



R

| REGION AND POPULATION | FREQUENCY OF HAPLOGROUP (%) | | | | | | | | | | | |
|-----------------------------------|-----------------------------|---------|--------|--------------------|----------------|---------------|----------------|----------------|-------|---------|-------|------|
| | E(xE3b) | E-M215* | E-M35* | E-M78 ^a | E-M78 α | E-M78 β | E-M78 γ | E-M78 δ | E-M81 | E-M123* | E-M34 | E-V6 |
| Near East: | | | | | | | | | | | | |
| Sephardi Turkish | 19 | ... | ... | ... | ... | ... | ... | ... | 5.3 | ... | 5.3 | ... |
| Istanbul Turkish | 35 | 2.9 | ... | 8.6 | 2.9 | ... | ... | 5.7 | 5.7 | ... | 2.9 | ... |
| Southwestern Turkish ^b | 40 | ... | ... | 2.5 | 2.5 | ... | ... | ... | 2.5 | ... | 2.5 | ... |
| Northeastern Turkish ^b | 41 | ... | ... | ... | ... | ... | ... | ... | 2.4 | ... | ... | ... |
| Central Anatolian ^b | 61 | ... | ... | 6.6 | 4.9 | ... | ... | 1.6 | ... | ... | 3.3 | ... |
| Southeastern Turkish ^b | 24 | ... | ... | 4.2 | 4.2 | ... | ... | ... | ... | ... | 4.2 | ... |
| Erzurum Turkish ^b | 25 | ... | ... | 4.0 | ... | ... | ... | 4.0 | ... | ... | 8.0 | ... |
| Turkish Cypriots ^b | 46 | 4.3 | ... | 13.0 | 10.9 | ... | ... | 2.2 | 8.7 | ... | 2.2 | ... |
| Bedouins ^b | 28 | 3.6 | ... | 3.6 | ... | ... | ... | 3.6 | 3.6 | ... | 7.1 | ... |
| Druze Arabs ^b | 28 | ... | ... | 10.7 | ... | ... | ... | 10.7 | ... | ... | 3.6 | ... |
| Palestinians ^b | 29 | 10.3 | ... | 10.3 | 3.4 | ... | ... | 6.9 | ... | ... | 3.4 | ... |
| United Arab Emirate ^b | 41 | 7.3 | ... | 2.4 | ... | ... | ... | 2.4 | ... | ... | 4.9 | ... |
| Omanite ^b | 13 | ... | ... | 7.7 | ... | ... | ... | 7.7 | ... | ... | 7.7 | ... |
| Caucasus: | | | | | | | | | | | | |
| Azeri ^b | 97 | ... | ... | 2.1 | ... | ... | ... | ... | ... | ... | 2.1 | ... |
| Adygei ^c | 18 | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Pakistani ^c | 176 | 0.6 | ... | 1.1 | ... | ... | ... | 1.1 | ... | ... | ... | ... |
| Eastern Asians ^{b,c} | 245 | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Oceanians ^c | 21 | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Native Americans ^c | 43 | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

^a E-M78 frequency includes chromosomes belonging to clusters α - δ and 14 additional chromosomes (12 chromosomes excluded from the four clusters in fig. 2B and two Azeri chromosomes for which complete microsatellite data were not available).

^b This sample (or a subset of it) was previously typed for a subset of the markers here analyzed (Scozzari et al. 1997, 2001; Malaspina et al. 2000, 2001; Cruciani et al. 2002).

^c Sample (or a subset of it) from the Human Genome Diversity Project/CEPH DNA panel (Cann et al. 2002).

^d This sample includes 16 DNA samples from the Human Genome Diversity Project/CEPH DNA panel and 26 previously reported samples (Scozzari et al. 2001).

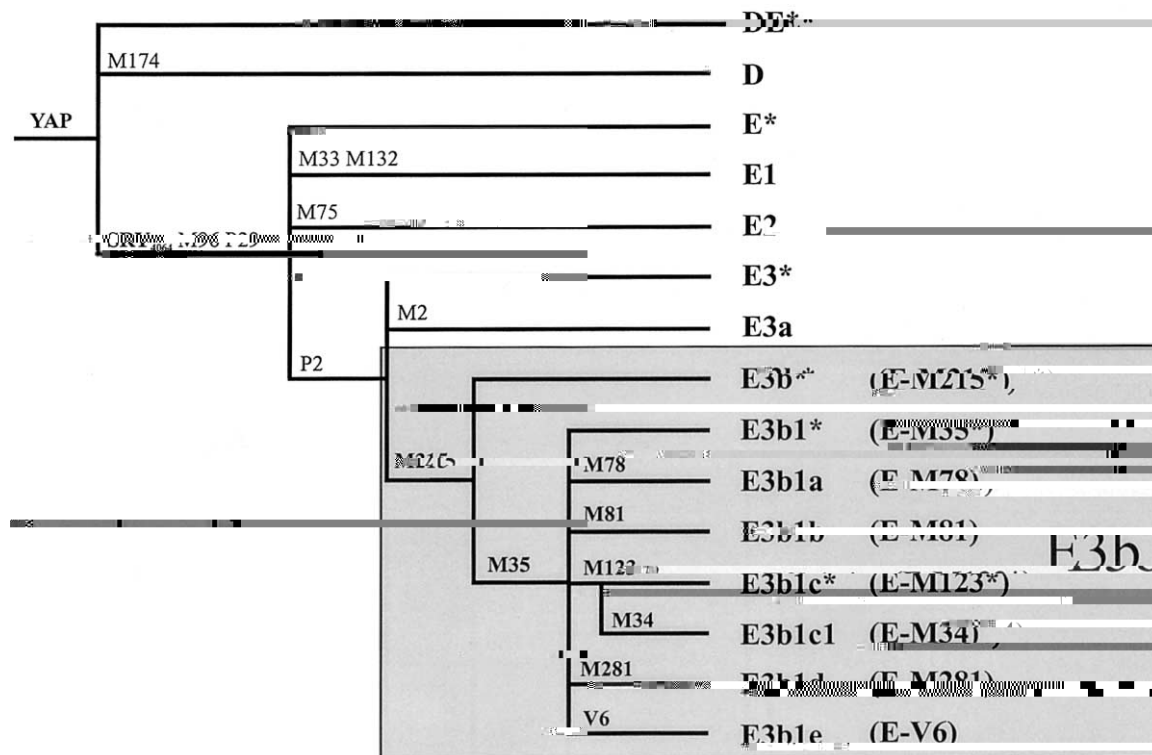
24/23 pattern and its one-step neighbors. When compared with the other clusters in the network, it displays marked starlike features, with three central haplotypes accounting for 26% of the entire cluster. This cluster is very common in the Balkans (with frequencies of 20%–32%), and its frequencies decline toward western (7.0% in continental Italy, 7.4% in Sicily, 1.1% in Sardinia, 4.3% in Corsica, 3.0% in France, and 2.2% in Iberia) and northeastern (2.6%) Europe. In the Near East, this cluster is essentially limited to Turkey (3.4%). The relatively high frequency of DYS413 24/23 haplogroup E chromosomes in Greece (A.N., unpublished data) suggests that cluster α of the E-M78 haplogroup is common in the Aegean area, too.

Cluster β , characterized by the DYS413 23/21 pattern and the rare 10-repeat allele at DYS439, is common in northwestern Africa (14.0%), representing 80% of E-M78 chromosomes in that area. Outside this region, E-M78 β was observed only in five European subjects.

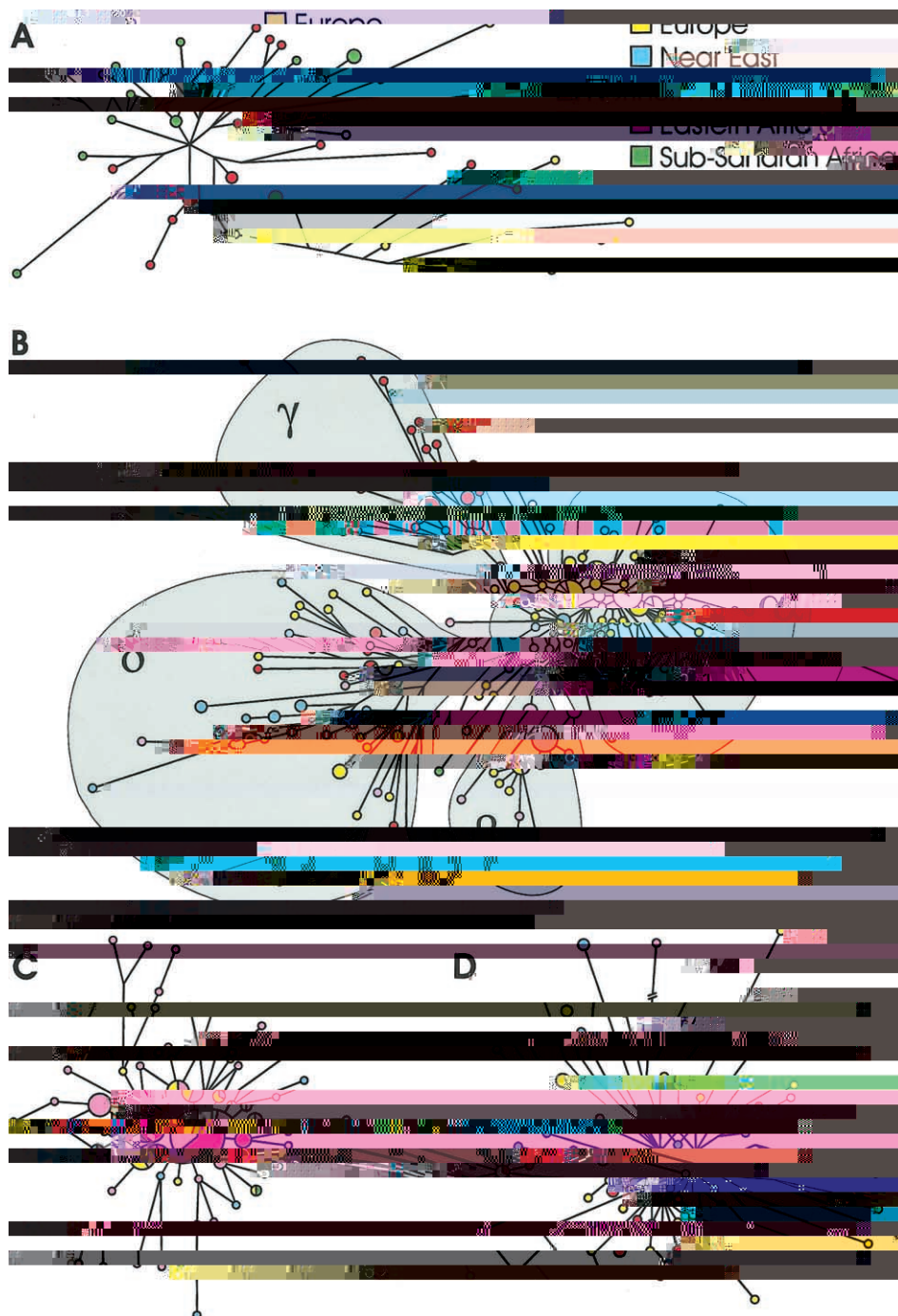
All of the chromosomes in cluster γ (fig. 2B) are identified by the rare short 11-repeat allele at the DYS19 STR locus. We did not find this allele in >2,000 Y(xE-M78) chromosomes analyzed (present study; R.S., unpublished data), and it is reported in only 9 of 13,447 subjects analyzed for this marker in the European Y-STR reference database (Y-STR Haplotype Reference Database Web site). The cluster E-M78 γ was found in eastern Africa at an average frequency of 17.7%, with

the highest frequencies in the three Cushitic-speaking groups: the Borana from Kenya (71.4%), the Oromo from Ethiopia (32.0%), and the Somali (52.2%). Outside of eastern Africa, it was found only in two subjects from Egypt (3.6%) and in one Arab from Morocco.

The fourth cluster (cluster δ in fig. 2B



Phylogenetic tree of haplogroup E3b. Markers typed in this study are in boldface letters. Haplogroups are designated according to the Y Chromosome Consortium (2002) and Jobling and Tyler-Smith (2003), by subclade and also by mutation (



Microsatellite networks of E3b haplogroups. A, E-M35*. B, E-M78. C, E-M81. D, E-M34. Reduced-median and median-joining procedures (Bandelt et al. 1995, 1999) were applied sequentially. A haplogroup-specific weight proportional to the reciprocal of microsatellite variance was used in the construction of the networks. The E-M78 unweighted network (not shown) gave the same quadripartite structure. Unassigned chromosomes 8e

language. Outside of Africa, E-M81 has been observed in all the six Iberian populations surveyed, with frequencies in the range of 1.6%–4.0% in northern Portuguese, southern Spaniards, Asturians, and Basques; 12.2% in southern Portuguese; and 41.1% in the Paisiegos from Cantabria. It has been suggested (Bosch et al. 2001) that recent gene flow may have brought E3b chromosomes from northwestern Africa into Iberia, as a consequence of the Islamic occupation of the peninsula, and that such gene flow left only a minor contribution to the current Iberian Y-chromosome pool. The relatively young TMRCA of 5.6 ky (95% CI 4.6–6.3 ky) that we estimated for haplogroup E-M81 and the lack of differentiation between European and African haplotypes in the network of E-M81 (fig. 2C) support the hypothesis of recent gene flow between northwestern Africa and Iberia. In this context, our data refine the conclusions of Bosch et al. (2001) in two ways. First, not all of the E3b chromosomes in Iberia can be regarded as a signature of African gene flow into the peninsula: in our data set, 8 of 15 E-M78 chromosomes belong to cluster α , denoting gene flow from mainland Europe (see above). Second, and more importantly, the degree of the African contribution is highly variable across different Iberian populations: the proportion of haplogroup E chromosomes of African origin (E[xE3b], E-M35*, and E-M81) was

chromosomes from eastern Africa into and out of Africa, as well as the introduction of the E-M34 subclade into Africa from the Near East. Later events involved short-range migrations within Africa (E-M78 γ and E-V6) and from northern Africa into Europe (E-M81 and E-M78 β), as well as an important range expansion from the Balkans to western and southern-central Europe (E-M78 α). This latter expansion was the main contributor to the present distribution of E3b chromosomes in Europe.

We would like to express our gratitude to all blood donors for their helpful collaboration, which made this study possible. We gratefully acknowledge Emanuele Guida, Jadwiga Jaruzelska, Kenneth K. Kidd, Judith R. Kidd, Damian Labuda, Jean-Paul Moisan, Valentino Romano, Laurent Varesi, Richard Vilems, and the National Laboratory for the Genetics of Israeli Populations for DNA samples. We also thank two anonymous reviewers for their helpful comments. This research received support from Grandi Progetti Ateneo Università di Roma "La Sapienza" (to R.S.) and the Italian Ministry of the University (Progetti Ricerca Interesse Nazionale 2002 and 2003) (to R.S., A.N., and A.T.).

The URL for data presented herein is as follows:

Y-STR Haplotype Reference Database, <http://www.ystr.org/>

-
- Ayub Q, Mohyuddin A, Qamar R, Mazhar K, Zerjal T, Mehdi SQ, Tyler-Smith C (2000) Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information. *Nucleic Acids Res* 28:e8
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mitochondrial portraits of human populations using median networks. *Genetics* 141:743–753
- 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 north-western Africa and the Iberian Peninsula. *Am J Hum Genet* 68:1019–1029
- Cann HM, de Toma C, Cazes L, Legrand M-F, Morel V, Piouffre L, Bodmer J, et al (2002) A human genome diversity cell line panel. *Science* 296:261–262
- Cinnioglu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA (2004) Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 114:127–148
- Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scoz-

