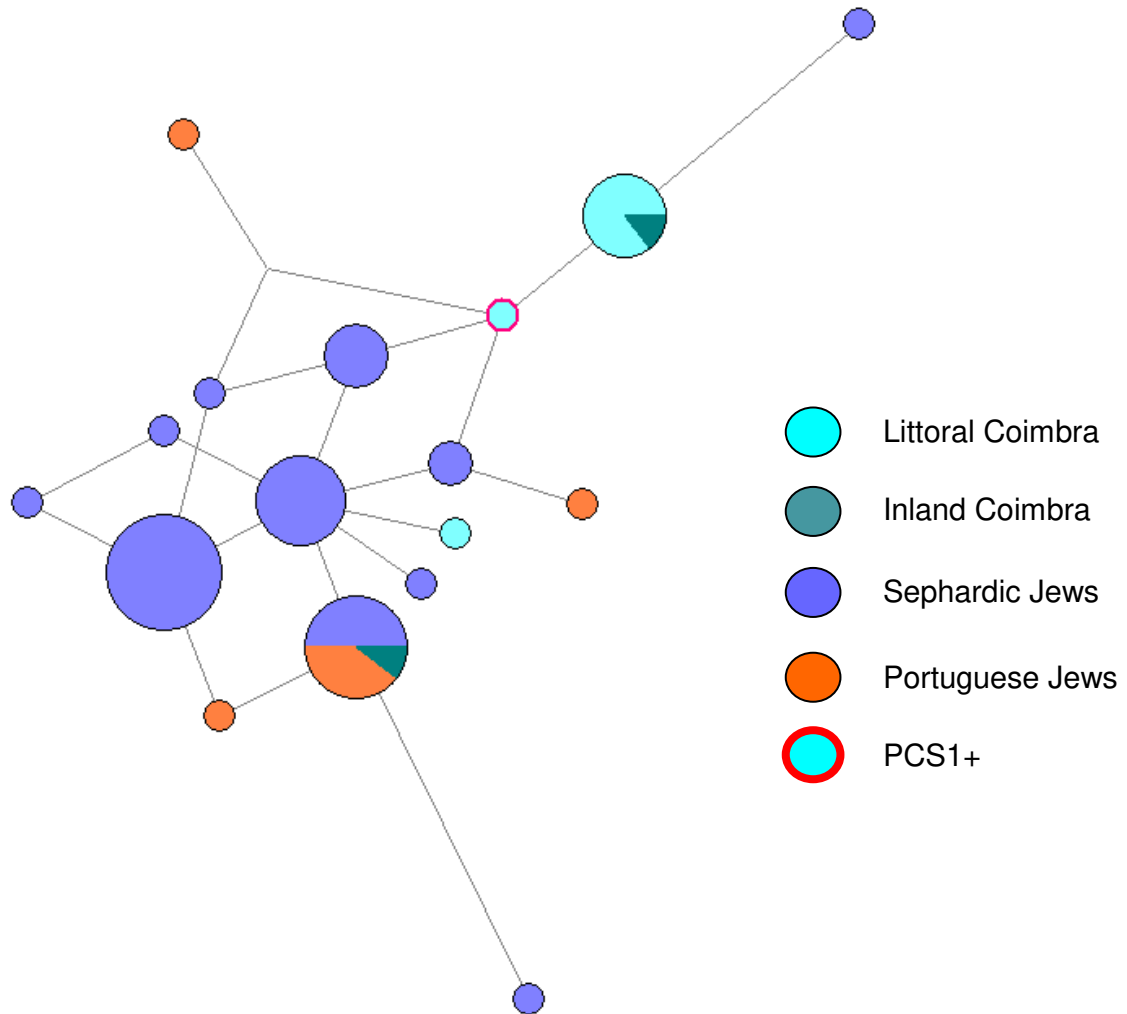


Figure 1 – Median-joining network of Y-SNP haplotypes within J1-M304 haplogroup observed in the study population of Coimbra district and in Portuguese and Sephardic Jews. The size of the circle is approximately proportional to the number of chromosomes observed. Population codes are as follows: PCS1+, Phoenician Colonization Signal.