

## THE IDENTIFICATION AND DATING OF THE Y CHROMOSOME OF AN AMERICAN ADAM<sup>1</sup>

*Néstor O. Bianchi\**

We analyzed the allelic polymorphisms in seven Y-specific microsatellite loci and a Y-specific aliphoid system with 27 variants ("hI-XXVII). A total of 53 Y chromosomes carrying the DYS199T allele and belonging to Amerindian (51) and Na-Dene (2) linguistic groups were studied. The information gathered allowed us to identify the ancestral founder haplotype (0A) and to recognize 7 derived haplogroups diverging from 0A by 1-7 mutational steps. The 0A haplotype was the most frequent and had the following allele constitution: DYS199T; "hII; DYS19/13; DYS389A/10; DYS389b/27; DYS390/24; DYS391/10; DYS392/14; DYS393/13 (microsatellite alleles are indicated as number of repeats). All native Americans had the DYS199T allele and the "hII form. Since there are no indications of recurrency for the DYS199C>T transition, we concluded that all DYS199T haplotypes were derived from a single individual in which the C>T mutation occurred for the first time; we call this individual a "New World Adam". We analyzed the Y-specific microsatellite mutation rate in 1743 father-son transmissions and we pooled our data with equivalent information in literature to obtain an average rate of 0.0018. We could estimate that the 0A haplotype has an average age of 33,750 years (minimum 20,250 and maximum 88,050 years). DYS199T allele is found in 85-90% of Amerindian chromosomes indicating that 0A haplotype is the most prevalent or perhaps the only founder paternal lineage of New World aborigines.

Anthropological, archaeological, linguistic, odontological and genetic tools have been used to reconstruct the history of the peopling of America. As a result of this multidisciplinary approach it is generally accepted that the first settlers of America came from Southwest Asia at the time of the last glaciation by way of a Bering land bridge connecting both continents. Moreover, it has been proposed that the colonization of America took place in three

successive chronological events giving rise to Amerindian, Na-Dene and Aleut-Eskimo linguistic groups respectively (Greenberg et al. 1986; Szathmari 1993).

Mitochondrial DNA is a molecule well suited for evolutionary studies due to its maternal mode of inheritance, lack of recombination and high frequency of polymorphisms. Yet, some of the hypotheses based on the interpretation of mt-haplotypes are conflictive. Mitochondrial analysis has been used to support a multiwave founder colonization of America (Torroni et al. 1993), while, on the other hand mt-DNA markers have been also identified as support for monophyletic colonization from Asia (Forster et al. 1996; Bonatto and Salzano 1997). Archaeological studies seem to indicate an antiquity of 11,000 to 12,000 years for the first settlements in Beringia and the New World. Different laboratories working with mt-polymorphisms, however, have proposed a time range of 14,000 to 55,000 YBP for this event (Horai et al. 1993; Torroni et al. 1994; Forster et al. 1997; Bonatto and Salzano 1997). Founder maternal Amerindian lineages were initially estimated to be four (Torroni et al. 1993). Now, it is assumed that there are no less than 10-13, although there is no agreement on the molecular typification of some of these founder haplogroups (Bailliet et al. 1994; Bianchi et al. 1997; Merriwether et al. 1995; Forster et al. 1996, 1997). This controversial information could perhaps be reinterpreted with the use of additional and complementary polymorphic DNA systems. In this regard, Y chromosome specific regions seem most promising.

The male-specific segment of the Y chromosome in mammals has no homologous counterpart, does not recombine, has all its genes in linkage disequilibrium and is paternally transmitted. In this regard, it is the male equivalent of mtDNA. Moreover, Y-specific genes are haploid while mt-genes are polyploid and the mutation rate of Y-DNA is much lower than that of mtDNA.

\* Instituto Multidisciplinario de Biología Celular IMBICE, Calle 526 e/ 19 y 11 CC 403, La Plata, Argentina.  
E-mail: imbice@satlink.com.

Recibido: diciembre 1998. Aceptado: diciembre 2000.

The association of two or more DNA markers defines a haplotype. Mitochondrial haplotypes are known to correlate with the ethnic origin of a population and this also seems to be the case for Y chromosome haplotypes (Bianchi et al. 1997; Pena et al. 1995; Sans et al. 1998; Underhill et al. 1996, 1997). A case in point is the native American. We reported elsewhere that the "hII form of the Y-specific alphoid satellite is associated with the allele A of the DYS19 microsatellite corresponding to a native American-specific haplotype that has not been so far detected in any other geographic population (Pena et al. 1995). In 1996, Underhill et al. found that Y chromosomes of Amerindians, Na-Dene and Eskimo-Aleut exhibited the association of the DYS19A allele with a C>T transition at base position 181 of the DYS199 locus. Bianchi et al. (1997) recently, showing linkage disequilibrium of "hII, DYS199T and DYS19A markers, defined the American aborigine Y chromosome with greater accuracy. The aim of this report is to use additional polyallelic Y-specific markers in order to reach a deeper understanding the origin and evolution of Amerindian chromosomes.

### Samples Analyzed

We analyzed a total of 53 Y chromosomes belonging to American aborigines and having DYS199T allele; 51 samples belonged to Amerindian and 2 to Na-Dene linguistic groups. In the populations analyzed, the number of individuals per population, languages and geographic origin of donors are indicated in Table 1. The samples included in this report were obtained from DNA banks of the IMBICE, La Plata, and Florida International University, Miami.

We also tested 40 certified paternal lineages provided by the CEPH ("Centre d' Etude du Polymorphismes Humains", Paris, France) comprising a total of 249 father-son events. The identification of the CEPH families included in this report is given in Table 3.

### Y-Specific Markers

We analyzed one biallelic and 7 polyallelic systems. The biallelic marker is a C>T transition in the 181 bp position of the locus DYS199 (Underhill et al. 1996). Polyallelic markers are the "hI-XVII) (Bianchi et al. 1997; Santos et al. 1996) and the following microsatellites DYS19 (tetranucleotide, 10 alleles), DYS389a (tetranucleotide, 7 alleles), DYS389b (tetranucleotide, 9 alleles), DYS390 (tetranucleotide, 10 alleles), DYS391 (tetranucleotide, 6 alleles), DYS392 (tetranucleotide, 8 alleles), DYS393 (trinucleotide, 6 alleles) (Kayser et al. 1997; de Knijff et al. 1997). The methods used for testing the above polymorphic markers are previously described (Kayser et al. 1997; Santos et al. 1993, 1996; Underhill et al. 1996).

In this report the two DYS199 alleles are identified as C or T; the ?h alleles are designated with Roman numbers; microsatellites are identified by the number of repeats as used in the reference table of Kayser et al. (1997) and de Knijff et al. (1997).

### Allelic Frequencies

Table 2 shows the allelic frequencies for each one of the polymorphic systems analyzed in the 40 CEPH paternal lineages and in 53 native American Y chromosomes. All CEPH samples had the

Table 1. Populations analysed

| Population | Abbrev. | Samples No. | Linguistic Group | Geographic Region    |
|------------|---------|-------------|------------------|----------------------|
| Navajo     | N       | 2           | Athabaskan       | New Mexico, USA      |
| Sioux      | S       | 2           | Siouan           | Minnesota, USA       |
| Maya       | My      | 3           | Yucatec          | Yucatan, Mexico      |
| Chimila    | Chi     | 4           | Chibcha          | Colombia             |
| Lengua     | L       | 3           | Moscoy           | South Paraguay       |
| Ayoreo     | A       | 4           | Zamuco           | South Paraguay       |
| Wichí      | W       | 12          | Mataco-Mataguayo | Salta, Argentina     |
| Toba       | T       | 4           | Mataco-Mataguayo | Salta, Argentina     |
| Chorote    | C       | 3           | Mataco-Mataguayo | Salta, Argentina     |
| Mapuche    | M       | 1           | Mapudugum        | Rio Negro, Argentina |
| Tehuelche  | Te      | 5           | AoniKen          | Chubut, Argentina    |
| Jujuy      | J       | 10          |                  | Jujuy, Argentina     |

Table 2. Allelic frequencies\*

|         | CEPH         | Amer.        |
|---------|--------------|--------------|
| DYS199  |              |              |
| C       | I            | -            |
| T       | -            | I            |
| ah I    | 0.050        | -            |
| II      | <b>0.500</b> | I            |
| III     | 0.300        | -            |
| IV      | 0.075        | -            |
| V       | 0.025        | -            |
| IX      | 0.025        | -            |
| XII     | 0.025        | -            |
| DYS19   |              |              |
| 13      | 0.025        | <b>0.830</b> |
| 14      | <b>0.550</b> | 0.170        |
| 15      | 0.300        | -            |
| 16      | 0.100        | -            |
| 17      | 0.025        | -            |
| DYS389a |              |              |
| 9       | 0.325        | 0.038        |
| 10      | <b>0.400</b> | <b>0.736</b> |
| 11      | 0.275        | 0.227        |
| DYS389b |              |              |
| 25      | 0.025        | -            |
| 26      | <b>0.325</b> | 0.132        |
| 27      | <b>0.325</b> | <b>0.377</b> |
| 28      | 0.150        | 0.321        |
| 29      | 0.100        | 0.170        |
| 30      | 0.075        | -            |
| DYS390  |              |              |
| 21      | 0.050        | -            |
| 22      | 0.200        | 0.075        |
| 23      | 0.200        | 0.302        |
| 24      | <b>0.375</b> | <b>0.415</b> |
| 25      | 0.175        | 0.208        |
| DYS391  |              |              |
| 9       | -            | 0.094        |
| 10      | <b>0.650</b> | <b>0.887</b> |
| 11      | 0.325        | 0.019        |
| 13      | 0.025        | -            |
| DYS392  |              |              |
| 11      | <b>0.525</b> | -            |
| 12      | 0.075        | 0.019        |
| 13      | 0.325        | 0.151        |
| 14      | 0.050        | <b>0.490</b> |
| 15      | 0.025        | 0.302        |
| 16      | -            | 0.038        |
| DYS393  |              |              |
| 12      | 0.100        | -            |
| 13      | <b>0.550</b> | <b>0.567</b> |
| 14      | 0.275        | 0.189        |
| 15      | 0.075        | -            |

\* Allele composition in each individual is given in Tables III and IV. Bold figures indicate most frequent allele.

DYS199C allele. Moreover, since the presence of DYS199T was one of the selection criteria to classify the sample as American, all chromosomes in this group exhibited that allele. All American samples were "h II, while CEPH chromosomes exhibited

seven different forms of "h, with "h II being the most frequent (Table 2). CEPH and American Ys showed the same predominant allele in DYS389a, DYS390, DYS391 and DYS393 loci and different predominant alleles in DYS19, DYS389b and DYS392 microsatellites (Table 2; details of allele distribution in each Y chromosome of our samples can be obtained from Tables 4 and 5).

The allelic frequencies found in the 40 CEPH lineages are similar to those reported by other authors in European and other human geographic populations (Deka et al. 1996; de Knijff et al. 1997; Kayser et al. 1997; Roewer et al. 1996). The predominance of DYS19/13 and DYS392/14 alleles in our series of American aborigines confirms previous reports from Santos et al. (1996 a, b), Bianchi et al. (1997), Kayser et al. (1997), and de Knijff et al. (1997).

### Mutation Rates

We assessed the mutation rate in the seven microsatellite loci of the 249 Y chromosomes contained in the 40 CEPH paternal lineages which is equivalent to a total of 1743 generation events. In two families (21 and 66) we found a mutation in a F2 male. However, all the male offspring from these two apparently mutated males had the same allele as the F1 ancestor and all other males in the lineage. It was therefore concluded that the two mutations observed were the result of the lymphoblast transformation used to immortalize CEPH lymphocytes (Weber and Wong 1993). Thus, the mutation rate was 0 with a 95% confidence interval limit of 0.0025. Two other direct estimations of mutation rates for Y-specific microsatellites are reported in the literature. Heyer et al. (1997) found three mutations in nine Y-specific microsatellite loci (comprising the seven loci analyzed in this report) and in 213 independent meiotic events, which in combination with the loci tested represent a total of 1917 generations tested. Moreover, Kayser et al. (1997) found two DYS19 slippage mutation events in 626 father-son pairs. If we pool data from this report with data from Heyer et al. (1997) and Kayser et al. (1997), the mutation rate obtained is 0.0012 with 95% confidence interval limits of 0.00046-0.0028.

As no mutation of ah forms was found in the 40 CEPH lineages, the mutation rate of this system should be lower than 0.004.

### Y-Haplotypes

Tables 3 and 4 detail the marker associations giving rise to Y-specific haplotypes. In the 40 CEPH lineages we found 38 haplotypes resulting from the existence of two pairs of lineages sharing the same haplotype (lineages 1350/13294 and 13291/13292, Table 3). Native Americans showed 40 haplotypes out of 53 individuals due to several Y chromosomes having the same haplotype (haplotype 0A, 4 cases; haplotypes 2g and 2h, 3 cases each; haplotype 1a, 1c, 1d, 2d, 3e, 3f, and 5b, 2 cases each; Table 5). The estimation of ge-

netic diversity for CEPH and American Y-specific markers is shown in Table 5. Average diversities ( $H$ ) (Nei 1986,1987) for CEPH and native American samples are 0.565 and 0.371, respectively.

By measuring the number of mutations separating the two most distant haplotypes, we can also estimate intragroup haplotype variability. For microsatellites, we accept the stepwise model of mutations (Otha and Kimura 1973); accordingly, a change in two repeats in the same locus is counted as two mutations. For the  $r$ -h system, we follow the pathway of mutations detailed in Santos et al.

Table 3. CEPH Haplotypes

| Haplotypes |          |       |         |         |        |        |        |        |            |
|------------|----------|-------|---------|---------|--------|--------|--------|--------|------------|
| DYS199     | $\alpha$ | DYS19 | DYS389a | DYS389b | DYS390 | DYS391 | DYS392 | DYS393 | Lineage    |
| C          | I        | 14    | 11      | 27      | 22     | 10     | 11     | 12     | 37         |
| C          | I        | 15    | 10      | 30      | 24     | 10     | 11     | 13     | 1375       |
| C          | II       | 13    | 10      | 27      | 24     | 11     | 13     | 13     | 1345       |
| C          | II       | 14    | 9       | 25      | 24     | 10     | 13     | 13     | 1346       |
| C          | II       | 14    | 9       | 26      | 25     | 10     | 13     | 14     | 1331       |
| C          | II       | 14    | 9       | 27      | 25     | 11     | 13     | 14     | 1377       |
| C          | II       | 14    | 10      | 26      | 24     | 10     | 13     | 13     | 1344       |
| C          | II       | 14    | 10      | 26      | 24     | 11     | 11     | 13     | 1333       |
| C          | II       | 14    | 10      | 26      | 24     | 11     | 13     | 13     | 02         |
| C          | II       | 14    | 10      | 26      | 24     | 13     | 14     | 14     | 1408       |
| C          | II       | 14    | 10      | 28      | 24     | 10     | 13     | 13     | 21         |
| C          | II       | 14    | 10      | 27      | 24     | 11     | 13     | 13     | 104        |
| C          | II       | 14    | 10      | 28      | 24     | 11     | 14     | 13     | 17         |
| C          | II       | 14    | 10      | 27      | 25     | 11     | 13     | 13     | 1418       |
| C          | II       | 14    | 11      | 27      | 23     | 11     | 13     | 13     | 35         |
| C          | II       | 14    | 11      | 27      | 24     | 10     | 11     | 13     | 884        |
| C          | II       | 14    | 11      | 28      | 23     | 10     | 13     | 13     | 1341       |
| C          | II       | 14    | 11      | 28      | 25     | 11     | 11     | 13     | 1416       |
| C          | II       | 15    | 9       | 27      | 21     | 10     | 11     | 14     | 13293      |
| C          | II       | 15    | 10      | 26      | 24     | 11     | 13     | 13     | 1334       |
| C          | II       | 16    | 11      | 27      | 23     | 10     | 11     | 13     | 1332       |
| C          | II       | 16    | 11      | 29      | 25     | 11     | 11     | 13     | 1349       |
| C          | III      | 14    | 9       | 26      | 22     | 11     | 11     | 13     | 1413       |
| C          | III      | 14    | 9       | 26      | 23     | 10     | 11     | 13     | 1350.13294 |
| C          | III      | 14    | 10      | 28      | 22     | 10     | 11     | 13     | 1362       |
| C          | III      | 15    | 9       | 26      | 22     | 10     | 11     | 13     | 1340       |
| C          | III      | 15    | 9       | 26      | 24     | 10     | 11     | 12     | 1424       |
| C          | III      | 17    | 9       | 26      | 25     | 11     | 11     | 14     | 12         |
| C          | III      | 15    | 10      | 29      | 24     | 10     | 11     | 12     | 1421       |
| C          | III      | 15    | 11      | 30      | 23     | 10     | 12     | 15     | 1347       |
| C          | III      | 15    | 11      | 29      | 23     | 10     | 15     | 14     | 1420       |
| C          | III      | 16    | 10      | 26      | 22     | 10     | 12     | 14     | 23         |
| C          | III      | 16    | 10      | 28      | 23     | 10     | 13     | 15     | 28         |
| C          | IV       | 15    | 9       | 27      | 22     | 10     | 11     | 14     | 3291.13292 |
| C          | IV       | 15    | 11      | 29      | 22     | 10     | 11     | 14     | 45         |
| C          | V        | 14    | 9       | 27      | 24     | 10     | 11     | 12     | 66         |
| C          | IX       | 15    | 10      | 27      | 21     | 10     | 12     | 15     | 102        |
| C          | XII      | 14    | 11      | 30      | 25     | 10     | 11     | 14     | 1423       |

(1996). A pairwise comparison of CEPH haplotypes shows that 102 and 1349 are the pair of most distant haplotypes. For microsatellites, the distance between these two lineages is 12 mutational steps. Moreover, haplotype 102 has the hIX form and haplotype 1349 the "hII form; as there is no direct conversion between II and IX, both forms have probably evolved independently from a V form via two deletions for the V>II conversion and two duplications for the V>IX conversion (Santos et al. 1996). Thus, the total number of changes between the 102 and 1349 CEPH lineages is 16. On the other hand, 0A and 7a, the two most distant native American lineages, are separated by 7 microsatellite allelic shiftings.

### The Y chromosome of a New World Adam

Thus far, the DYS199T allele has been only found in American populations belonging to Amerindian, Na-Dene or Aleut-Eskimo linguistic groups (Underhill et al. 1996). Since it is generally accepted that these populations derive from a common pool of ancestral Asiatic-Beringian populations and since there are no indications of recurrence for the DYS199C>T transition (Underhill et al. 1997), it seems reasonable to assume that all Y chromosomes exhibiting this allele derive from a single ancestor who carried the mutation for the first time. We call this individual the "New World Adam"; the 53 Y chromosomes of Table 5 derive from him. In this regard, the DYS199T allele is equivalent to the Y-specific Alu insert which is known to have occurred in a single individual in Africa and from which all extant YAP+ Y chromosomes found worldwide derive (Hammer 1995).

It is usually accepted that in a group of phylogenetically related individuals, the most frequent molecular markers represent shared-ancestral characters which are due to retention of traits found in the common ancestor (Stewart 1993). Thus, we attempted to reconstruct the Y haplotype of the New World Adam by combining the predominant alleles observed in each one of the seven loci analysed (Table 2). The haplotype obtained was 0A, which in fact was the most frequent (7.4%) of haplotypes in Table 4. All other haplotypes in that table derived from 0A and could be sorted in 7 haplogroups diverging from 0A by 1- to 7 mutational steps. The

letters following the numbers in the designation column of Table 4 identify the haplotypes within each haplogroup.

By using the Y microsatellite mutation rate of 0.0012 we could calculate the age of the ancestral haplotype. The probability of observing a microsatellite allele shifting in the 7 loci analyzed was given by the binomial distribution:  $P_{(k)} = \binom{n}{k} P^k (1-P)^{n-k}$  where n is the loci number and k the number of mutations to occur (in our case k=1); the figure obtained was 0.0084. The quotient of 7 (mutational distance between 0A and 7a) by P (0.0084) gave 833, that is the number of generations expected to produce the most divergent haplogroup. Assuming an average generation time of 27 years (Underhill et al. 1996), the antiquity of 0A can be estimated to be 22,770 years with a minimum and a maximum of 13,500- and 58,700 years. There is however a caveat to be taken into account regarding the model used to estimate the age of DYS199T chromosomes. Once a microsatellite has taken an upstream (gain of one repeat) or downstream (loss of one repeat) mutational pathway, a reversal in the allele shifting direction will go unnoticed as the haplotype will move from its haplogroup to a haplogroup less distant from 0A. If we assume that microsatellite mutation rates are the same for repeat gains and losses, we may tentatively conclude that approximately 50% of mutations will go undetected. Therefore, the age indicated above can, in a first approximation, be increased in 50% to reach the figure of 33,750 years with a minimum and maximum of 20,250 and 88,050 respectively. These figures are about the same as estimates of entry into America based on classical genetic markers (Cavalli-Sforza et al. 1994), mitochondrial DNA (Bonatto and Salzano 1997) and archaeological remains (Dillehay and Collins 1998).

It is worth mentioning here that the finding of "hII in all native American haplotypes indicates a mutation rate lower than 1/33,750 or  $3 \times 10^{-4}$  for this form.

### Is the 0A the only New World Ancestral Paternal Lineage?

Several authors have proposed a multiwave early colonization of America. Neves and Pucciarelli (1991) compared the cranial morphology of early South American remains with worldwide human remains of Late Pleistocene and Holocene

and reached the conclusion that the Americas were occupied by undifferentiated premongoloid populations before the migration and spreading of differentiated Mongoloid colonizers. Roosevelt et al. (1996) analyzed Paleoindian campsites in the Brazilian Amazon and found evidence of a cultural tradition contemporary with but different from that of the Clovis Paleoindian culture of North

America. The conclusion drawn from these results is that big-game hunters were probably not the only source of migration into America. Moreover, linguistic, dental and nuclear DNA markers have been interpreted as supporting a three wave migration into America giving rise to Amerindians, Na-Denes and Aleut-Eskimos (Greenberg et al. 1986).

Table 4. Native American Haplotypes

| Haplotypes |    |           |           |           |           |           |           |           |       |             |                             |
|------------|----|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-------|-------------|-----------------------------|
| D199       | αh | D19       | D389a     | D389b     | D390      | D391      | D392      | D393      | Cases | Popul.<br>* | Hapl.<br>Notatio<br>n<br>** |
| T          | II | 13        | 10        | 27        | 24        | 10        | 14        | 13        | 4     | W,T         | 0A                          |
| T          | II | <b>14</b> | 10        | 27        | 24        | 10        | 14        | 13        | 2     | W,Te        | 1a                          |
| T          | II | 13        | 10        | 27        | 24        | 10        | <b>15</b> | 13        | 1     | W           | 1b                          |
| T          | II | 13        | 10        | <b>28</b> | 24        | 10        | 14        | 13        | 2     | T,A         | 1c                          |
| T          | II | 13        | 10        | 27        | <b>23</b> | 10        | 14        | 13        | 2     | L           | 1d                          |
| T          | II | 13        | 10        | <b>26</b> | 24        | 10        | 14        | 13        | 1     | My          | 1e                          |
| T          | II | 13        | 10        | 27        | 24        | 10        | <b>13</b> | 13        | 1     | W           | 1f                          |
| T          | II | <b>14</b> | 10        | <b>26</b> | 24        | 10        | 14        | 13        | 1     | M           | 2a                          |
| T          | II | 13        | 10        | <b>26</b> | 24        | 10        | <b>15</b> | 13        | 1     | W           | 2b                          |
| T          | II | 13        | 10        | 27        | <b>25</b> | 10        | <b>15</b> | 13        | 1     | W           | 2c                          |
| T          | II | 13        | 10        | <b>28</b> | 24        | 10        | <b>15</b> | 13        | 2     | T,C         | 2d                          |
| T          | II | 13        | <b>11</b> | 27        | <b>23</b> | 10        | 14        | 13        | 1     | N           | 2e                          |
| T          | II | 13        | 10        | 27        | <b>23</b> | <b>9</b>  | 14        | 13        | 1     | S           | 2f                          |
| T          | II | 13        | 10        | <b>29</b> | 24        | 10        | 14        | 13        | 3     | C,A         | 2g                          |
| T          | II | 13        | 10        | 27        | <b>25</b> | <b>9</b>  | 14        | 13        | 3     | Te,Chi      | 2h                          |
| T          | II | 13        | <b>11</b> | 27        | <b>25</b> | 10        | 14        | 13        | 1     | J           | 2i                          |
| T          | II | 13        | 10        | <b>29</b> | 24        | 10        | <b>15</b> | 13        | 1     | W           | 3a                          |
| T          | II | 13        | 10        | <b>28</b> | <b>23</b> | 10        | <b>15</b> | 13        | 1     | C           | 3b                          |
| T          | II | 13        | 10        | <b>28</b> | <b>25</b> | 10        | <b>15</b> | 13        | 1     | W           | 3c                          |
| T          | II | <b>14</b> | 10        | <b>28</b> | 24        | 10        | <b>15</b> | 13        | 1     | T           | 3d                          |
| T          | II | 13        | <b>11</b> | <b>28</b> | <b>25</b> | 10        | 14        | 13        | 2     | W,J         | 3e                          |
| T          | II | 13        | 10        | <b>28</b> | <b>25</b> | 10        | <b>13</b> | 13        | 2     | Chi         | 3f                          |
| T          | II | 13        | <b>11</b> | <b>28</b> | <b>23</b> | 10        | 14        | 13        | 1     | S           | 3g                          |
| T          | II | 13        | 10        | 26        | <b>23</b> | 10        | <b>13</b> | <b>14</b> | 1     | My          | 4a                          |
| T          | II | 13        | <b>11</b> | <b>29</b> | <b>25</b> | 10        | 14        | 13        | 1     | W           | 4b                          |
| T          | II | <b>14</b> | 10        | <b>29</b> | 24        | 10        | <b>13</b> | 13        | 1     | W           | 4c                          |
| T          | II | 13        | 10        | <b>26</b> | <b>22</b> | 10        | <b>13</b> | 13        | 1     | L           | 4d                          |
| T          | II | <b>14</b> | 10        | <b>26</b> | <b>23</b> | 10        | <b>15</b> | 13        | 1     | Te          | 4e                          |
| T          | II | 13        | <b>11</b> | 27        | <b>23</b> | 10        | <b>15</b> | <b>14</b> | 1     | J           | 4f                          |
| T          | II | 13        | 10        | <b>29</b> | <b>23</b> | 10        | 14        | <b>14</b> | 1     | J           | 4g                          |
| T          | II | 13        | <b>9</b>  | <b>26</b> | 24        | 10        | <b>16</b> | 13        | 1     | N           | 4h                          |
| T          | II | 13        | <b>9</b>  | <b>28</b> | <b>23</b> | <b>9</b>  | <b>13</b> | 13        | 1     | M           | 5 <sup>a</sup>              |
| T          | II | 13        | <b>11</b> | <b>28</b> | <b>23</b> | 10        | <b>15</b> | <b>14</b> | 2     | J           | 5b                          |
| T          | II | 13        | <b>11</b> | 27        | <b>22</b> | 10        | <b>15</b> | <b>14</b> | 1     | J           | 5c                          |
| T          | II | <b>14</b> | 10        | <b>28</b> | <b>23</b> | 10        | <b>15</b> | <b>14</b> | 1     | J           | 5d                          |
| T          | II | <b>14</b> | 10        | 27        | <b>22</b> | 10        | <b>12</b> | <b>14</b> | 1     | Te          | 6a                          |
| T          | II | <b>14</b> | 10        | <b>29</b> | <b>22</b> | 10        | <b>13</b> | 13        | 1     | A           | 6b                          |
| T          | II | 13        | <b>11</b> | <b>28</b> | <b>23</b> | <b>11</b> | <b>15</b> | <b>14</b> | 1     | J           | 6c                          |
| T          | II | 13        | <b>11</b> | <b>29</b> | <b>23</b> | 10        | <b>16</b> | <b>14</b> | 1     | J           | 7a                          |

Horizontal lines separate the ancestral haplotype (0A) and derived haplogroups. Bold numbers identify mutations.

\* Abbreviation for populations are indicated in Table 1.

\*\* Numbers in this column (1-7) identify haplogroups and mutation distances from 0A. Letters in each haplogroup identify haplotypes.

Recent studies using mtDNA markers to resolve the timing and number of prehistoric migrations into America propose founding times of 20000-25000 YBP (Forster et al. 1996) or between 22,000-55,000 YBP (Bonatto and Salzano 1997). In the out-of-Asia hypothesis, Siberia is considered the geographic region of origin of American populations and Beringia is given the role of a corridor (Forster et al. 1996). In the out-of-Beringia proposal, Beringia is assumed to be the place where American ancestors differentiated before migrating into the New World (Bonatto and Salzano 1997). In spite of the above disagreements, both groups of authors coincide in suggesting that after an early colonization event, the Laurentidae corridor that allowed the passage from Beringia to North America collapsed producing the isolation and sub-

sequent differentiation of Northern and Southern populations, and that Northern populations gave rise to Na-Dene and Aleut-Eskimo linguistic groups.

The DYS199T is found in 85-90% of Amerindians and is also present in Na-Denes and Aleut-Eskimos (Bianchi et al. 1997; Underhill et al. 1996). Accordingly, it is tempting to speculate that all native Americans derive from a single paternal lineage that evolved intracontinentally to give rise to the haplogroups depicted in Table 5. Yet, there are some alternatives to be considered before accepting the single paternal lineage proposal. First, although most of the DYS199C chromosomes detected in Amerindians are the consequence of recent genetic mixture with Europeans and Africans (Bianchi et al. 1997) we cannot rule out some of

Table 5. Genetic Diversity in CEPH and Native American Y Chromosomes\*

| Locus      | Population | Total diversity | Total intra-population | Inter-population diversity | Percentage of inter population diversity (Gst%) | Subpopulation diversity        | Significance (Student test) |
|------------|------------|-----------------|------------------------|----------------------------|-------------------------------------------------|--------------------------------|-----------------------------|
|            |            | (Ht)            | (Hs)                   | (Dst/Dsm)                  |                                                 | (h)                            | (t)                         |
| DYS199     | CEPH Amer. | 0.500           | 0.000                  | 0.500/1.00                 | 100                                             | 0.000<br>0.000                 |                             |
| $\alpha$ h | CEPH Amer. | 0.637           | 0.325                  | 0.312/0.625                | 49                                              | 0.668 ± 0.000<br>0.039 ± 0.000 | 16.66                       |
| DYS19      | CEPH Amer. | 0.659           | 0.438                  | 0.221/0.442                | 33.5                                            | 0.588 ± 0.298<br>0.307 ± 0.047 | 4.82                        |
| DYS389a    | CEPH Amer. | 0.594           | 0.554                  | 0.039/0.079                | 6.6                                             | 0.667 ± 0.014<br>0.453 ± 0.046 | 4.40                        |
| DYS389b    | CEPH Amer. | 0.773           | 0.707                  | 0.026/0.053                | 3.6                                             | 0.717 ± 0.026<br>0.711 ± 0.019 | 0.19                        |
| DYS390     | CEPH Amer. | 0.722           | 0.713                  | 0.009/0.019                | 1.3                                             | 0.746 ± 0.022<br>0.694 ± 0.020 | 1.74                        |
| DYS391     | CEPH Amer. | 0.375           | 0.336                  | 0.039/0.078                | 10.4                                            | 0.477 ± 0.039<br>0.202 ± 0.048 | 4.39                        |
| DYS392     |            | 0.774           | 0.635                  | 0.139/0.278                | 17.9                                            | 0.617 ± 0.037<br>0.666 ± 0.029 | -1.02                       |
| DYS393     |            | 0.478           | 0.459                  | 0.019/0.039                | 4.0                                             | 0.613 ± 0.041<br>0.314 ± 0.040 | 5.18                        |

\* Estimated according to Nei (1986,1987)

them to correspond to a second founder lineage. Second, in the case of Na-Denes and Aleut-Eskimos it is possible to assume that DYS199T is the founder lineage and DYS199C is the consequence of genetic flow, or viceversa; we can surmise that the founding lineage is C and that T indicates Amerindian Y chromosome infiltration into the other two linguistic groups. Thus, we cannot disregard the existence of other founder lineages carrying the DYS199C allele.

The probability of fixation of an advantageous mutation is approximately equal to twice the fitness advantage (Crow and Kimura 1970; Haldane 1937). Furthermore, the effective population size ( $N_e$ ) is inversely related to the probability that copies of a same gene in two randomly chosen genomes are derived from the same ancestral copy (Crow and Kimura 1970). Thus, selective advantage and a very small  $N_e$  are the most cogent explanations for the marked predominance of DYS199T alleles in native Americans. Due to linkage disequilibrium and genetic hitch-hiking (Kaplan et al. 1989; Maynard-Smith and Haig 1974), any advantageous mutation linked to DYS199T would give rise to a rapid expansion of this allele with a subsequent decrease of DYS199C allele. Moreover, the effect of chance resulting from the small  $N_e$  of Y chromosomes (1/4 and 1/3 the  $N_e$ s of autosomes and X chromosomes, respectively) should have been enhanced in native American ancestral populations by the increased male mortality during big-game hunting and warfare, and by the practice of polysy-

ny (Bianchi et al. 1997; Jobling and Tyler-Smith 1995; Santos et al. 1996).

The hypothesis of Neves and Pucciarelli (1991) regarding the replacement of an undifferentiated premongoloid American population by the subsequent migration and expansion of a Mongoloid population entering the New World from Asia/Beringia is an interesting possibility as it introduces new approach in interpreting Y-chromosome and mtDNA results. Cultural instead of genetic advantages may have been the cause of the replacement of preexisting premongoloid populations by Mongoloid newcomers into America. When two mammal populations of the same species compete for an ecologic niche, the gene expansion of the dominating group is usually mediated through matings of dominating males with females from the dominating and subdued populations. If that was the case in America, the high frequency of DYS199T alleles would indicate the expansion of Y chromosomes of the culturally prevalent Mongoloid population. Conversely, DYS199C alleles would be the Y chromosome relic of the earliest premongoloid American settlers, while the wide variability of mtDNA haplotypes would reflect a more balanced mixture of mt-genomes from the two American colonizations.

**Acknowledgments.** This work was supported with grants from CONICET, CIC, ANPCYT and CABBIO of Argentina. We wish to thank C. Horgan (IMBICE, Argentina) for his technical assistance.

### References Cited

- Bailliet, G., F. Rothhammer, F.R. Carnese, C.M. Bravi, N.O. Bianchi  
1994 Founder Mitochondrial Haplotypes in Amerindian Populations. *Am J Hum Genet* 55: 27-33.
- Bianchi, N.O., G. Bailliet, C.M. Bravi, F.R. Carnese, F. Rothhammer, V.I. Martinez-Marignac, S.D.J. Pena  
1997 Origin of Amerindian Y-Chromosomes as Inferred by the Analysis of Six Ppolymorphic Markers. *Am J Phys Anthropol* 102: 79-89.
- Bonato, S.L., F.M. Salzano  
1997 A Single and Early Migration for the Peopling of the Americas Supported by Mitochondrial DNA Sequence Data. *Proc Natl Acad Sci USA* 94: 1866-1971
- Cavalli-Sforza L.L., P. Menozzi, A. Piazza  
1994 *The History and Geography of Human Genes*. Princeton Univ. Press, New Jersey.
- Crow, J.F., M. Kimura  
1970 *An Introduction to Population Genetics Theory*. Harper & Row.
- Deka, R., L. Jin, M.D. Shriver, L. Mei Yu, N. Saha, R. Barrantes, R. Chakraborty, R.E. Ferrell  
1996 Dispersion of Human Y Chromosome Haplotypes Based on Five Microsatellites in Global Populations. *Genome Res* 6: 1177-1184.
- De Knijff, P., M. Kayser, A. Caglia, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog et al.  
1997 Chromosome Y Microsatellites: Population Genetic and Evolutionary Aspects. *Int J Legal Med* 110: 134-140.
- Dillehay, T.D., M.B. Collins  
1988 Early Cultural Evidence from Monte Verde in Chile. *Nature* 332: 150-152.
- Forster, P., R. Harding, A. Torroni, H.J. Bandelt  
1996 Origin and Evolution of Native American mtDNA Variation: a Reappraisal. *Am J Hum Genet* 59: 935-945
- Forster, P., R. Harding, A. Torroni, H.J. Bandelt  
1997 Reply to Bianchi and Bailliet. *Am J Hum Genet* 61: 246-247.
- Greenberg, J.H., C.G. Turner, L.Z. Zegura  
1986 The Settlement of the Americas: a Comparison of the



- Linguistic, Dental and Genetic Evidence. *Curr Anthropol* 27: 477-497.
- Haldane, J.B.S.  
1937 A Mathematical Theory of Natural and Artificial Selection. Part V. Selection and Mutation. *Proc Camb Phil Soc* 28: 838-844.
- Hammer, M.F.  
1995 A Recent Common Ancestry for Human Y Chromosomes. *Nature* 378: 376-378.
- Heyer, E., J. Puymirat, P. Dieltjes, E. Bakker, P. De Knijff  
1997 Estimating Y Chromosome Specific Microsatellite Mutation Frequencies Using Deep Rooting Pedigrees. *Hum Mol Genet* 6: 799-803.
- Horai, S., R. Kondo, Y. Nakagawa-Hattori, S. Hayashi, S. Sonoda, K. Tajima  
1993 Peopling of the Americas, Founded by Four Major Lineages of Mitochondrial DNA. *Mol Biol Evol* 10: 23-47.
- Jobling, M.A., C. Tyler-Smith  
1995 Fathers and Sons: the Y Chromosome and Human Evolution. *Trends Genet* 11: 449-456.
- Kaplan, N.L., R.R. Hudson, C.H. Langley  
1989 The "Hitchhiking Effect" Revisited. *Genetics* 123: 887-889.
- Kayser, M., A. Caglia, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Hidding et al  
1997 Evaluation of Y-Chromosomal STRs: a Multicenter Study. *Int J Legal Med* 110: 125-133.
- Maynard-Smith J., J. Haigh  
1974 The Hitchhiking Effect of a Favorable Gene. *Genet Res* 23: 23-35.
- Merrivether, D.A., F. Rothhammer, F. Ferrell  
1995 Distribution of Four Founding Lineage Haplotypes in Native Americans Suggests a Single Wave of Migration for the New World. *Am J Phys Anthropol* 98: 411-430.
- Nei, M.  
1986 Definition and Stimation of Fixation Indexes. *Evolution* 40: 643-645.
- Nei, M.  
1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Neves, W.A., H.M. Puciarelli  
1991 Morphological Affinities of the First Americans: an Exploratory Analysis Based on Early South American Human Remains. *J Hum Evol* 21: 261-273.
- Otha, T., M Kimura  
1973 A Model of Mutation Appropriate to Estimate the Number of Electrophoretically Detectable Alleles in a Finite Population. *Genet Res* 22: 201-204.
- Pena, S.D.J., F.R. Santos, N.O. Bianchi, C.M. Bravi, F.R. Carnese, F. Rothhammer, T. Gerelsaikhan, B. Munkhtuja, T. Oyunsuren  
1995 A Major Founder Y-Chromosome Haplotype in Amerindians. *Nature Genet* 11: 15-16.
- Roewer, L., M. Kayser, P. Dieltjes, M. Nagy, E. Bakker, M. Krawczak, P. De Knijff  
1996 Analysis of Molecular Variance (AMOVA) of Y-Chromosome-Specific Microsatellites in two Closely Related Human Populations. *Hum Mol Genet* 5: 1029-1033.
- Roosevelt, A.C., M. Lima da Costa, C. Lopes Machado, M. Michab, N. Mercier, H. Valladas, J. Feathers, W. Barnett, M. Imazio da Silveira, A. Henderson et al.  
1996 Paleoindian Cave Dwellers in the Amazon: the Peopling of the Americas. *Science* 272: 373-384.
- Sans, M., T.A. Weiner, S.C. Fracac, M.H. Franco, F.M. Salzano, N. Betancor, I. Alvarez, N.O. Bianchi, R. Chakraborty  
1998 Unequal Contributions of Male and Female Gene Pools from Parental Populations in the African Descendants of the City of Melo, Uruguay. *Am J Phys Anthropol* (in press).
- Santos, F.R., S.D.J. Pena, J.T. Epplen  
1993 Genetic and Population Study of a Y-Linked Tetra-nucleotide Repeat DNA Polymorphism with a Simple Non-isotopic Technique. *Hum Genet* 90: 655-656.
- Santos, F.R., N.O. Bianchi, S.D.J. Pena  
1996 Worldwide Distribution of Human Y-chromosome Haplotypes. *Genome Res* 7: 601-611.
- Santos, F.R., T. Gerelsaikhan, B. Munkhtuja, T. Oyunsuren, J.E. Epplen, S.D.J. Pena  
1996a Geographic Differences in the Allelic Frequencies of the Human Y-linked Tetra-nucleotide Polymorphism DYS19. *Hum Genet* 97: 309-313.
- Santos, F.R., L. Rodriguez-Delfin, S.D.J. Pena, J. Moore, K.M. Weiss  
1996b North and South Amerindians may have the Same Major Founder Y chromosome. *Am J Hum Genet* 58: 1369-1370.
- Stewart, C.B.  
1993 The Powers and Pitfalls of Parsimony. *Nature* 361: 603-607.
- Szathmary, E.J.E.  
1993 Genetics of Aboriginal North Americans. *Evolut. Anthropol.* 1: 202-220.
- Torroni, A., R.I. Sukernik, T.G. Schurr, Y.B. Starikovskaya, M.F. Cabell, M.H. Crawford, A.G. Comuzzie, D.C. Wallace  
1993 MtDNA Variation of Aboriginal Siberians Reveals Distinct Genetic Affinities with Native Americans. *Am J Hum Genet* 53: 591-608.
- Torroni, A., J.V. Neel, R. Barrantes, T.G. Schurr, D.C. Wallace  
1994 Mitochondrial DNA "clock" for the Amerinds and its Implications for Timing their Entry into North America. *Proc Natl Acad Sci USA* 91: 1158-1162.
- Underhill, P.A., L. Jin, R. Zemans, P.J. Oefner, L.L. Cavalli-Sforza  
1996 A pre-Columbian Y Chromosome-specific Transition and its Implications for Human Evolutionary History. *Proc Natl Acad Sci USA* 93: 196-200.
- Underhill, P.A., L. Jin, A.A. Lin, S. Qasim Mehdi, T. Jenkins, D. Vollrath, R.W. Davis, L.L. Cavalli-Sforza, P.J. Oefner  
1997 Detection of Numerous Y Chromosome Biallelic Polymorphisms by Denaturing High-performance Liquid Chromatography. *Genome Res* 7: 996-1005.
- Weber, J.L., C. Wong  
1993 Mutation of Human Short Tandem Repeats. *Hum Mol Genet* 2: 1123-1128.

## Notas

- 1 Conferencia Inaugural del III Congreso Mundial de Estudios sobre Momias. leída el 18 de Mayo de 1998.