Discordant Phylogeographic Patterns Between the *Y* Chromosome and Mitochondrial DNA in the House Mouse: Selection on the *Y* Chromosome?

Stéphane Boissinot and Pierre Boursot

Laboratoire Génome et Populations (CNRS UPR 9060), Université Montpellier II, F-34095 Montpellier, France

Manuscript received October 18, 1996
Accepted for publication March 17, 1997

ABSTRACT

We have compared patterns of geographic variation and molecular divergence of mitochondrial DNA (mtDNA) and *Y* chromosome over the range of the different subspecies of *Mus musculus*. MtDNA was typed for 305 nucleotides in the control region, the *Y* chromosome for 834 base pairs (bp) in *Zfy* introns and 242 bp in *Sry*, a *Zfy2* 18-bp deletion, and two microsatellites. Apparent discrepancies exist between the distributions of the lineages of mtDNA and of the two major *Y*-chromosome lineages thus defined: some subspecies share the same mtDNA lineage but have different *Y*-chromosome lineages or vice versa. One microsatellite reveals a geographically clustered variation inside the distribution of each *Y*-chromosome lineage, showing that new *Y*-chromosome variants can rapidly spread locally. The two major *Y*-chromosome lineages have a divergence time only about one fourth of that between mtDNA lineages. Although this recent coalescence of the *Y* chromosomes between subspecies could partly be due to a lower ancestral polymorphism of the *Y* chromosome, it suggests that secondary introgression after the radiation of the subspecies might have occurred. There is evidence that the differentiation of the *Y*-chromosome lineages contributes to partial reproductive isolation between subspecies, and patterns of molecular evolution suggest that selection has played a role in the rapid spread across subspecies.

THE importance of sex chromosomes in the development of postmating isolation between species has long been recognized and is implicit in the expression of HALDANE's rule (HALDANE 1922). Because of their partial or total hemizygosity, sex chromosomes are expected to evolve faster than the autosomes under the influence of positive Darwinian selection (CHARLESWORTH *et al.* 1987) and to be relatively more often involved than the autosomes in interspecific incompatibilities, provided most mutations causing the divergence between species are recessive (ORR 1995; TURELLI and ORR 1995). Furthermore, in the case of the *Y* chromosome, mutation pressure alone is expected to lead to a rapid evolution because this chromosome is transmitted through the male germ line, which undergoes more cell divisions than the female line. A rapid evolution of genes on the *Y* chromosome has in fact been observed in both rodents and primates (MIYATA *et al.* 1987; SHIMMIN *et al.* 1993; CHANG *et al.* 1994; CHANG and LI 1995). Interspecific comparisons of the sequences of the sex-determining gene *Sry/SRY* in mice (TUCKER and LUNDRIGAN 1993) and primates (WHITFIELD *et al.* 1993) have not only shown high rates of evolution but also exceptionally high ratios of nonsynonymous over synonymous nucleotide substitutions, which either suggests that it evolves by positive Darwinian selection or that selective constraints are relaxed (TUCKER and LUNDRIGAN 1995). If positive Darwinian selection is important, it should lead to reduced intraspecific polymorphism of the nonrecombining part of the *Y* chromosome because of hitchhiking provoked by favorable mutations (MAYNARD SMITH and HAIGH 1974; KAPLAN *et al.* 1989). Both in mice (NACHMAN and AQUADRO 1994) and humans (HAMMER 1995; WHITFIELD *et al.* 1995) the *Y* chromosome was found to be less polymorphic than mitochondrial DNA (mtDNA), although in neither case was this trend strong enough to provide significant evidence for selective sweeps in the HKA test for selective neutrality (HUDSON *et al.* 1987).

Another consideration is that because the two sex chromosomes have different modes of transmission, they are potential targets for the expression of conflicts of interest between paternal and maternal genes. It has been proposed that HALDANE's rule results from the presence of a coevolved system of segregation distorters and responders on the sex chromosomes in each species, which has evolved in response to such a conflict (FRANK 1991a,b; HURST and POMIANKOWSKI 1991; POMIANKOWSKI and HURST 1993). Although this view has been strongly criticized for various reasons, including the lack of experimental evidence for segregation distortion of sex chromosomes (JOHNSON and WU 1992; CHARLESWORTH *et al.* 1993; COYNE and ORR 1993), if such a mechanism operated, it would also be expected to lead to rapid evolution of sex chromosomes in re-

Corresponding author: Pierre Boursot, Laboratoire Génome et Populations, CC 063, Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 5, France. E-mail: boursot@univ-montp2.fr

sponse to the expression of the conflict. Sexual selection is another mechanism that could promote rapid divergence of genetic systems, although not necessarily those on sex chromosomes. However since *Y*-chromosome genes are involved in male reproductive functions, some of them could participate in the expression of sexual selection (WU *et al.* 1996).

Until now, however, experimental tests of these predictions have been rare concerning the *Y* chromosome, partly because the molecular organization and gene content of this chromosome were not well understood. The house mouse, *Mus musculus*, offers an interesting natural situation to study *Y*-chromosome evolution in relation to differentiation and reproductive isolation. This species appears to have undergone a geographic radiation that started in the Northern Indian subcontinent less than a million years ago. From there it is thought to have gradually colonized the periphery of the continent, where three well-differentiated subspecies are currently found: *M. m. domesticus* in Western Europe and around the Mediterranean Basin, *M. m. musculus* from Central Europe to Northern China, and *M. m. castaneus* in Southeast Asia (for reviews see BOURSOT *et al.* 1993 and SAGE *et al.* 1993; BOURSOT *et al.* 1996; DIN *et al.* 1996). The peripheral subspecies of *M. musculus* can still hybridize in natural conditions wherever their distribution areas meet, so that a clear-cut hybrid zone exists between *M. m. domesticus* and *M. m. musculus* in Europe, while in Southeast Asia *M. m. musculus* and *M. m. castaneus* intermix apparently to a greater extent (see BOURSOT *et al.* 1993 for review).

Previous studies of mtDNA differentiation by RFLP had shown that, at this scale, three major lineages can be recognized, one that is typical of *M. m. domesticus*, a second that is typical of *M. m. musculus*, and a third diversified one that is found in *M. m. castaneus*, in the Indian subcontinent and in the Middle East (YONEKAWA *et al.* 1988; BOURSOT *et al.* 1996). The study of a repeated *Y*-specific sequence by RFLP revealed two major lineages of *Y* chromosomes, one of which was found in *M. m. domesticus* and the other in *M. m. musculus* and in the old inbred laboratory strains (BISHOP *et al.* 1985). Analysis of the hybrid zone between *M. m. domesticus* and *M. m. musculus* in Europe has revealed the absence of introgression of the sex chromosomes, suggesting a major role of these chromosomes in the partial reproductive isolation between these subspecies (VANLERBERGHE *et al.* 1986; TUCKER *et al.* 1992b; DOD *et al.* 1993). The *M. m. musculus*-type *Y* chromosome is also fixed in *M. m. castaneus* (BOURSOT *et al.* 1989). Studies based on other repeated sequences have all confirmed the existence of these two major lineages and found concordant geographic and strain distributions of the two lineages (NISHIOKA and LAMOTHE 1986; NISHIOKA 1987; PLATT and DEWEY 1987; NAGAMINE *et al.* 1992; TUCKER *et al.* 1992a). The *M. m. musculus Y* chromosome has an 18 base pair (bp) deletion in the *Zfy2* gene, which distinguishes it from the *M. m. domesticus* type (NAGAMINE *et al.* 1992). Furthermore, an *Sry* RFLP (NAGAMINE *et al.* 1992; NAGAMINE *et al.* 1994b), as well as RFLP for a repeated sequence (TUCKER *et al.* 1992a), identifies a variant *M. m. musculus Y* chromosome that is only found in Japan and on the continent near Japan, as well as in most old inbred laboratory strains.

The reconstruction of the histories of the different parts of the genome during the radiation process of *M. musculus* and their comparison should provide insight into the relative roles that historical contingencies and selective factors have played in determining present patterns of differentiation and reproductive isolation between the subspecies. In this article we do this by comparing the phylogeographic patterns of mtDNA and the *Y* chromosome at a species-wide scale. We extend the geographical survey of mtDNA and the *Y* chromosome by including new samples from the center of the continent (Northern India, the Middle East, and the Caucasus), and we compare the age of the coalescence of the *Y* and mtDNA lineages. We show that the present distribution of the two major *Y*-chromosome lineages results from their rapid spread from a recent common ancestor, across populations already differentiated for mtDNA. We discuss the evidence for the influence of selection on the spread of *Y*-chromosome lineages and the relationship between *Y*-chromosome differentiation and reproductive isolation between subspecies.

MATERIALS AND METHODS

Animals: All the animals studied here were caught in the wild or were from wild-derived strains maintained in Montpellier. Details will be given later.

DNA extraction: Most of the analyses were done on total DNA from either fresh, frozen, or ethanol-preserved tissues but some of the mtDNA samples analyzed used mtDNA-enriched preparations from fresh tissues as described in BOURSOT *et al.* (1987).

MtDNA control region amplification and sequencing: The amplification primers used flank the control region. The forward amplification primer was a 21-mer with its 3′ end at position 15392 on the L strand of the complete sequence in BIBB *et al.* (1981) and was 5′ phosphorylated. The reverse primer was H21 (18-mer). The amplification conditions were 30 cycles with the following steps: 92° (1 min), 60° (1 min), and 71° (2 min). The following steps were as described in GRAVEN *et al.* (1995) and included digestion of one strand with exonuclease and sequencing using primer L15392.

***Zfy2* deletion:** Male DNAs were typed by PCR for the presence/absence of an 18-bp deletion in the last exon of *Zfy2*, described in NAGAMINE *et al.* (1992). The primers used were 5′CATTAAGACAGAAAAGACCACCG3′ and 5′GTGAGGAAATTTCTTCCTGTGG3′. The PCR conditions were 2 mM $MgCl_2$ and 30 cycles with 1 min at 94°, 1 min at 60°, and 1 min at 72°. The primers amplify equally *Zfy1* and *Zfy2*, so that when the *Zfy2* deletion is absent, a single 202-bp fragment is obtained. When the deletion is present, an extra 184-bp fragment is observed.

***Sry* sequences:** We amplified and sequenced a part of the *Sry* gene between positions 8491 and 8732 (positions as in GUBBAY *et al.* 1992) using the protocols described in LUNDRIGAN and TUCKER (1994).

***Zfy* intron sequences:** Primers that amplify two introns were designed using the intron/exon map of *Zfy1* kindly provided by JÉROME COLLIGNON. Intron II lies between positions 918 and 919 of the complementary DNA sequence of *Zfy1* (EMBL Database accession no. X14382) and intron IV lies between positions 1181 and 1182. Both introns are ~2 kilobases (kb) long in *Zfy1*. The introns we studied are outside the 5′ region of the gene where alternative splicing has been described (KOOPMAN *et al.* 1989; ZAMBROWICZ *et al.* 1994). The numbering of introns that we use is arbitrary. Amplification primers were designed from the sequence of the *Zfy1* flanking exons. For intron II they were as follows: Forward, 5′TAAATTGGATGAAGCATCTCCA3′ and Reverse, 5′AATGTTGACTCTGGGAACACG3′, and for intron IV they were as follows: Forward, 5′AGCTTCCTACCTATTGCATGG3′ and Reverse, 5′AGAGAGCACTGGCAGTGACA3′. Amplifications were performed in 3 mM $MgCl_2$ with 30 cycles of 1 min at 94°, 1.5 min at either 62 or 65°, and 1.5 min at 72°. At 62° the amplification gave two fragments (~2 and 1 kb long) with both the intron II and intron IV primers. The *Zfy1* introns, of predicted size 2 kb, could be selectively amplified by raising the temperature to 65°, because the reverse primers for both introns were chosen in exonic regions where sequence differences between the two genes exist. The two bands obtained with each set of primers using the less stringent conditions were each recovered from the gel and sequenced. Intron II was sequenced with its reverse primer, and intron IV was sequenced with its forward primer. In the case of intron II, the nucleotide differences that exist between *Zfy1* and *Zfy2* in the coding region between the sequencing primer and the intron allowed us to confirm that the 2-kb fragment came from *Zfy1* and the 1-kb fragment came from *Zfy2*. We supposed that the same was true for intron IV, for which no such diagnostic exonic site existed.

***Y*-specific microsatellites:** One of the microsatellites we studied is located in the 5′ end of the second intron of *Zfy2*, the sequence of which we have determined (not shown). It is a TTTTG pentanucleotide repeat. The primers used to amplify it were 5′AGCTGACTCAGAAGTGGATGA3′ and 5′CCAGGGCTATACAGAGGAACT3′. Amplification conditions were 30 cycles of 92° (1 min), 58° (1 min), and 71° (1 min). The products were analyzed by 5% PAGE and visualized by ethidium bromide staining. The other microsatellite is an imperfect repeat with stretches of di- and trinucleotides that is located in the large inverted repeat of the sex-determining region sequenced by GUBBAY *et al.* (1992). The primers used were 5′CCTCATTGATCCTTTGGCAT3′ and 5′TCGAAAGCTCTCTTGCACAA3′. They are expected to amplify the fragments at positions 5768–5985 and 11694–11911 of the sequence determined by GUBBAY *et al.* (1992). Analysis was as above except that the gels were 6% PAGE.

Data analyses: The MEGA program package (KUMAR *et al.* 1993) was used to calculate nucleotide divergences and their standard errors between the *Zfy* intron sequences. The computer package PHYLIP 3.5c (FELSENSTEIN 1993) was used to analyze the mtDNA sequence data. Each of 500 bootstrapped data sets (generated with program SEQBOOT) were used to calculate pairwise distances between haplotypes (using DNADIST with the maximum likelihood option and a transition:transversion ratio of 10). The 500 distance matrices obtained were used to generate 500 trees with the Neighbor-Joining method (program NEIGHBOR), from which the majority rule consensus tree was determined (program CONSENSE). The topology of this consensus tree was used to calculate branch lengths by the maximum likelihood method (program DNAML). The parameters used in the ML model were a transition:transversion ratio of 10; three categories of sites with frequencies 0.6, 0.3, and 0.1; and relative probabilities of mutation of 0.0, 1.0, and 5.0, respectively. Initial parameters for the ML model were chosen from the examination of the patterns of mutation inferred from the results of a parsimony analysis (program DNAPARS). The different parameters were then adjusted successively to maximize the likelihood of the data given the consensus tree obtained as described above. Calculation time limitations prevented simultaneous optimization of the parameters of the model and of the tree topology. Trees were drawn using the programs RETREE and TREEVIEW. Published mtDNA restriction maps (SHE *et al.* 1990) and the computer package RESTSITE (NEI and MILLER 1990) were used to calculate mtDNA divergences and their standard errors by bootstrap resampling of restriction enzymes. FISHER's exact tests on contingency tables were performed using the Markov chain algorithm of GUO and THOMPSON (1992) as implemented in the computer package GENEPOP 1.2 (RAYMOND and ROUSSET 1995).

RESULTS

Mitochondrial data: We sequenced 305 nucleotides in the 5′ end of the control region of mtDNA, corresponding to the region between positions 15443 and 15742 in the complete sequence of BIBB *et al.* (1981), in 131 animals mainly from the Mediterranean basin, the Caucasus, Iran, Pakistan, and India (see Table 1 for a complete list). The sequences were aligned by hand, together with a sequence of *M. spretus* determined by NACHMAN and AQUADRO (1994). The variable sites in the resulting alignment are shown in Figure 1. Among the 91 sites given, seven correspond to sites with no polymorphism but with undetermined sequence for one or a few haplotypes; six are monomorphic in our *M. musculus* sample but different in *M. spretus*; 70 vary by point substitution among *M. musculus*; one varies by both point substitution and single nucleotide insertion/deletion, three by single nucleotide insertion/deletion alone, and four correspond to variations of copy number of a tandemly repeated TA motif (microsatellite, positions 15546–15547 and 15553a–15553Cb). If sequences with undetermined positions are included, altogether 71 different haplotypes can be defined and are described in Figure 1, where they are given numbers. Table 1 shows in which samples the different haplotypes were found. Eleven mice were found to possess a tandem duplication of the region spanning positions 15538–15615 in the sequence of BIBB *et al.* (1981). The haplotypes that have this duplication are indicated in Figure 1 (first column entitled "dupli"). They include haplotype 130 (two mice from Moscow) and haplotypes 131–138 (nine mice from India, Pakistan, and Iran). For Figure 1 only the 5′ copy of the duplication was considered in the alignments with the other haplotypes that do not have the duplication. Figure 2 shows the alignment of the 5′ and 3′ copies of the duplication in the eight haplotypes possessing it. Due to an insertion/deletion, the duplication is 75 bp in haplotype 130 and 76 bp in the others. It can be seen that because of this and character states at two other nucleotide positions

TABLE 1

List of the samples studied for mtDNA control region sequence

Haplotype[a]	n^b	Locality	Country	Lineage[c]	Locality number[d]
84 NE	1	Saasenheim	France	*dom*	1
82 ME	1	Montpellier	France	*dom*	2
83 ME	1	Orcetto	Italy	*dom*	3
81 ME	2	Rabanal del Camino	Spain	*dom*	4
80 ME	1	Sta Colomba de la Vega	Spain	*dom*	4
65 NAF	1	Bembla	Tunisia	*dom*	7
68 NAF	1	Bembla	Tunisia	*dom*	7
75 NAF	8	Bembla	Tunisia	*dom*	7
76 NAF	3	Bembla	Tunisia	*dom*	7
72 NAF	1	Bembla ext.	Tunisia	*dom*	7
75 NAF	1	Bembla ext.	Tunisia	*dom*	7
73 NAF	1	Beni-Hassen	Tunisia	*dom*	7
78 NAF	6	Djemmal	Tunisia	*dom*	7
79 NAF	1	Djemmal	Tunisia	*dom*	7
68 NAF	1	Henzel Ennour	Tunisia	*dom*	7
65 NAF	1	Khnès	Tunisia	*dom*	7
76 NAF	5	Lamta	Tunisia	*dom*	7
76 NAF	1	Maatmer	Tunisia	*dom*	7
67 NAF	5	Monastir	Tunisia	*dom*	7
71 NAF	1	Monastir	Tunisia	*dom*	7
70 NAF	6	Sahline	Tunisia	*dom*	7
69 NAF	2	Sfax	Tunisia	*dom*	7
74 NAF	1	Smiret	Tunisia	*dom*	7
76 NAF	1	Smiret	Tunisia	*dom*	7
77 NAF	1	Smiret	Tunisia	*dom*	7
64 NAF	2	Sousse	Tunisia	*dom*	7
78 NAF	2	Sousse	Tunisia	*dom*	7
66 NAF	1	Zembra	Tunisia	*dom*	7
72 NAF	1	Zeramdine	Tunisia	*dom*	7
88 CA	1	Breti	Georgia	*dom*	9
87 CA	1	Gardabani	Georgia	*dom*	9
86 CA	1	Gori	Georgia	*dom*	9
85 CA	1	Kareli	Georgia	*dom*	9
87 CA	1	Shiraskaya	Georgia	*dom*	9
93 CA/EE	3	Megri	Armenia	*mus*	9
95 CA	1	Megri	Armenia	*mus*	9
94 CA	1	Alazani	Georgia	*mus*	9
96 CA	2	Didich-Shiraki	Georgia	*mus*	9
97 CA	1	Didich-Shiraki	Georgia	*mus*	9
99 CA	4	Gardabani	Georgia	*mus*	9
102 CA	1	Lagodekhi	Georgia	*mus*	9
98 CA	1	Lake Lissi	Georgia	*mus*	9
103 CA	1	Sagandsilé	Georgia	*mus*	9
94 CA	1	Shiraskaya	Georgia	*mus*	9
91 CA	1	Tbilissi	Georgia	*mus*	9
101 CA	2	Abkasia	Georgia	*mus*	10
100 CA	1	Derbent	Russia (Daghestan)	*mus*	11
100 CA	1	Mhachkala	Russia (Daghestan)	*mus*	11
93 CA/EE	1	Riga	Lettonia	*mus*	12
92 EE	1	Riga	Lettonia	*mus*	12
130 EE	2	Moscow	Russia	*mus*	13
119 IR	1	Kakhk	Iran	*mus*	17
118 IR	1	Mashhad	Iran	*mus*	18
129 IR	1	Mashhad	Iran	*mus*	18
137 IR	1	Birdjan	Iran	*ori*	17
111 IR	4	Tehran	Iran	*ori*	19
112 IR	1	Tehran	Iran	*ori*	19
113 IR	1	Tehran	Iran	*ori*	19
114 IR	1	Tehran	Iran	*ori*	19

TABLE 1

Continued

Haplotype[a]	n^b	Locality	Country	Lineage[c]	Locality number[d]
138 IP	1	Gujarkhan	Pakistan	*ori*	20
108 IP	1	Islamabad	Pakistan	*ori*	20
109 IP	1	Islamabad	Pakistan	*ori*	20
110 IP	1	Islamabad	Pakistan	*ori*	20
104 IP	1	Rawalpindi	Pakistan	*ori*	20
105 IP	2	Rawalpindi	Pakistan	*ori*	20
106 IP	1	Rawalpindi	Pakistan	*ori*	20
132 IP	1	Rawalpindi	Pakistan	*ori*	20
133 IP	1	Rawalpindi	Pakistan	*ori*	20
105 IP	1	Tahmasapabad	Pakistan	*ori*	20
107 IP	1	Angah	Pakistan	*ori*	21
122 IP	1	Chaboraha	India	*ori*	22
123 IP	1	Chaboraha	India	*ori*	22
124 IP	1	Chaboraha	India	*ori*	22
110 IP	1	Delhi	India	*ori*	23
116 IP	1	Delhi	India	*ori*	23
117 IP	1	Delhi	India	*ori*	23
125 IP	1	Delhi	India	*ori*	23
127 IP	1	Delhi	India	*ori*	23
131 IP	2	Delhi	India	*ori*	23
132 IP	1	Delhi	India	*ori*	23
135 IP	1	Delhi	India	*ori*	23
136 IP	1	Koorg	India	*ori*	24
120 IP	2	Leh	India	*ori*	25
121 IP	1	Leh	India	*ori*	25
115 IP	1	Masinagudi	India	*ori*	26
128 EO	1	He-Mei	Taiwan	*ori*	27

[a] Haplotypes are defined in Figure 1. They are identified by a number (from 64 to 138) and letters that refer to the geographical region where they were found: CA, Caucasus; EE, Eastern Europe; EO, Extreme Orient (Far East); IP, Indo-Pakistan; IR, Iran; ME, Mediterranean Europe; NAF, North Africa; NE, Northern Europe.

[b] Sample size.

[c] Refers to the three lineages defined in Figure 3.

[d] Refers to the numbers on Figure 4A.

the 5′ and 3′ copies of haplotype 130 are more similar to each other than they are to either the 5′ or the 3′ copies belonging to the other haplotypes. This indicates that two independent duplications have occurred (one in haplotype 130 and one in the others), as confirmed by the phylogenetic analysis of the whole region sequenced, presented below. Duplications similar to that of haplotype 130 were found by PRAGER *et al.* (1996) in 28 *M. m. musculus* mice from various geographic origins from Austria to Abkhasia. A detailed discussion of the structure and possible molecular mechanisms at the origin of this duplication can be found in PRAGER *et al.* (1996).

The phylogenetic analysis of the mtDNA sequences covering part of the control region is presented in Figure 3. The results are concordant with those obtained previously by RFLP (BOURSOT *et al.* 1996), and several lineages can be defined. One is typical of *M. m. domesticus* and, as expected, is found in our samples from Northern Africa and Mediterranean Europe. Another is typical of *M. m. musculus* and is found in our samples from Eastern Europe. The other branches of the tree harbor samples from the Far East (*castaneus*), the Indian subcontinent, and Iran. This group of haplotypes was referred to as the "oriental" lineage in the RFLP analysis of BOURSOT *et al.* (1996), but its monophyly is poorly supported by the present data. The map in Figure 4A summarizes the distribution of the major mtDNA lineages that have been found in this and some previously published studies (BOURSOT *et al.* 1996 and references therein). For the purpose of representation, the lineages other than *domesticus* and *musculus* are given a single symbol in Figure 4.

The separation of the haplotypes with the 75- and 76-bp duplications on the phylogeny shown in Figure 3 (these haplotypes are indicated by filled circles) shows that the two duplications occurred independently. Given the uncertainties of the phylogenetic reconstruction, the fact that one nonduplicated haplotype lies on the same branch as the haplotypes carrying the 76-bp

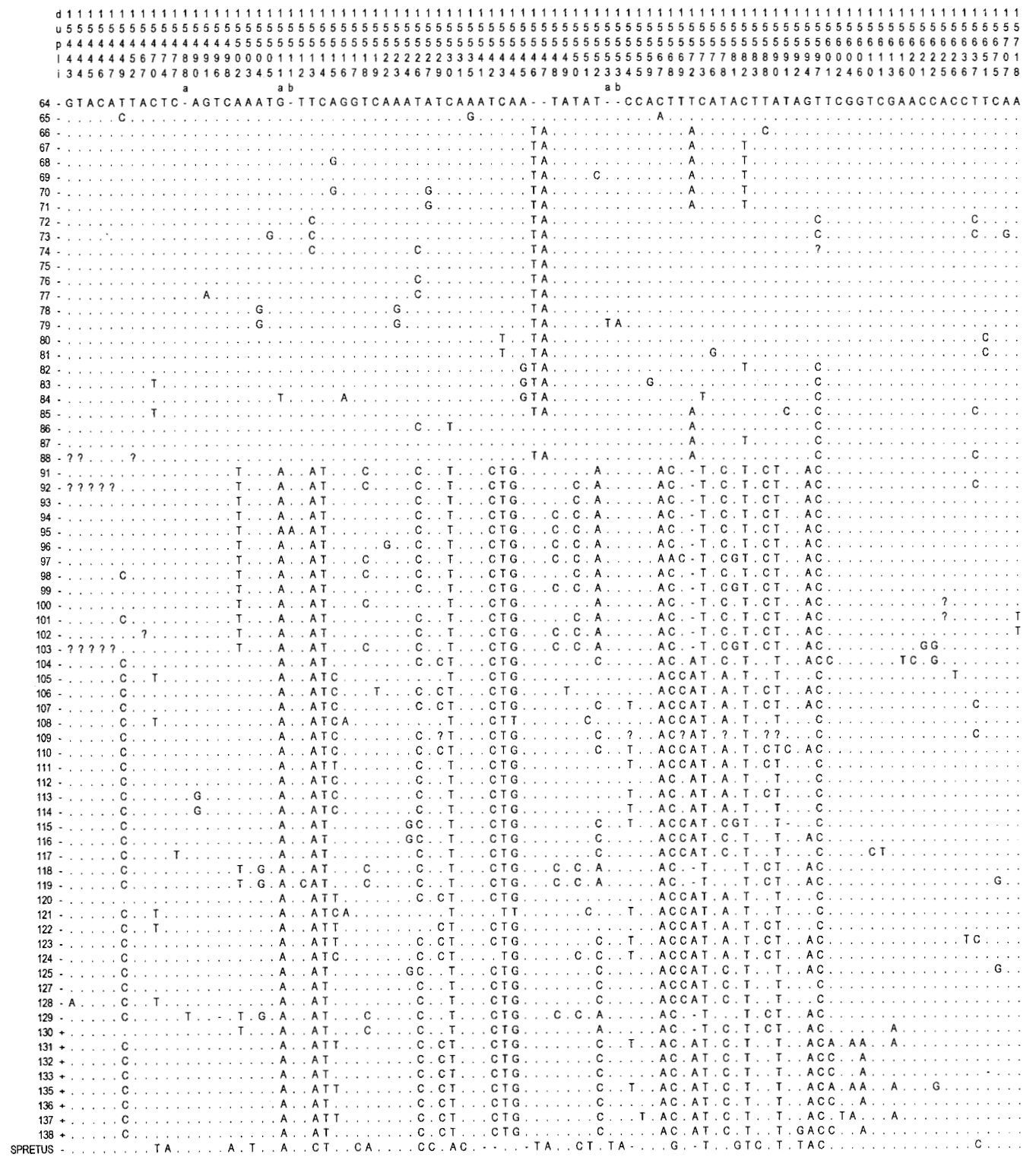

FIGURE 1.—Alignment of the mtDNA control region sequences of the 71 haplotypes found in our *M. musculus* sample, together with one sequence of *M. spretus* (NACHMAN and AQUADRO 1994). Dots indicate identity with the first sequence, question marks indicate undeterminations of the sequence, and dashes indicate deletions. Only variable nucleotide positions are shown and numbered according to the complete sequence of BIBB *et al.* (1981). Insertions compared to this sequence are given the number of the nucleotide 5′ of the insertion, followed by a letter. Alignment in the region of the TA microsatellite (15546–15553b) is arbitrary among *M. musculus* haplotypes and was optimized by comparison with *M. spretus*. The first column (dupli) shows the haplotypes that have the 75–76 bp duplication (+) and those that do not (−).

```
130 5' TAAACTGATA TAAACCATGA ATATTATACT AA-TACATCA AATTAATGCT TTAAAGACAT ATCTGTGTTA TCTAAC
131 5' .......... ..C.T..... .......... ..A..... ........T. .......... .A..A.A... ......
132 5' .......... ..C..... .......... ..A..... ........T. .......... .C....A... ...G..
133 5' .......... ..C..... .......... ..A..... ........T. .......... .C....A... ...G..
135 5' .......... ..C.T..... .......... ..A..... ........T. .......... .A..A.A... ......
136 5' .......... ..C..... .......... ..A..... ........T. .......... .C....A... ...G..
137 5' .......... ..C..T.... .......... ..A..... ........T. .......... ..T.A..... ......
138 5' .......... ..C..... .......... ..A..... ........T. ..G..... .C....A... ...G..
130 3' .......... C......... .......... .......... .......... ..C..... .......... ...G..
131 3' .......... ..C.T..... .......... ..A.....T. ........T. .......... .......... ...G..
132 3' .......... ..C..... .......... ..A..... ........T. .......... .......... ...G..
133 3' .......... ..C..... .......... ..A..... ........T. .......... .......... ...G..
135 3' .......... ..C.T..... .......... ..A.....T. ........T. .......... .......... ...G..
136 3' .......... ..C.T..... .......... ..A..... ........T. .......... .......... ...G..
137 3' .......... ..C.T..... .........C ..G.C..... ........T. .......... .......... ...G..
138 3' .......... C.C.T..... .......... ..A..... ........T. .......... .......... ...G..
```

FIGURE 2.—Alignment of the 5′ and 3′ copies of the 75–76 duplication found in eight mtDNA haplotypes. Haplotype numbers, and symbols are as in Figure 1.

duplication cannot be taken as evidence for two independent duplications among these haplotypes.

***Zfy2* deletion:** In previous reports, NAGAMINE *et al.* (1992, 1994b) showed that the 18-bp deletion in the last exon of *Zfy2* was fixed in *M. m. castaneus* and *M. m. musculus* and was absent from *M. m. domesticus*, as well as from all the individuals from the Indian subcontinent that they analyzed (seven from Delhi, India and one from Lahore, Pakistan). To complete their survey of its distribution we typed 73 additional males, and the results are given in Table 2, together with previous results on 22 males reported in NAGAMINE *et al.* (1992). Our data confirm the previous conclusions: as *M. m. domesticus* mice, the mice from the Indian subcontinent (seven from Pakistan and seven from India) do not have the deletion, and all others do. The latter include mice that were caught in the Caucasus and in northeastern Iran, where we have found admixtures of mtDNA lineages.

***Sry* sequences:** LUNDRIGAN and TUCKER (1994) showed that *M. m. domesticus* could be distinguished from *M. m. musculus* and *M. m. castaneus* by two nonsynonymous nucleotide substitutions at positions 8701 and 8731 of the sex determination gene *Sry*. The old inbred strain sequenced by GUBBAY *et al.* (1992), which, as shown by NAGAMINE *et al.* (1992, 1994a) and TUCKER *et al.* (1992a) harbors a *molossinus Y* chromosome, differs from *musculus* and *castaneus* by one nonsynonymous mutation at position 8491 and one synonymous at position 8711 over the region sequenced by LUNDRIGAN and TUCKER (1994). To extend this survey, we sequenced the same region, in a mouse from Delhi, India, one from Tehran, Iran, and one from Taiwan. A new polymorphism was found at site 8706, at which the mouse from Tehran was different from all others (A instead of C). At sites 8701 and 8731 the mouse from Delhi was identical to *M. m. domesticus*, while the mouse from Tehran was identical to *M. m. musculus* and *M. m. castaneus* (including the mouse from Taiwan that we sequenced). These data show a perfect concordance with those on the *Zfy2* deletion and confirm that the mice from the Indian subcontinent have a *Y* chromosome that is more related to that of *M. m. domesticus* than to that of *M. m. musculus* and *M. m. castaneus*. In Table 2 we have also reported the data on a *Zfy* RFLP of NAGAMINE *et al.* (1992) and those on RFLP for the repeated sequence pY353 of BOURSOT *et al.* (1989). Although the different markers were studied on different sets of animals, the combination of geographic and molecular informations are fully compatible between them and show that two major lineages of *Y* chromosomes exist in the house mouse. One of them is found in *M. m. domesticus* and in the Indian subcontinent, while the other is shared by *M. m. musculus* and *M. m. castaneus*, as summarized in Figure 4B, which gives the geographic distribution of these two *Y*-chromosome lineages.

***Zfy* introns divergence:** To measure the molecular divergence between the two main *Y*-chromosome lineages defined, we sequenced 834 nucleotides from the introns of the *Zfy1* and *Zfy2* genes in a mouse belonging to each *Y*-chromosome lineage. The same region of a *M. spretus Y* chromosome was also sequenced and served as an outgroup. We sequenced 212 nucleotides in the second intron of *Zfy2*, 299 in the fourth intron of *Zfy1*, and 323 in the fourth intron of *Zfy2* as described in MATERIALS AND METHODS. Only 17 variable sites were found (Table 3), revealing one nucleotide difference between *M. m. musculus* and *M. m. domesticus* and 16 and 17 nucleotide differences between *M. spretus* and *M. m. musculus* and *M. m. domesticus*, respectively.

Comparison of *Y*-chromosome and mtDNA divergences: In Table 4 we computed the nucleotide divergences of the *Zfy* introns between these three taxa and compared them with the values obtained for mtDNA. We used the RFLP mtDNA data from SHE *et al.* (1990) rather than the mtDNA sequences of the hypervariable control region determined in the present study to estimate divergence as the high level of homoplasy prevents reliable estimates being made, particularly in comparisons with the outgroup. While the mtDNA divergence found between the subspecies of *M. musculus* is ~40% of that found between these subspecies and *M. spretus*, in the case of the *Y* chromosome it is only 10%. The *Y*-chromosome divergences in Table 4 are remarkably similar to those obtained by NACHMAN and AQUADRO (1994) who sequenced 1277 bp of noncoding DNA 5′ of *Sry* and found 0.27% divergence between *M. m. domesticus* and *M. m. musculus* and 1.9% divergence be-

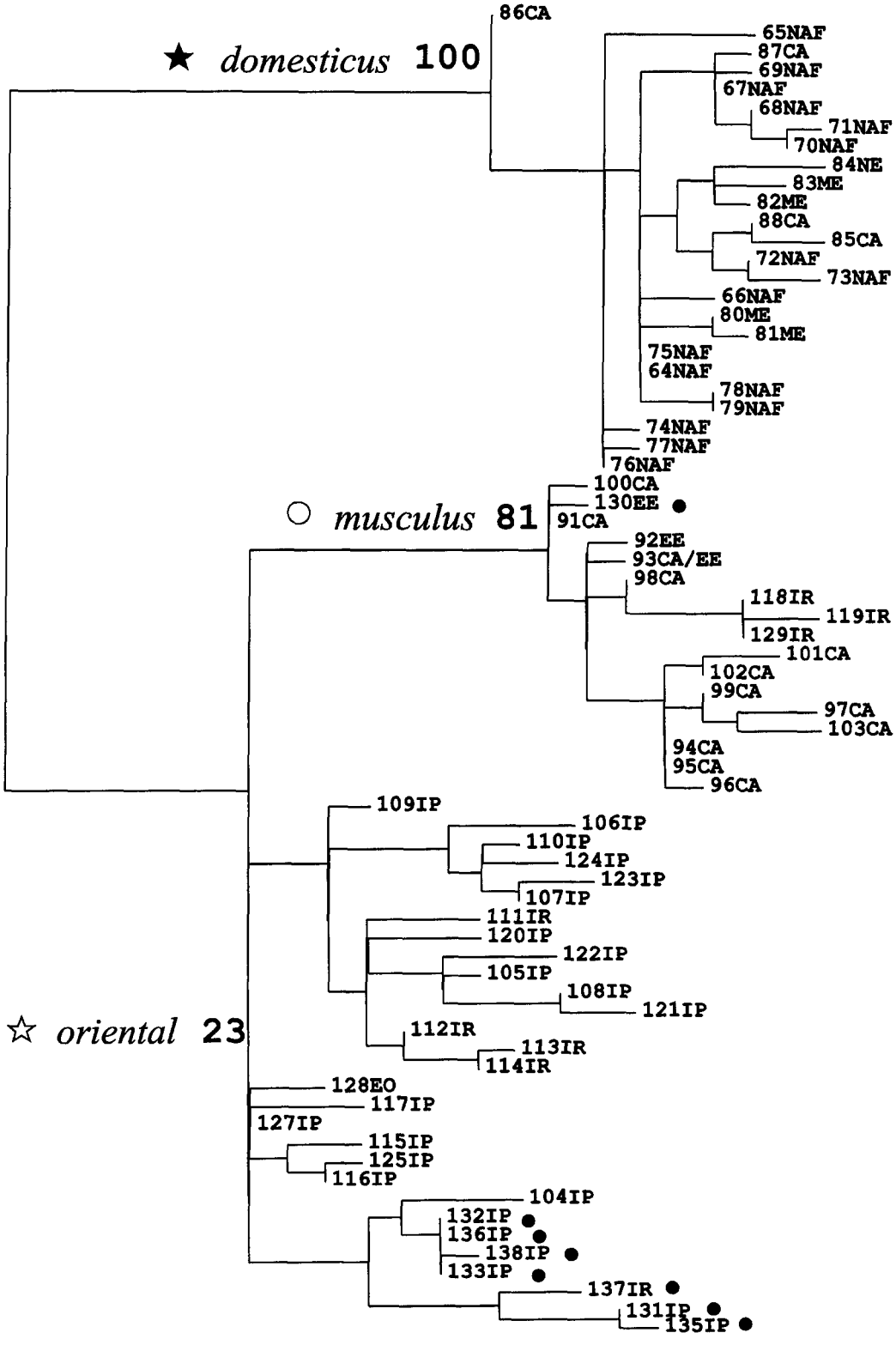

FIGURE 3.—Phylogeny of the 71 mtDNA haplotypes found in *M. musculus*. The topology of the tree is that of the majority rule consensus of trees obtained from 500 bootstrap samples. Branch lengths were calculated by maximum likelihood (see text). The four sites corresponding to variations in copy number of the TA repeat were not considered in this analysis. Haplotype numbers are as in Table 1 and Figure 1. The three major lineages defined are identified, given symbols, and named as in BOURSOT *et al.* (1996) and their bootstrap scores indicated (in percentage). ●, haplotypes with the 75–76 bp duplication. Only the 5′ copy of their duplication was used to build this phylogeny. The tree was rooted with the *M. spretus* sequence presented in Figure 1.

tween *M. m. domesticus* and *M. spretus*. To test the significance of the difference between the *Y*-chromosome and mtDNA trees, we computed the number of mutation events on the branches of the unrooted trees linking the mtDNAs and *Y* chromosomes of *M. spretus*, *M. m. domesticus*, and *M. m. musculus*. We used two different data sets to compute mtDNA branch lengths: the RFLP data of SHE *et al.* (1990, Figure 5A) and sequence data of the complete control region in PRAGER *et al.* (1996, Figure 5B). For the latter, we used the branch lengths given in their Figure 8B, which were estimated using a parsimony analysis that gave transversions five times more weight than transitions. The *Y*-chromosome data in Figure 5C are based on our *Zfy* intron sequences

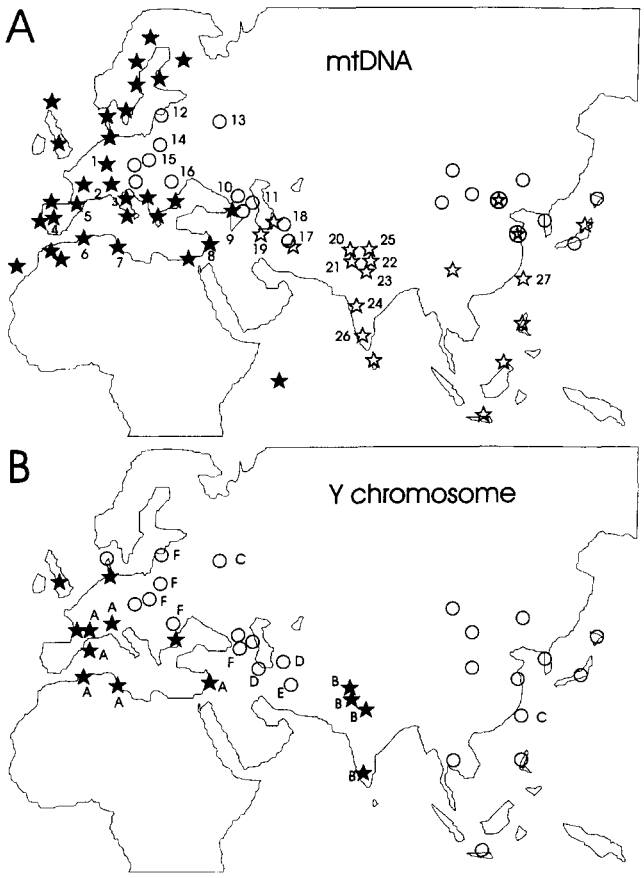

FIGURE 4.—Distribution of the major lineages of mtDNA (A) and *Y* chromosome (B) in Eurasia and North Africa. (A) ★, *domesticus* mtDNA lineage; ○, *musculus*; ☆, *oriental* (see text and Figure 3). (B) ★, *domesticus*-type *Y* chromosome; ○, *musculus*-type. The localities surveyed in the present study are numbered in A according to Table 1. Data for the other localities come from the literature for mtDNA (MORIWAKI *et al.* 1984; YONEKAWA *et al.* 1988; SAGE *et al.* 1990; PRAGER *et al.* 1993; BOURSOT *et al.* 1996) and for the *Y* chromosome (VANLERBERGHE *et al.* 1986; BOURSOT *et al.* 1989; NAGAMINE *et al.* 1992; TUCKER *et al.* 1992a; TUCKER *et al.* 1992b). Many close localities were grouped in the same data point on the maps. The letters in B indicate the alleles found at the SDR microsatellite (Table 2 and text).

(834 bp) and the sequences of the 5′ flanking region of *Sry* (1277 bp) of NACHMAN and AQUADRO (1994), with the *musculus* sequence taken from GUBBAY *et al.* (1992). Using FISHER's exact test, we calculated the probability that the numbers of events on the branches of either mtDNA tree are drawn from the same distribution as those for the *Y*-chromosome tree. We summed the lengths of the *domesticus* and *musculus* branches and performed the tests on 2 × 2 contingency tables. In the comparison between trees A (mtDNA RFLP) and C (Y chromosome) $P = 0.010 \pm 0.001$, while in that between trees B (mtDNA sequences) and C (Y chromosome) $P = 0.025 \pm 0.002$. Thus it can be considered that in the trees of Figure 5, the sum of the branches leading to *domesticus* and *musculus* is significantly shorter in the mtDNA trees than in the *Y*-chromosome tree, when compared in both cases to the *spretus* branch. Provided that the substitutions are neutral and that the mutation rates have not varied significantly along the branches of the trees, this shows that the divergence of the *Y*-chromosome lineages of the two subspecies of *M. musculus* is more recent than that of their mtDNA lineages.

Zfy intron IV variation: Because site 141 of *Zfy1* intron IV differs between the *domesticus* (DJO) and *musculus* (MAI) mice analyzed (Table 3), we further investigated its geographic variation by sequencing this intron in a panel of eight wild derived strains of various origins: BEM from Mallorca, Spain (*M. m. domesticus*), DEL from Delhi, India, TEH from Tehran, Iran, MBT from Toshevo, Bulgaria (*M. m. musculus*), and CTA from Hei-Mei, Taiwan (*M. m. castaneus*). The two *M. m. domesticus* studied were different from all others, including the mouse from Delhi. This shows that, although the *Y* chromosomes of the Indian and *domesticus* mice share a recent common ancestor, they may have diverged to a certain extent. A larger sample would be necessary to assess whether the A at site 141 is fixed in *domesticus* and the G in India.

Y-chromosome microsatellites: To detect further *Y*-chromosome variation the 73 males that were typed for the *Zfy2* deletion were also typed for two *Y*-chromosome microsatellites, one in the second intron of *Zfy2* and the other in the sex-determining region (SDR in Table 2). Only two alleles were found for the *Zfy2* microsatellite (Table 2): allele A, which has seven repetitions of the pentanucleotide (determined by sequence data not shown), is found in all *M. m. domesticus* mice, and allele B, which has six repetitions is found everywhere else, the only exception being the sample from Mallorca, which is *M. m. domesticus* but has the B allele. Whether the presence of the B allele in Mallorca is due to the persistence of ancestral polymorphism, convergent mutation, or migration is impossible to determine on the basis of the present data. Thus barring this exception, the A allele appears to be a fixed autapomorphy of *M. m. domesticus*, and this microsatellite locus shows the same pattern of variation as site 141 in the forth intron of *Zfy1*, giving further evidence that the *Y* chromosomes of *M. m. domesticus* and the Indian mice have accumulated some degree of differentiation despite their belonging to the same major lineage.

The results obtained with the microsatellite defined in the SDR are shown in Table 2. Although the primers were defined in the inverted repeat that surrounds *Sry*, single-band patterns were observed in all animals tested. We thus do not know whether both copies of the repeat were amplified. On the basis of the size of the fragment obtained, six *Y*-chromosome haplotypes could be defined, the geographic distribution of which is shown on Figure 4B. Apart from allele C, which is found in both Moscow and Taiwan, the distribution of the other alleles tends to be clustered: allele A is found in *M. m. domesticus*; allele B is found in India and Pakistan; alleles

TABLE 2

Polymorphism of the Y chromosome for several markers

Locality number[a]	Locality	Country	n[b]	*Zfy2* del[c]	SDR μsat[d]	*Zfy* μsat[e]	*Zfy* RFLP	pY353 RFLP	Reference
	Pomorie	Bulgaria	1	—	—	—	—	D	BOURSOT *et al.* (1989)
	Cambridge	England	1	—	—	—	—	D	BOURSOT *et al.* (1989)
	Windsor	Canada	1	—	—	—	—	D	BOURSOT *et al.* (1989)
	Toulouse	France	1	D	—	—	D	—	NAGAMINE *et al.* (1992)
	Montpellier	France	1	D	A	A	D	D	NAGAMINE *et al.* (1992); BOURSOT *et al.* (1989)
	Tirano	Italy	1	D	—	—	D	—	NAGAMINE *et al.* (1992)
3	Orcetto	Italy	1	D	A	A	—	—	This report
5	Mallorca Is	Spain	1	D	A	B	—	—	This report
	Monastir	Tunisia	1	D	—	—	—	—	NAGAMINE *et al.* (1992)
7	Seven localities	Tunisia	13	D	A	A	—	—	This report
6	Oran	Algeria	1	D	A	A	—	—	This report
8	Kefar Galim	Israel	1	D	A	A	—	—	This report
	Lahore	Pakistan	1	D	—	—	D	D	NAGAMINE *et al.* (1992); BOURSOT *et al.* (1989)
20	Islamabad	Pakistan	2	D	B	B	—	—	This report
20	Rawalpindi	Pakistan	2	D	B	B	—	—	This report
20	Tahmasapabad	Pakistan	1	D	B	B	—	—	This report
21	Angah	Pakistan	1	D	B	B	—	—	This report
23	Delhi	India	7	D	B	B	D	—	NAGAMINE *et al.* (1992); this report
26	Masinagudi	India	4	D	B	B	—	—	This report
	Two localities	Denmark	2	M	—	—	M	—	NAGAMINE *et al.* (1992)
13	Moscow	Russia	1	M	C	B	—	—	This report
12	Riga	Lettonia	2	M	F	B	—	—	This report
14	Bialowieza	Poland	1	M	F	B	—	—	This report
15	Illmitz	Austria	1	M	F	B	—	—	This report
	Bratislava	Czech Rep	1	M	—	—	M	—	NAGAMINE *et al.* (1992)
	Moravia	Czech Rep	1	M	—	—	—	—	NAGAMINE *et al.* (1992)
16	Toshevo	Bulgaria	1	M	F	B	—	M	BOURSOT *et al.* (1989); this report
9	Alazani	Georgia	1	M	F	B	—	—	This report
9	Didich-Shiraki	Georgia	2	M	F	B	—	—	This report
9	Gardabani	Georgia	3	M	F	B	—	—	This report
9	Gori	Georgia	1	M	F	B	—	—	This report
9	Kareli	Georgia	1	M	F	B	—	—	This report
9	Lagodekhi	Georgia	4	M	F	B	—	—	This report
9	Lissi	Georgia	4	M	F	B	—	—	This report
9	Shiraskaya	Georgia	1	M	F	B	—	—	This report
9	Vashlavan	Georgia	2	M	F	B	—	—	This report
9	Megri	Armenia	6	M	F	B	—	—	This report
18	Mashhad	Iran	1	M	—	—	M	—	NAGAMINE *et al.* (1992)
18	Mashhad	Iran	2	M	D	B	—	—	This report
19	Tehran	Iran	1	M	D	B	—	—	This report
17	Birdjan	Iran	2	M	E	B	—	—	This report
17	Kakhk	Iran	2	M	E	B	—	—	This report
	Beijing	China	1	M	—	—	—	—	NAGAMINE *et al.* (1992)
	Five localities	China	6	—	—	—	—	M	BOURSOT *et al.* (1989)
	Changchun	China	1	M	—	—	—	M	NAGAMINE *et al.* (1992); BOURSOT *et al.* (1989)
27	He-Mei	Taiwan	1	M	C	B	—	—	This report
	Taichun	Taiwan	1	M	—	—	—	M	NAGAMINE *et al.* (1992); BOURSOT *et al.* (1989)
	Chonburi	Thailand	1	M	—	—	—	—	NAGAMINE *et al.* (1992)
	Seven localities	Japan	7	M	—	—	M	M	NAGAMINE *et al.* (1992); BOURSOT *et al.* (1989)
	Bandung	Indonesia	1	M	—	—	—	—	NAGAMINE *et al.* (1992)
	Three localities	SE Asia	3	—	—	—	—	M	BOURSOT *et al.* (1989)
	Two localities	Korea	4	—	—	—	—	M	BOURSOT *et al.* (1989)

[a] Locality numbers as on the map of Figure 4A.
[b] Sample size.
[c] M, presence of the 18-bp deletion in the last exon of *Zfy2* (*musculus*-type). D, no deletion (*domesticus*-type).
[d] Microsatellite defined in the SDR (see text).
[e] Microsatellite defined in the second intron of *Zfy2* (see text).

TABLE 3

Variations observed in the introns of *Zfy1* and *Zfy2*

Taxon	Strain[a]	*Zfy1* intron IV							*Zfy2* intron II	*Zfy2* intron IV								
		40[b]	96	141	197	283	286	299	141	7	68	85	89	106	197	255	278	279
M. m. domesticus	DJO	T	A	**A**	A	C	C	C	T	G	A	G	G	C	T	G	T	A
M. m. musculus	MAI	T	A	**G**	A	C	C	C	T	G	A	G	G	C	T	G	T	A
M. spretus	SFM	A	G	**G**	C	A	T	A	G	T	T	A	T	A	C	A	A	T

[a] Strain name: DJO, Orcetto, Italy; MAI, from Ilmitz, Austria; and SFM, Montpellier, France.
[b] Numberings of positions are those of the entries in the EMBL database, accession nos. U31311 and U31418–U31425.

D and E are found in Iran; allele F is found in *M. m. musculus* (except Moscow) and in the Caucasus. Thus this locus reveals subdivisions inside the major divisions defined on the basis of the other markers analyzed. This indicates that the polymorphisms for this microsatellite locus are more recently derived than those for the other markers and that new *Y*-chromosome variants tend to spread and get fixed locally in mouse populations.

DISCUSSION

MtDNA phylogeny: Our analysis of mtDNA and *Y*-chromosome variation in *M. musculus* has led to the identification of phylogeographic assemblages for both markers. The data did not allow full resolution of the mtDNA phylogeny but showed that at least three anciently diverged lineages could be recognized (Figure 3). However we could not exclude the fact that what we called the *oriental* lineage consists of two or more lineages, among which lies the *musculus* lineage. Even if this is the case, it seems likely that the *musculus* haplotypes are monophyletic and that their root lies deep among the phylogeny of the *oriental* haplotypes. Similar conclusions can be drawn from RFLP data on the whole mtDNA molecule (BOURSOT *et al.* 1996). Whether the branching of the *musculus* lineage is slightly older or younger than the ancestor of the *oriental* haplotypes is thus not known, but this will not affect the discussion that follows.

MtDNA phylogeography: The present study provides new information about two important geographical regions that were poorly sampled in the previous RFLP study but that were thought to be potential zones of overlap of the distributions of two or more of the mtDNA lineages (BOURSOT *et al.* 1996). The first is in the Caucasus, where we found both *domesticus* and *musculus* haplotypes. These results, as well as data on nuclear gene polymorphisms, which are described in greater detail elsewhere (ORTH *et al.* 1996), show that the Caucasus is a zone of secondary intergradation between *domesticus* and *musculus* as suggested by MILISHNIKOV *et al.* (1990) and FRISMAN *et al.* (1990). PRAGER *et al.* (1996) also found both *domesticus* and *musculus* mtDNA in Abkhasia. The second region of interest is northeastern Iran. We had previously found an *oriental* haplotype from Mashhad, but the new samples from this locality analyzed were found to have *musculus* haplo-

TABLE 4

Nucleotide divergence between *M. m. musculus*, *M. m. domesticus* and *M. spretus*

	M. m. musculus (%)	*M. m. domesticus* (%)	*M. spretus* (%)
M. m. musculus		0.23 ± 0.16	1.84 ± 0.46
M. m. domesticus	4.1 ± 1.2		2.06 ± 0.49
M. spretus	11.0 ± 2.4	10.0 ± 1.7	

Nucleotide divergence (±SE) is inferred for the *Zfy* introns sequences (above diagonal, this study) and for mtDNA (below diagonal, RFLP data of SHE *et al.* 1990).

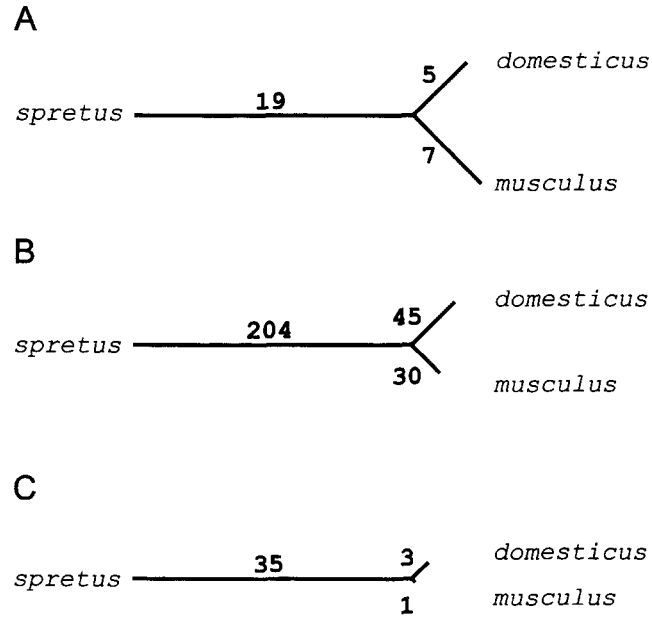

FIGURE 5.—Test of the differences between the mtDNA and *Y*-chromosome trees. The trees represent the unrooted phylogeny of *M. m. domesticus*, *M. m. musculus*, and *M. spretus*. The numbers along the branches represent (A) the number of mtDNA restriction site changes, inferred from the data in SHE *et al.* (1990); (B) the number of nucleotide changes in the control region of mtDNA, as inferred in PRAGER *et al.* (1996); (C) the number of *Y*-chromosome nucleotide changes, inferred from our *Zfy* intron data, pooled with those of NACHMAN and AQUADRO (1994) on the 5′ flanking region of *Sry*.

types. Approximately 250 km south of Mashhad, in Kakhk, we found mice with the *musculus* lineage, and 150 km further south, in Birjan, we found the *oriental* lineage. It therefore appears that in northeastern Iran the distributions of the *musculus* and *oriental* lineages overlap, but whether this corresponds to secondary admixture or primary differentiation requires further investigation. Because we cannot exclude that the *musculus* lineage is in fact included in the group of *oriental* lineages (see above), the presence of *musculus* haplotypes in northern Iran could either be a remnant of the ancestral population that was at the origin of the expansion of *M. m. musculus* north of the Himalayas and the Caspian Sea or the result of secondary admixture.

Y-chromosome phylogeography: By combining molecular and geographic variation, we were able to define two major *Y*-chromosome lineages. One lineage includes a *Y*-chromosome type found in the Indian subcontinent, and another found in *M. m. domesticus*, and differing at site 141 of *Zfy1* intron IV, and at the microsatellite in *Zfy2* intron II. The second lineage differs from the first one by an RFLP for the pY353 repeated sequence, the 18-bp deletion in *Zfy2* last exon, a *Zfy TaqI* RFLP, and substitutions at sites 8701 and 8731 in *Sry*. This lineage comprises a type found in *M. m. musculus* and *M. m. castaneus*, as well as in the Caucasus and in northeastern Iran (including a subtype from Tehran with an autapomorphy at site 8706 of *Sry*) and a second type found in Japanese mice, *M. m. molossinus*, as well as in old inbred laboratory strains. The *molossinus* type differs from the *musculus-castaneus* type by an *Sry TaqI* RFLP, and substitutions at sites 8491 and 8711 in *Sry*. The data for the SDR microsatellite (Table 2) subdivide the Indian type into two subtypes (alleles B and E) and the *musculus-castaneus* type into three subtypes (alleles C, D, and F). No subdivision is revealed in *M. m. domesticus*, in which a single allele was found (allele A).

Compared geographic variations of mtDNA and *Y* chromosomes: A major point coming out of this study is the apparent discrepancy between the geographic distributions of the lineages in *M. musculus* that can be defined using either *Y*-chromosome or mtDNA variation. Comparison of the two maps in Figure 4, which give the distributions of the major mtDNA and the *Y*-chromosome types, shows that several different combinations occur. Although *M. m. castaneus* and *M. m. musculus* share the same *Y*-chromosome lineage, they harbor different mtDNAs. The Indian and Pakistani populations harbor a *domesticus*-like *Y* chromosome, but share mtDNA lineages with *M. m. castaneus*.

There are two possible ways to account for these apparent discrepancies in the distributions of the mtDNA and *Y*-chromosome lineages. It could be that the two *Y*-chromosome lineages that are now found in *M. musculus* had already differentiated in the ancestral population and that the different combinations that occur today were sorted together when the radiation process began (ancestral sorting hypothesis). Such a process would not necessarily result in the same phylogeographic pattern for all the genes (WU 1991; HUDSON 1992). Alternatively, the present distribution of the *Y*-chromosome lineages could have been acquired after the differentiation of the subspecies, through secondary sweeps between subspecies (secondary sweep hypothesis).

Recent divergence of *Y*-chromosome lineages: The ancestral sorting hypothesis would imply that distinct ancestors of the two *Y*-chromosome lineages and the three mtDNA lineages already existed in the ancestral population before the geographical radiation, because this is the only way by which sister populations can share lineages in the absence of secondary introgression. However our *Zfy* intron sequence data and the *Sry* flanking region sequences of NACHMAN and AQUADRO (1994) show that the divergence time of the two *Y*-chromosome lineages is about one fourth of that found for the mtDNA lineages of *M. musculus* subspecies, if their respective divergences with *M. spretus* are taken as a standard (Table 4). The data provide strong support for the idea that the divergence of the *Y*-chromosome lineages is more recent than that of the mtDNA lineages (Figure 5), which could mean that the *Y*-chromosome divergence is posterior to the isolation between the subspecies, thus invalidating the ancestral sorting hypothesis.

Although the lower divergence of the *Y*-chromosome lineages could be an artifact due to an underestimation of the mtDNA divergence between *M. musculus* and *M. spretus*, because of saturation of the fast-evolving mtDNA, this does not seem likely, as the overall mtDNA divergence estimated by RFLP between the two species is only of the order of 10%, and saturation is not expected to be a problem in this range. Even when comparing species that are twice as divergent as *M. musculus* and *M. spretus*, SHE *et al.* (1990) found no evidence of saturation of mtDNA compared to scnDNA. PRAGER *et al.* (1996), however, found evidence for the saturation of transition substitutions between *Mus musculus* and *M. spretus* in the data set that we used to test the differences between the mtDNA and *Y*-chromosome trees (Figure 5B). At least some of this saturation was corrected for in the parsimony analysis they used, but we cannot be sure that all of it was. However comparison with nuclear genes suggest that the difference of divergence time between the mtDNA and *Y*-chromosome lineages is due to an exceptionally low divergence for the *Y* chromosomes rather than to an exceptionally high divergence for mtDNA. The comparison of the among-subspecies divergence *vs.* the among-species divergence estimated for nuclear DNA by scnDNA/DNA hybridization (SHE *et al.* 1990) and 5S ribosomal DNA spacer sequences (SUZUKI *et al.* 1994) gives ratios of 0.32 and 0.36, respectively, which are similar to the 0.37–0.41 ratio observed for mtDNA in this study (see Table 4).

The ancestral sorting hypothesis: It could be that differences of divergence time of the *Y*-chromosome and mtDNA lineages result from different levels of ancestral polymorphism existing prior to the radiation of the subspecies. A greater ancestral polymorphism for mtDNA could have caused the divergence of mtDNA lineages to predate the radiation event to a greater extent than in the case of the *Y* chromosomes, thus producing the pattern observed. Lower polymorphism of the *Y* chromosome as compared to mtDNA was observed in *M. m. domesticus* by NACHMAN and AQUADRO (1994). As discussed by these authors, one explanation for this could be that effective population size is smaller for males than for females because of possible polygyny in mice. In the northern part of the Indian subcontinent, which is thought to be the cradle of the species and where the level of diversity is greatest, BOURSOT *et al.* (1996) found levels of mtDNA nucleotide diversities calculated from RFLP data of the order of 2%. If we suppose that the diversity in the ancestral population of *M. musculus* was of the same order, then only 2% of the roughly 4% difference between mtDNAs of *M. m. domesticus* and *M. m. musculus* (Table 5) would be attributable to divergence after the ancestral population split. However even if we also take into account a similar amount of polymorphism present before the split between *M. spretus* and *M. musculus*, the ratio of mtDNA intersubspecific:interspecific divergence decreases from roughly 0.4 (Table 4) to roughly 0.2. Although this is closer to the value obtained for the *Y* chromosome (~0.1, Table 4) it is still twice as high and may not explain the different patterns completely.

The secondary sweep hypothesis: Therefore, the extremely low divergence of the *Y*-chromosome lineages as compared to mtDNA and nuclear genes between subspecies invites us to consider the possibility that the divergence of the two *Y*-chromosome lineages occurred after the subspecies had started their radiation and differentiation, *i.e.*, when the ancestors of the present peripheral subspecies already existed (the secondary sweep hypothesis). We must determine in what region the ancestors of the two *Y*-chromosome lineages appeared and to which regions their descendants spread. *M. m. domesticus* is believed to have originated in the fertile crescent (BOURSOT *et al.* 1996; DIN *et al.* 1996). The *domesticus*-type *Y* chromosome could have appeared there and spread to the northern Indian subcontinent or vice versa. Nuclear gene polymorphism provides evidence for some sort of genetic continuum between India and *M. m. domesticus* (DIN *et al.* 1996). However, no such evidence exists for mtDNA (present article and BOURSOT *et al.* 1996), and the populations from the northern Indian subcontinent are more distantly related to *M. m. domesticus* (Nei's genetic distance $D = 0.15$, based on allozyme data), with which they share the same *Y*-chromosome lineage, than to *M. m. castaneus* ($D = 0.06$) with which they do not (DIN *et al.* 1996). Thus if such a *Y*-chromosome sweep occurred, it was apparently not accompanied by massive introgression of mtDNA or autosomal genes. The ancestral *musculus-castaneus Y* chromosomes could have appeared in either of these two subspecies and spread to the other. Since the presumed centers of differentiation of *M. m. musculus* and *M. m. castaneus* are far apart (west and east of the Himalayas, respectively, BOURSOT *et al.* 1996; DIN *et al.* 1996), this would have been by introgression through the zone of secondary contact in China. Alternatively, the *musculus-castaneus Y* chromosome would have appeared in the Indian subcontinent and spread to *musculus* and *castaneus* before it was replaced in India by the *domesticus*-type *Y* chromosome. Whatever the scenario, the secondary sweep hypothesis implies from one to three sweeps over major geographic subdivisions of the species, presumably driven by strong selection.

Selection on the *Y* chromosome: Evidence of a recent selective sweep across subspecies of *M. musculus* exists for the *t* haplotype. There is molecular evidence that all *t* haplotypes found in the different subspecies of *M. musculus* are descendant from a haplotype that spread only recently into house mouse populations when they were already differentiated (SILVER *et al.* 1987; MORITA *et al.* 1992; HAMMER and SILVER 1993; SILVER 1993). Because the *t* haplotype is a transmission distorter (SILVER 1985), the driving force leading to its apparent sweep is clearly identified. As argued in the Introduction, there are theoretical reasons to think that the *Y* chromosome could be particularly involved in conflicts of interest promoting transmission distortion (FRANK 1991a; HURST and POMIANKOWSKI 1991; HURST 1994) or some other sort of competition driving rapid coevolution of the genes involved (RICE 1996). In fact, although the two lineages of *Y* chromosomes in *M. musculus* have apparently evolved from a very recent common ancestor, they have accumulated significant differences. As mentioned in the Introduction, several families of repeated sequences have been extensively and differentially reorganized in the two lineages (BISHOP *et al.* 1985; NISHIOKA 1987, 1988; BOURSOT *et al.* 1989; TUCKER *et al.* 1989; NISHIOKA *et al.* 1993, 1994; LUNDRIGAN and TUCKER 1994). Whether this is linked to functional differences is unclear because the function of these repeated sequences is not known. One of them, pY353, is transcribed in the testis and suspected to play a role in sperm head development (CONWAY *et al.* 1994).

Selection acting on any gene or sequence on the *Y* chromosome should reduce neutral intraspecific polymorphism everywhere else on this nonrecombining chromosome (HUDSON *et al.* 1987). NACHMAN and AQUADRO (1994) found that the ratio of polymorphism within *M. m. domesticus* to the divergence with *M. spretus* was approximately four times lower for the 5′ flanking region of *Sry* than for mtDNA control region. Their result is compatible with our finding that the divergence time of the *Y*-chromosome lineages is one fourth

of that for the mtDNA lineages in *M. musculus* but does not provide strong evidence for hitch-hiking, as argued by the authors themselves.

Evidence for selection can also be looked for in patterns of nucleotide substitution in protein coding regions. TUCKER and LUNDRIGAN (1993) found a high ratio of replacement to synonymous substitutions when comparing partial nucleotide sequences of *Sry* between species of Murinae. However such a pattern is not seen in the HMG box, which is thought to code for the active site of the protein (COWARD *et al.* 1994). The high proportion of replacement substitutions outside the HMG box might reflect a relaxation of the selection pressure, the substitutions being neutral, or slightly deleterious and accumulating because of drift and the absence of recombination. That selection pressure is relaxed is suggested by the substantial length polymorphism of amino acid repeats reported 3′ of HMG (MILLER *et al.* 1995; CARLISLE *et al.* 1996). If the mutations were neutral, divergence between subspecies of *M. musculus* compared to the divergence with *M. spretus* should be as low for regions of *Sry* outside the HMG box, as it is for the noncoding *Y*-chromosome regions we studied (Figure 5C). The data of LUNDRIGAN and TUCKER (1994) on the 5′ end of *Sry* suggest that this is not the case. They found from two to four substitutions between subspecies of *M. musculus*, but only from four to eight between these subspecies and *M. spretus*. This apparent acceleration of *Sry* evolution among *M. musculus* subspecies is suggestive of some form of selection, but obviously needs confirmation on a larger data set.

Conclusion: Overall the available data show that the two major lineages of *Y* chromosomes in *M. musculus* have differentiated very recently, as compared to mtDNA lineages, and spread over vast regions of the species in a short period of time. The SDR microsatellite data show that further differentiation has occurred since this initial spread and that the new variants have spread locally, leading to the clustered geographic distribution of alleles observed. Although the extremely recent divergence of the *Y*-chromosome lineages suggests that the spread of *Y*-chromosome lineages could have been posterior to the geographic radiation of the species, the uncertainties in the estimations of nucleotide divergence, as well as errors related to the stochasticity of the mutation process, do not allow to formally exclude the ancestral sorting hypothesis. At least part of the difference in coalescence time between mtDNA and the *Y* chromosome could be accounted for by differences between male and female effective population sizes. However we have also noted that some observations on the patterns of molecular evolution of the *Y* chromosome are suggestive of the influence of natural selection, irrelevant of whether or not secondary sweeps have occurred between subspecies.

Molecular *vs.* functional differentiation of the *Y* chromosome: A related and fundamental question is to understand the link between observed molecular differences and functional differences between the *Y*-chromosome types, in relation to their role in the development of reproductive isolation. One way to address the question that has been successful in previous studies is to look at the exact distribution of *Y*-chromosome types in zones of contact between subspecies. In the case of the hybrid zone between *M. m. domesticus* and *M. m. musculus* in Europe, the low level of introgression of sex chromosomes compared to other markers provides evidence for their major role in the genetic isolation between the two subspecies (VANLERBERGHE *et al.* 1986; TUCKER *et al.* 1992b; DOD *et al.* 1993). Whether a similar absence of introgression prevails in other zones of contact remains to be determined. In the Caucasus and northeastern Iran we found two mtDNA lineages, which indicates that these are regions of admixture between subspecies. However in both cases our samples displayed only the *musculus Y* chromosome, and we failed to sample the zones of contact between the *Y*-chromosome lineages, which must be looked for further south in both cases. Our data also allow us to predict that a zone of contact between *Y*-chromosome lineages must exist in the northeastern Indian subcontinent. The analysis of patterns of exchange of the sex chromosomes and other nuclear genes across these contact zones, as well as more sequences of the coding *Y*-chromosome DNA, should give clues as to the link between molecular and functional differentiation of the *Y* chromosome.

We are grateful to R. ANAND, P. A. CAZENAVE, D. DARVICHE, W. DIN, A. KANDAUROV, E. LYAPUNOVA, and N. VORONTSOV for invaluable help in obtaining the samples. J. COLLIGNON kindly provided the unpublished intron/exon map of *Zfy1*. We appreciate the skillful technical assistance of A. ORTH. We thank C. H. LANGLEY and M. SLATKIN for helpful comments on an earlier version of the manuscript. B. DOD made numerous suggestions to improve the manuscript. We thank F. BONHOMME for his constant support, encouragement, and numerous discussions. S.B. was supported by a fellowship from the French Ministère de l'Education Nationale. This research was partially supported by the Biotechnology Programme of the Commission of European Communities, contract BIO2CT-920476 to P.B., and by a grant from the Ministère de l'Education Nationale (ACCSV no.7) to P.B.

LITERATURE CITED

BIBB, M. J., R. A. VAN ETTEN, C. T. WRIGHT, M. W. WALBERG and D. A. CLAYTON, 1981 Sequence and gene organization of mouse mitochondrial DNA. Cell **26:** 167–180.

BISHOP, C. E., P. BOURSOT, F. BONHOMME and D. HATAT, 1985 Most classical *Mus musculus domesticus* laboratory mouse strains carry a *Mus musculus musculus* Y chromosome. Nature **315:** 70–72.

BOURSOT, P., H. YONEKAWA and F. BONHOMME, 1987 Heteroplasmy in mice with deletion of a large coding region of mitochondrial DNA. Mol. Biol. Evol. **4:** 46–55.

BOURSOT, P., F. BONHOMME, J. CATALAN and K. MORIWAKI, 1989 Variation of a Y chromosome repeated sequence across subspecies of *Mus musculus*. Heredity **63:** 289–297.

BOURSOT, P., J.-C. AUFFRAY, J. BRITTON-DAVIDIAN and F. BONHOMME, 1993 The evolution of house mice. Annu. Rev. Ecol. Syst. **24:** 119–152.

BOURSOT, P., W. DIN, R. ANAND, D. DARVICHE, B. DOD *et al.*, 1996

Origin and radiation of the house mouse: mitochondrial DNA phylogeny. J. Evol. Biol. **9:** 391–415.

CARLISLE, C., H. WINKING, D. WEICHENHAN and C. M. NAGAMINE, 1996 Absence of correlation between *Sry* polymorphisms and XY sex reversal caused by the *M. m. domesticus* Y chromosome. Genomics **33:** 32–45.

CHANG, B. H.-J., and W.-S. LI, 1995 Estimating the intensity of male-driven evolution in rodents by using X-linked *Ube 1* genes and pseudogenes. J. Mol. Evol. **40:** 70–77.

CHANG, B. H.-J., L. C. SHIMMIN, S.-K. SHYUE, D. HEWETT-EMMETT and W.-H. LI, 1994 Weak male-driven molecular evolution in rodents. Proc. Natl. Acad. Sci. USA **91:** 827–831.

CHARLESWORTH, B., J. A. COYNE and N. H. BARTON, 1987 The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. **130:** 113–146.

CHARLESWORTH, B., J. A. COYNE and H. A. ORR, 1993 Meiotic drive and unisexual hybrid sterility. Genetics **133:** 421–424.

CONWAY, S. J., S. K. MAHADEVAIAH, S. M. DARLING, B. CAPEL, A. M. RATTIGAN *et al.*, 1994 Y353B: a candidate multiple-copy spermiogenesis gene on the mouse Y chromosome. Mamm. Genome **5:** 203–210.

COWARD, P., K. NAGAI, D. CHEN, H. D. THOMAS, C. M. NAGAMINE *et al.*, 1994 Polymorphism of a CAG trinucleotide repeat within *Sry* correlates with B6.YDom sex reversal. Nat. Genet. **6:** 245–250.

COYNE, J. A., and H. A. ORR, 1993 Further evidence against meiotic-drive models of hybrid sterility. Evolution **47:** 685–687.

DIN, W., R. ANAND, P. BOURSOT, D. DARVICHE, E. JOUVIN-MARCHE *et al.*, 1996 Origin and radiation of the house mouse: clues from nuclear genes. J. Evol. Biol. **9:** 519–539.

DOD, B., L. S. JERMIIN, P. BOURSOT, V. M. CHAPMAN, J. T. NIELSEN *et al.*, 1993 Counterselection on sex chromosomes in the *Mus musculus* European hybrid zone. J. Evol. Biol. **6:** 529–546.

FELSENSTEIN, J., 1993 *PHYLIP (Phylogeny Inference Package).* University of Washington, Seattle.

FRANK, S. A., 1991a Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. Evolution **45:** 262–267.

FRANK, S. A., 1991b Haldane's rule: a defense of the meiotic drive theory. Evolution **45:** 1714–1717.

FRISMAN, L. V., K. V. KOROBITSINA, L. V. YAKIMENKO, F. M. BOKSHTEIN and A. I. MUNTYANU, 1990 Genetic differentiation of the U.S.S.R. house mice: electrophoretic study of proteins. Biol. J. Linn. Soc. **41:** 65–72.

GRAVEN, L., G. PASSARINO, O. SEMINO, P. BOURSOT, S. SANTACHIARA-BENERECETTI *et al.*, 1995 Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. Mol. Biol. Evol. **12:** 334–345.

GUBBAY, J., N. VIVIAN, A. ECONOMOU, D. JACKSON, P. GOODFELLOW *et al.*, 1992 Inverted repeat structure of the *Sry* locus in mice. Proc. Natl. Acad. Sci. USA **89:** 7953–7957.

GUO, S. W., and E. A. THOMPSON, 1992 Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics **48:** 361–372.

HALDANE, J. B. S., 1922 Sex ratio and unisexual sterility of hybrid animals. J. Genet. **12:** 101–109.

HAMMER, M. F., 1995 A recent common ancestry for human Y chromosomes. Nature **378:** 376–378.

HAMMER, M. F., and L. M. SILVER, 1993 Phylogenetic analysis of the Alpha-Globin pseudogene-4 (*Hba-ps4*) locus in the house mouse species complex reveals a stepwise evolution of *t* haplotypes. Mol. Biol. Evol. **10:** 971–1001.

HUDSON, R. R., 1992 Gene trees, species trees and the segregation of ancestral alleles. Genetics **131:** 509–512.

HUDSON, R. R., M. KREITMAN and M. AGUADÉ, 1987 A test of neutral evolution based on nucleotide data. Genetics **116:** 153–159.

HURST, L. D., 1994 Embryonic growth and the evolution of the mammalian Y chromosome. I. The Y as an attractor for selfish growth factors. Heredity **73:** 223–232.

HURST, L. D., and A. POMIANKOWSKI, 1991 Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of HALDANE's rule and related phenomena. Genetics **128:** 841–856.

JOHNSON, N. A., and C.-I. WU, 1992 An empirical test of the meiotic drive models of hybrid sterility: sex-ratio data from hybrids between *Drosophila simulans* and *Drosophila sechellia.* Genetics **130:** 507–511.

KAPLAN, N. L., R. R. HUDSON and C. H. LANGLEY, 1989 The "hitchhiking effect" revisited. Genetics **123:** 887–899.

KOOPMAN, P., J. GUBBAY, J. COLLIGNON and R. LOVELL-BADGE, 1989 *Zfy* gene expression patterns are not compatible with a primary role in mouse sex determination. Nature **342:** 940–942.

KUMAR, S., K. TAMURA and M. NEI, 1993 *MEGA: Molecular Evolutionary Genetics Analysis.* The Pennsylvania State University, University Park, PA.

LUNDRIGAN, B. L., and P. K. TUCKER, 1994 Tracing paternal ancestry in mice, using the Y-linked, sex determining locus, *Sry.* Mol. Biol. Evol. **11:** 483–492.

MAYNARD SMITH, J., and J. HAIGH, 1974 The hitchhiking effect of a favorable gene. Genet. Res. **23:** 23–35.

MILISHNIKOV, A. N., A. N. RAFIEV, L. A. LAVRENCHENKO and V. N. ORLOV, 1990 High level of introgression of *Mus domesticus* genes in populations of *Mus musculus* S. str. of Transcaucasia. Dok. Akad. Nauk SSSR **311:** 764–768.

MILLER, K. E., B. L. LUNDRIGAN and P. K. TUCKER, 1995 Length variation of CAG repeats in *Sry* across populations of *Mus domesticus.* Mamm. Genome **6:** 206–208.

MIYATA, T., H. HAYASHIDA, K. KUMA, K. MITSUYASU and T. YASUNAGA, 1987 Male-driven molecular evolution: a model and nucleotide sequence analysis. Cold Spring Harbor Symp. Quant. Biol. **52:** 863–867.

MORITA, T., H. KUBOTA, K. MURATA, M. NOZAKI, C. DELARBRE *et al.*, 1992 Evolution of the mouse *t* haplotype: recent and worldwide introgression to *Mus musculus.* Proc. Natl. Acad. Sci. USA **89:** 6851–6855.

MORIWAKI, K., H. YONEKAWA, O. GOTOH, M. MINEZAWA, H. WINKING *et al.*, 1984 Implications of the genetic divergence between European wild mice with Robertsonian translocations from the viewpoint of mitochondrial DNA. Genet. Res. **43:** 227–287.

NACHMAN, M. W., and C. F. AQUADRO, 1994 Polymorphism and divergence of the 5′ flanking region of the sex-determining locus, *Sry,* in mice. Mol. Biol. Evol. **11:** 539–547.

NAGAMINE, C. M., Y. NISHIOKA, K. MORIWAKI, P. BOURSOT, F. BONHOMME *et al.*, 1992 The musculus-type Y chromosome of the laboratory mouse is of Asian origin. Mamm. Genome **3:** 84–91.

NAGAMINE, C. M., P. BOURSOT, Y.-F. C. LAU and K. MORIWAKI, 1994a Evolution of the Y chromosome in the wild mouse, pp. 41–55 in *Genetics in Wild Mice—Its Application to Biomedical Research,* edited by K. MORIWAKI, T. SHIROISHI and H. YONEKAWA. Japan Sci. Soc. Press, Tokyo; S. Karger, Basel, Switzerland.

NAGAMINE, C. M., T. SHIROISHI, N. MIYASHITA, K. TSUCHIYA, H. IKEDA *et al.*, 1994b Distribution of the molossinus allele of *Sry,* the testis-determining gene, in wild mice. Mol. Biol. Evol. **11:** 864–874.

NEI, M., and J. MILLER, 1990 A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. Genetics **125:** 873–879.

NISHIOKA, Y., 1987 Y-chromosomal DNA polymorphism in mouse inbred strains. Genet. Res. **50:** 69–72.

NISHIOKA, Y., 1988 Evolutionary characterization of a Y chromosomal sequence conserved in the genus *Mus.* Genet. Res. **52:** 145–150.

NISHIOKA, Y., and E. LAMOTHE, 1986 Isolation and characterization of a mouse Y chromosomal repetitive sequence. Genetics **113:** 417–432.

NISHIOKA, Y., B. M. DOLAN, V. F. PRADO, L. ZAHED and H. TYSON, 1993 Comparison of mouse Y-chromosomal repetitive sequences isolated from *Mus musculus, M. spicilegus,* and *M. spretus.* Cytogenet. Cell Genet. **64:** 54–58.

NISHIOKA, Y., B. M. DOLAN, L. ZAHED, V. PRADO and H. TYSON, 1994 Molecular evolution of a Y-chromosomal repetitive sequence family in the genus *Mus.* Mol. Biol. Evol. **11:** 146–153.

ORR, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics **139:** 1805–1813.

ORTH, A., E. LYAPUNOVA, A. KANDAUROV, S. BOISSINOT, P. BOURSOT *et al.*, 1996 L'espèce polytypique *Mus musculus* en Transcaucasie. C. R. Acad. Sci. Paris **319:** 435–441.

PLATT, T. H., and M. J. DEWEY, 1987 Multiple forms of male-specific simple repetitive sequences in the genus *Mus.* J. Mol. Evol. **25:** 201–206.

POMIANKOWSKI, A., and L. D. HURST, 1993 Genomic conflicts underlying HALDANE's rule. Genetics **133:** 425–432.

PRAGER, E. M., R. D. SAGE, U. GYLLENSTEN, W. K. THOMAS, R. HÜBNER *et al.*, 1993 Mitochondrial DNA sequence diversity and the colonization of Scandinavia by house mice from East Holstein. Biol. J. Linn. Soc. **50:** 85–122.

PRAGER, E. M., H. TICHY and R. D. SAGE, 1996 Mitochondrial DNA variation in the eastern house mouse, *Mus musculus*: comparison with other house mice and report of a 75-bp tandem repeat. Genetics **143:** 427–446.

RAYMOND, M., and F. ROUSSET, 1995 Genepop (Version 1.2) - Population Genetics Software For Exact Tests and Ecumenicism. J. Heredity **86:** 248–249.

RICE, W. R., 1996 Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. Nature **381:** 232–234.

SAGE, R. D., E. M. PRAGER, H. TICHY and A. C. WILSON, 1990 Mitochondrial DNA variation in the house mouse, *Mus domesticus* (Rutty). Biol. J. Linn. Soc. **41:** 105–123.

SAGE, R. D., W. R. ATCHLEY and E. CAPANNA, 1993 House mice as models in systematic biology. Syst. Biol. **42:** 523–561.

SHE, J. X., F. BONHOMME, P. BOURSOT, L. THALER and F. CATZEFLIS, 1990 Molecular phylogenies in the genus *Mus*: comparative analysis of electrophoretic, scnDNA hybridisation, and mtDNA RFLP data. Biol. J. Linn. Soc. **41:** 83–103.

SHIMMIN, L. C., B.-J. CHANG and W.-H. LI, 1993 Male-Driven Evolution of DNA Sequences. Nature **362:** 745–747.

SILVER, L. M., 1985 Mouse t haplotypes. Annu. Rev. Genet. **19:** 179–208.

SILVER, L. M., 1993 The peculiar journey of a selfish chromosome: mouse *t* haplotypes and meiotic drive. Trends Genet. **9:** 250–254.

SILVER, L. M., M. HAMMER, H. FOX, J. GARRELS, M. BUCAN *et al.*, 1987 Molecular evidence for the rapid propagation of mouse t haplotypes from a single, recent, ancestral chromosome. Mol. Biol. Evol. **4:** 473–482.

SUZUKI, H., K. MORIWAKI and S. SAKURAI, 1994 Sequences and evolutionary analysis of mouse 5S rDNAs. Mol. Biol. Evol. **11:** 704–710.

TUCKER, P. K., and B. L. LUNDRIGAN, 1993 Rapid evolution of the sex determining locus in Old World mice and rats. Nature **364:** 715–717.

TUCKER, P. K., and B. L. LUNDRIGAN, 1995 The nature of gene evolution on the mammalian Y chromosome: lessons from *Sry*. Philos. Trans. R. Soc. Lond. B **350:** 221–227.

TUCKER, K. P., B. K. LEE and E. M. EICHER, 1989 Y chromosome evolution in the subgenus *Mus* (Genus *Mus*). Genetics **122:** 169–179.

TUCKER, P. K., B. K. LEE, B. L. LUNDRIGAN and E. M. EICHER, 1992a Geographic origin of the Y chromosomes in "old" inbred strains of mice. Mamm. Genome **3:** 254–261.

TUCKER, P. K., R. D. SAGE, J. WARNER, A. C. WILSON and E. M. EICHER, 1992b Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. Evolution **46:** 1146–1163.

TURELLI, M., and H. A. ORR, 1995 The dominance theory of HALDANE's rule. Genetics **140:** 389–402.

VANLERBERGHE, F., B. DOD, P. BOURSOT, M. BELLIS and F. BONHOMME, 1986 Absence of Y-chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. Genet. Res. **48:** 191–197.

WHITFIELD, L. S., R. LOVELL-BADGE and P. N. GOODFELLOW, 1993 Rapid evolution of the mammalian sex-determining gene SRY. Nature **364:** 713–715.

WHITFIELD, L. S., J. E. SULSTON and P. N. GOODFELLOW, 1995 Sequence variation of the human Y chromosome. Nature **378:** 379–380.

WU, C.-I., 1991 Inference of species phylogeny in relation to segregation of ancient polymorphisms. Genetics **127:** 429–435.

WU, C.-I., N. A. JOHNSON and M. F. PALOPOLI, 1996 Haldane's rule and its legacy: why are there so many sterile males? Trends Evol. Ecol. **11:** 281–284.

YONEKAWA, H., K. MORIWAKI, O. GOTOH, N. MIYASHITA, Y. MATSUSHIMA *et al.*, 1988 Hybrid origin of Japanese mice "*Mus musculus molossinus*": evidence from restriction analysis of mitochondrial DNA. Mol. Biol. Evol. **5:** 63–78.

ZAMBROWICZ, B. P., S. D. FINDLEY, E. M. SIMPSON, D. C. PAGE and R. D. PALMITER, 1994 Characterization of the murine *Zfy1* and *Zfy2* promoters. Genomics **24:** 406–408.

Communicating editor: G. B. GOLDING