Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 6851–6855, August 1992
Evolution

Evolution of the mouse *t* haplotype: Recent and worldwide introgression to *Mus musculus*

(*Tcp-1*/transmission ratio distortion/*t* complex/mouse chromosome 17/meiotic drive)

TAKASHI MORITA*†, HIROSHI KUBOTA*, KEIKO MURATA*, MASAMI NOZAKI*, CHRISTIANE DELARBRE‡, KEITH WILLISON§, YOKO SATTA¶, MITSURU SAKAIZUMI‖, NAOYUKI TAKAHATA¶, GABRIEL GACHELIN‡, AND AIZO MATSUSHIRO*

*Department of Microbial Genetics, Research Institute for Microbial Diseases, Osaka University, 3-1, Yamadaoka, Suita, Osaka 565, Japan; ‡Unité de Biology Moléculaire du Gène, Institut Pasteur, 25, rue du Dr. Roux, 75724, Paris, Cedex 15, France; §Institute of Cancer Research, Chester Beatty Laboratories, Fulham Road, London SW3 6JB, United Kingdom; ¶National Institute of Genetics, 1111, Yata, Mishima, Shizuoka 411, Japan; and ‖The Tokyo Metropolitan Institute of Medical Science, 3-18-22, Honkomagome, Bunkyo-ku, Tokyo 113, Japan

Communicated by Motoo Kimura, April 20, 1992

ABSTRACT **Mouse *t* haplotypes are variants of chromosome 17, consisting of four inversions. Despite the homozygous lethality and pleiotropic effect on embryonic development, sperm production, and recombination, they have widely spread in natural populations of the house mouse (10–40% in frequency) because of the meiotic drive advantage. We sequenced 14 *Tcp-1* (*t*-complex polypeptide 1) genes from four *t* haplotypes, nine wild mice, and a rat as a reference. From a comparison of intron sequences of 610 base pairs, we dated the origin of *t* haplotypes to 2.9 ± 0.7 million years ago, which predates the splitting of *Mus musculus* subspecies (≈1 million years ago). However, the *Tcp-1* intron sequences of *t* haplotypes from different *M. musculus* subspecies from various parts of the world show no divergence, indicating the recent introgression (no earlier than 0.8 million years ago) of a single ancestral type. Nucleotide changes in coding regions are also consistent with this conclusion. Hence, polymorphisms among *t* haplotypes including lethality factors have accumulated during this short time period independently in each *M. musculus* subspecies.**

Mouse *t* haplotypes are variant forms of chromosome 17 that involve a large number of loci (1, 2). The *t* haplotype was originally detected in a natural population by its interaction with mutation Brachyury (*T*) to produce tailless mice (*T*/*t*). Most complete *t* haplotype mice have recessive lethality factors and die during embryonic development when homozygous. However, there are variations in embryonic lethalities among various wild mice and some of *t* haplotype hybrids bearing different lethality factors can survive. These *t* haplotype chromosomes are considered to belong to different complementation groups. Thus far, 16 complementation groups have been identified. Despite homozygous lethality, heterozygous male mice can transmit >95% of the *t* chromosomes to their progeny. Because of this bias in segregation, *t* chromosomes are found in high frequencies in natural populations and provide a well-known example of meiotic drive systems with segregation distorter and sex ratio in *Drosophila* (3). Transmission ratio distortion in *t* haplotypes is caused by four distorter loci (*Tcd-1*, *-2*, *-3*, and *-4*) and one responder locus (*Tcr*) (4). These loci are in strong linkage disequilibrium by recombination suppression that is presumably caused by four chromosomal inversions, *In(17)1* through *In(17)4* (5) in the *t* complex (Fig. 1*A*).

It has been suggested that *t* haplotypes were derived from a single common ancestral type (1–3, 5–7). To date the spreading of *t* haplotypes in wild *Mus musculus* populations and to understand the origin of *t* haplotypes, we have sequenced *Tcp-1* genes from *t* haplotype and wild mice and made the phylogenetic analysis. Mouse TCP-1 was first identified by Silver *et al.* (8) by two-dimensional gel electrophoresis. The *t* haplotype mice examined thus far exhibit a *t* haplotype-specific electromorph at the *Tcp-1* locus, TCP-1A, whereas most of inbred and wild-type mice exhibit a different electromorph, TCP-1B. Within *t* haplotypes, the extent of DNA polymorphism is high in the fourth inversion, which is likely to have resulted from occasional recombination between *t* haplotype and wild-type chromosomes. On the other hand, the extent of DNA polymorphism is low in the second inversion (9). Since the *Tcp-1* locus is located in the second inversion (10) and, presumably, frozen in recombination, it seems that the locus has useful information on the origin and evolution of *t* haplotypes. The TCP-1 polypeptide has high homology with "chaperonin" proteins (11), which are thought to function in protein folding and assembly. The gene is in fact highly conserved at the amino acid sequence level not only within mammals but also within eukaryotes, including *Drosophila* and yeast (11–13). We show that *t* haplotype-specific *Tcp-1* genes are unusual in terms of both silent and amino acid replacement substitutions and use this DNA sequence information to elucidate the origin of *t* haplotypes.**

MATERIALS AND METHODS

Mice and Rat. The taxonomy and geographical origins of wild and *t* haplotype mice used in this study are as follows (14, 15, 34): 129/Sv, *Mus musculus domesticus*; CAS, *Mus musculus castaneus* from Bogor (Indonesia); MOL, *Mus musculus molossinus* from Japan; MBT, *Mus musculus musculus* from Toshevo (Bulgaria); BAC, *Mus musculus bactrianus* from Afghanistan; XBS, *Mus spretoides* from Slantchev Briag (Bulgaria); ZBN, *Mus spicilegus* from Kranero (Bulgaria); SEI, *Mus spretus* from Ibiza (Spain); CRP, *Mus cervicolor* from Thailand; t^{w32}, *M. m. domesticus* from Clinton, Montana (U.S.A.); t^{w73}, *M. m. musculus* from South Jutland (Denmark); t^{Rin1}, *M. m. molossinus* from Omiya (Japan); t^{BAC}, *M. m. bactrianus* from Afghanistan (Table 1). The rat is the Fischer rat *Rattus norvegicus*.

PCR. DNA was extracted from mouse and rat liver specimens and amplified by 30 PCR cycles using T-18 (5′-AGCTGAATTCCCAGTGAAAGATAACACTGCATTGT-3′) and T-19 (5′-AGCTGGATCCCCCAATATTGTTT-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: mya, million years ago.

†To whom reprint requests should be addressed.

**The sequences reported in this paper have been deposited in the GenBank data base (accession nos. X61211–X61222).

Table 1. List of mice

Strain or *t* haplotype	Species or subspecies	Origin
129/Sv	*M. m. domesticus*	
CAS	*M. m. castaneus*	Indonesia (Bogor)
MOL	*M. m. molossinus*	Japan
MBT	*M. m. musculus*	Bulgaria (Toshevo)
BAC	*M. m. bactrianus*	Afghanistan
XBS	*M. spretoides*	Bulgaria (Slantchev Briag)
ZBN	*M. spicilegus*	Bulgaria (Kranero)
SEI	*M. spretus*	Spain (Ibiza)
CRP	*M. cervicolor*	Thailand
t^{w32}	*M. m. domesticus*	USA (Montana)
t^{w73}	*M. m. musculus*	Denmark (South Jutland)
t^{BAC}	*M. m. bactrianus*	Afghanistan
t^{Rin1}	*M. m. molossinus*	Japan (Omiya)

TATCCCATGCGT-3′) as primers for the 5′ region and T-35 (5′-AGCTGAATTCGTAAGTTACAGAGCAATGCAGAT-TGTG-3′) and T-37 (5′-AGCTGGATCCACAGCACCTC-CACCTGGGACCACAG-3′) as primers for the 3′ region. DNA amplifications were performed in 25-μl volumes of 10 mM Tris·HCl (pH 8.3) containing 2.5 mM $MgCl_2$, all four dNTPs (each at 0.2 mM), each primer at 0.25 μM, 100 ng of template DNA, and 1 unit of Taq polymerase. The temperature profile for 30 cycles of DNA amplification was 1 min at 93°C (denaturation), 1 min at 65°C (annealing of primers), 1 min 30 sec at 72°C (polymerization). The amplified DNA fragments were cut with *Bam*HI and *Eco*RI, fractionated by polyacrylamide gel electrophoresis, and visualized by ethidium bromide staining. They were cloned into M13 phage vectors and sequenced with an automated DNA sequencer (ABI 370A) over two independent clones for each strain or haplotype.

RESULTS AND DISCUSSION

We first cloned and partially sequenced the *Tcp-1* structural gene of the 129/Sv mouse (13). The sequenced region includes three exons coding for amino acids 274–407 (7) and contains two introns (Fig. 1*B*). We made oligonucleotides and amplified genomic DNAs from various mice by the PCR. Amplified DNA fragments were cloned into M13 phage and more than two independent clones were sequenced for each chromosome. To study the genealogical relationships of *Tcp-1* genes, we analyzed nine wild-type mice from various *Mus* species, four *t* haplotypes from different *M. musculus* subspecies, and one gene from rat as a reference. All mouse DNAs contain a B1 repetitive sequence (16) in intron 9 of *Tcp-1*, linked with a direct repeat of 15 nucleotides (5′-ATGC[T]CTAAATAC[T]CTT-3′) derived from the host-targeted sequence. Because the coding region is highly conserved and phylogenetically less informative (only one synonymous and nonsynonymous difference within the *Tcp-1* genes of four *t* haplotypes), we compared the sequences of introns 8 and 9 (610 base pairs) excluding the B1 repeat, which is absent in the rat *Tcp-1* gene.

The sequences of 14 *Tcp-1* genes are shown in Fig. 2 and the pairwise nucleotide differences in the two introns are given in Table 2 (see Fig. 3 for a reconstructed phylogenetic tree). The sequences of $Tcp\text{-}1^a$ genes carried by the *t* haplotypes are very different from those of $Tcp\text{-}1^b$ genes in wild-type *M. musculus* mice. The mean pairwise distance per site between $Tcp\text{-}1^a$ and $Tcp\text{-}1^b$ is 0.023 ± 0.010 at the synonymous sites, 0.014 ± 0.005 at the nonsynonymous sites, and 0.026 ± 0.006 in the intron region (Table 3). This shows that the synonymous sites and introns have evolved at nearly the same rate and that the degree (*f*) of selective constraints against the nonsynonymous sites is ≈60%. The *f* = 0.6 is unexpectedly high in contrast to the well-conserved amino acid sequence of TCP-1 polypeptide (11–13). Moreover, comparison of either *t* haplotype or wild-type sequences with the rat homologue indicates that the synonymous sites have actually evolved more than twice as fast as the intron region. In other words, compared with the substitution rate in the intron region, the synonymous rate is faster between the *t* haplotypes and wild types but slower between mice and rat. This accelerated synonymous rate in mice may be related to the differences in base compositional biases at the third codon positions. In particular, the T+C content is significantly different between the mouse and rat genes. In mice, thymidine is 36.9% and cytidine is 12.7%, whereas in rat thymidine is 31.6% and cytidine is 16.5%. Interestingly, however, there is no conspicuous difference in the intron region (thymidine is 35.5% vs. 36.4% and cytidine is 15.6% vs. 13.9%). This discordance suggests that base compositional bias at the third codon positions is not due to changes in mutation pressure. Rather, some relaxation of the selective constraint against synonymous changes of mouse *Tcp-1* genes appears to be more responsible, and in fact the synonymous rate (11.0 ± 1.0 per site per 10^9 year) is as high as that of a processed pseudogene of *Tcp-1* (19).

An advantage of using the intron sequences to estimate the divergence between the *t* haplotype and wild type is that the region is much longer than the synonymous sites compared. We first used the relative rate test and neighbor-joining method (20, 21) to examine whether the intron sequences of *Tcp-1* have evolved regularly in time. Since we obtained evidence for a molecular clock (data not shown), we assumed

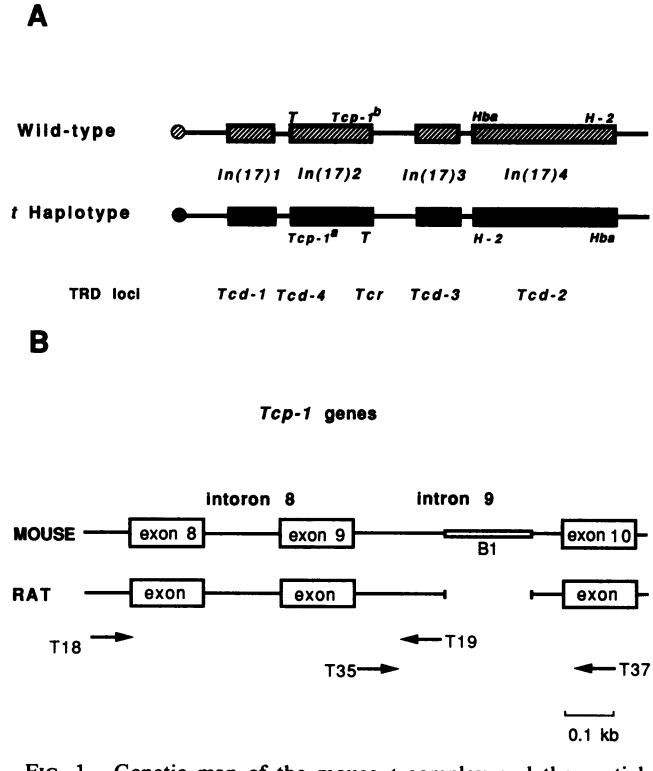

FIG. 1. Genetic map of the mouse *t* complex and the partial structure of *Tcp-1* genes. (*A*) Genetic map of chromosome 17 of wild-type and *t* haplotype mice is shown. The four chromosomal inversions [*In(17)1–4*] are shown by the shaded or solid boxes. The symbols *T*, *Tcp-1*, *Hba*, and *H-2* represent Brachyury, t-complex polypeptide-1, hemoglobin α pseudogene 4, and histocompatibility 2, respectively. The five loci involved in transmission ratio distortion are also shown. (*B*) Partial structure of the *Tcp-1* genes of mouse and rat, showing the intron–exon organization of the region used for the sequence comparison. Three exons are shown as open boxes. The mouse B1 repeat is shown as a narrow box. The PCR primers are indicated by arrows.

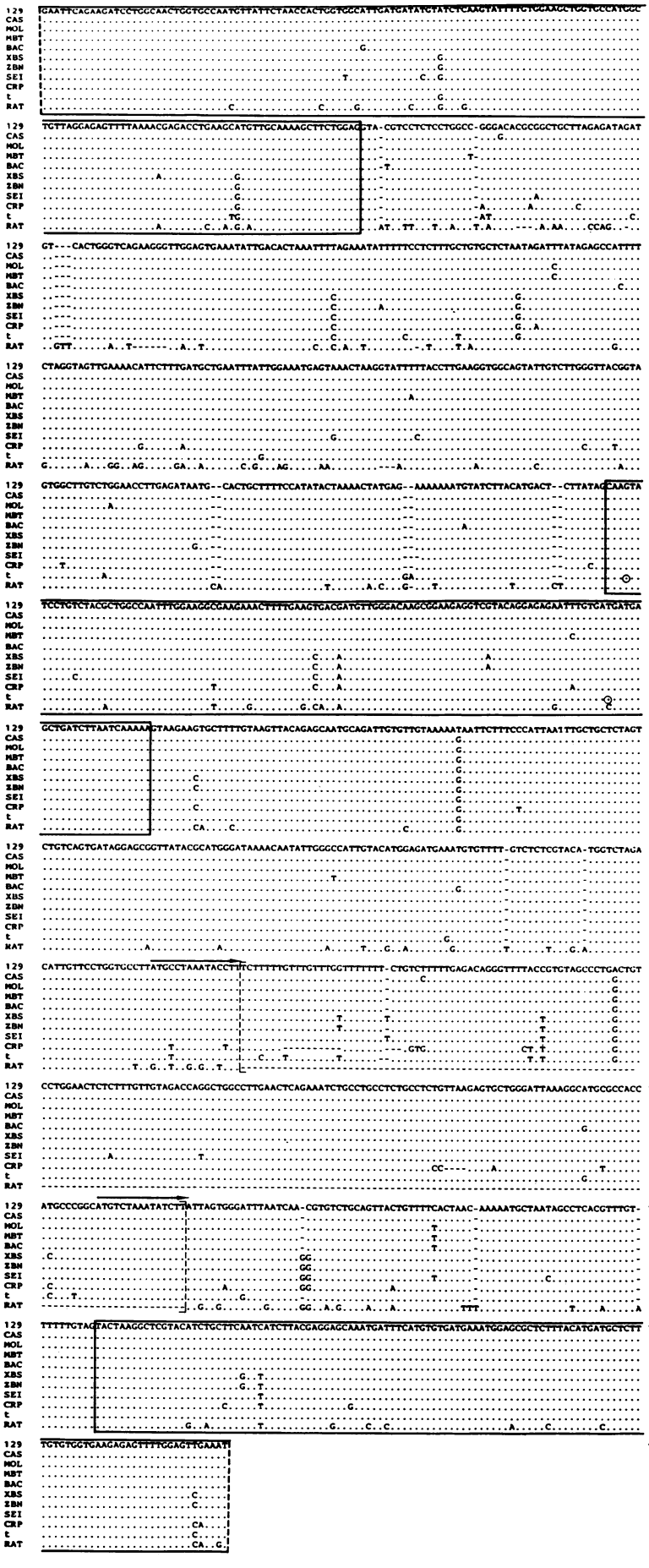

FIG. 2. *Tcp-1* DNA sequences from *Mus* and *Rattus* species. Nucleotide sequences (1231 base pairs) of three exons coding for amino acid 274–407 and two introns of mouse and rat *Tcp-1* genes are shown. The exon sequences are boxed. Taxonomies of wild and *t* haplotype mice are as follows: 129, *M. m. domesticus* (129/Sv); CAS, *M. m. castaneus*; MOL, *M. m. molossinus*; MBT, *M. m. musculus*; BAC, *M. m. bactrianus*; XBS, *M. spretoides*; ZBN, *M. spicilegus*; SEI, *M. spretus*; CRP, *M. cervicolor*; t^{w32}, *M. m. domesticus*; t^{w73}, *M. m. musculus*; t^{Rin1}, *M. m. molossinus*; t^{BAC}, *M. m. bactrianus*. The t^{w32}, t^{w73}, and t^{Rin1} complement each other concerning their lethality factors. The complementation group for t^{BAC} has not been determined. The *Tcp-1* sequence of rat (RAT) is from the Fischer rat *Rattus norvegicus*. All four *t* haplotypes have an identical sequence except at two nucleotide positions (nucleotide 498 in t^{w32} and nucleotide 595 in t^{BAC}, circled) where the t^{w32} haplotype has G → C and the t^{BAC} has T → C substitutions, respectively. The former is nonsynonymous change (Ser → Thr) and the latter is synonymous one. Dots indicate nucleotides identical with 129/Sv (*M. m. domesticus*) and dashes are deletions. The mouse-specific B1 repeat is in dotted brackets. Direct repeats of the targeted sequence for the B1 repeat insertion are shown by arrows.

Table 2. Number of nucleotide differences and percent divergence of *Tcp-1* introns among 14 mouse and a rat species

	129	CAS	MOL	MBT	BAC	XBS	ZBN	SEI	CRP	t^{w32}	t^{w73}	t^{BAC}	t^{Rin1}	Rat
129	—	2	4	6	6	5	7	9	19	13	13	13	13	81
CAS	0.3	—	4	6	6	5	7	9	19	13	13	13	13	81
MOL	0.7	0.7	—	4	6	7	9	9	21	15	15	15	15	83
MBT	1.0	1.0	0.7	—	8	9	11	11	23	17	17	17	17	85
BAC	1.0	1.0	1.0	1.3	—	9	11	11	23	17	17	17	17	83
XBS	0.8	0.8	1.2	1.5	1.5	—	2	6	14	12	12	12	12	78
ZBN	1.2	1.2	1.5	1.8	1.8	0.3	—	8	16	14	14	14	14	80
SEI	1.5	1.5	1.5	1.8	1.8	1.0	1.3	—	20	16	16	16	16	84
CRP	3.2	3.2	3.5	3.9	3.9	2.3	2.7	3.4	—	22	22	22	22	84
t^{w32}	2.2	2.2	2.5	2.8	2.8	2.0	2.3	2.7	3.7	—	0	0	0	85
t^{w73}	2.2	2.2	2.5	2.8	2.8	2.0	2.3	2.7	3.7	0	—	0	0	85
t^{BAC}	2.2	2.2	2.5	2.8	2.8	2.0	2.3	2.7	3.7	0	0	—	0	85
t^{Rin1}	2.2	2.2	2.5	2.8	2.8	2.0	2.3	2.7	3.7	0	0	0	—	85
Rat	14.6	14.6	15.0	15.4	15.0	14.0	14.4	15.2	15.2	15.4	15.4	15.4	15.4	—

Tcp-1 intron sequences of 610 base pairs were compared. The observed numbers of substitutions are indicated above the diagnonal. Nucleotide divergences (%) are below the diagonal, corrected by the Jukes and Cantor method (17). The wild-type strains and *t* haplotypes compared are described in Fig. 2.

that the divergence time between mouse and rat is 17 million years ago (mya) (22) and calibrated the mean divergence rate of the *Tcp-1* intron to be 5×10^{-9} per site per year (23). The phylogenetic tree of these intron sequences also concurred well with that of mitochondrial DNAs (24) and of biochemical markers (25). Using this rate, we obtained that $Tcp\text{-}1^a$ and $Tcp\text{-}1^b$ lineages diverged 2.9 ± 0.7 mya. The divergence time is as old as that between *M. cervicolor* and other *Mus* species. It is consistent with the result of a phylogenetic analysis of four cutter restriction sites in the α-globin pseudogene (Hb-4ψ) located in *In(17)4* (3). The $Tcp\text{-}1^a$ intron sequences of *t* haplotypes are not closely related to the *Tcp-1* intron sequences from any other *Mus* species including *M. spretus*, *M. spicilegus*, *M. spretoides*, and *M. cervicolor* (Table 1 and Fig. 3).

Surprisingly, despite a number of substitutions within wild-type mice, all *t* haplotypes have an identical *Tcp-1* intron sequence (Fig. 2). With the above substitution rate of introns, we computed a probability that no nucleotide change has occurred among the four *t* haplotypes since they first diverged. Although this probability depends on the ancestral relationships of these *t* haplotypes, we can use a simple formula for the case of two genes to make an estimate of the divergence time. Such an estimate is conservative because we assume implicitly that the other two *t* haplotypes have diverged recently. If the *t* haplotype had evolved before the divergence of *M. musculus*, the probability that no substitution has occurred among two independently evolving *Tcp-1* introns becomes extremely small. Under the assumption of the Poisson process of substitution with a mean rate of 5×10^{-9} per site per year, the divergence time of two orthologous *Tcp-1* introns must be <0.8 million years with 99% confidence. If on the other hand we assume that all four *t* haplotypes simultaneously diverged, the divergence time must be <0.4 million years with 99% confidence. In this case, it must be later than the divergence of the *M. musculus* subspecies (1 mya; see Fig. 3). We therefore conclude that the *t* haplotypes have most likely had a single recent origin and spread rapidly in most of *M. musculus* subspecies (introgression), including *M. m. castaneus* (M.S., unpublished data).

Table 3. Pairwise distance per nucleotide site

Sequence	Within groups	Between groups
Synonymous sites (90)	$\pi_T = 0.0056 \pm 0.0056$	$d_{TW} = 0.023 \pm 0.010$
	$\pi_W = 0.0344 \pm 0.0131$	$d_{RT} = 0.379 \pm 0.079$
		$d_{RW} = 0.372 \pm 0.077$
Nonsynonymous sites (309)	$\pi_T = 0.0016 \pm 0.0016$	$d_{TW} = 0.014 \pm 0.005$
	$\pi_W = 0.0110 \pm 0.0037$	$d_{RT} = 0.011 \pm 0.006$
		$d_{RW} = 0.014 \pm 0.005$
Introns (610)	$\pi_T = 0 \pm 0$	$d_{TW} = 0.026 \pm 0.006$
	$\pi_W = 0.0161 \pm 0.0034$	$d_{RT} = 0.154 \pm 0.017$
		$d_{RW} = 0.148 \pm 0.017$

Sampling errors are calculated as described by Takahata and Tajima (18). Within groups, the mean pairwise distance (π_X) within a group, where the subscript indicates either *t* haplotype (T) or wild type (W); between groups, the mean pairwise distance (d_{XY}) between groups *X* and *Y*, which are mouse wild type (W), mouse *t* haplotype (T), or rat (R). The number of base pairs examined is in parentheses. Data are the mean ± the sampling error.

The four *t* haplotype chromosomes used here were derived from different *M. musculus* subspecies. The t^{w32} of *M. m. domesticus* occurs in Western Europe and America, whereas the t^{w73} of *M. m. musculus* occurs in Eastern Europe, North Asia, and Eastern Siberia (27). The t^{Rin1} of *M. m. molossinus* and the t^{BAC} of *M. m. bactrianus* were isolated from Japan and Afghanistan, respectively. Among these four *t* haplotypes, three (t^{w32}, t^{w73}, and t^{Rin1}) were shown to complement each other for their lethality factors (t^{BAC} was not characterized) so that the four *t* haplotypes consist of at least three complementation groups. In addition, many *t* haplotypes thus far isolated from *M. m. domesticus* show various complementation groups (28) and they do complement the t^{w73} in *M. m. musculus*. Even in the same subspecies, some *t* haplotypes do have different lethality factors (low allelism rate). From these observations we suppose that lethality factors have accumulated after introgression to respective *M. musculus* subspecies and that the introgression must have been such a rapid process that most of the lethality factors could not be shared by different subspecies. Such introgression in the worldwide scale would have been impossible by natural migration alone. It therefore supports the hypothesis that the spreading of *t* haplotypes has been facilitated by modern human activities. This might have occurred as recently as 7000–9000 years ago (7, 29) around which agriculture and trade routes became disseminated. Polymorphisms among *t* haplotypes, if present, must have resulted from independent mutations after the introgression.

From these results and previous findings (1, 5–7, 15, 30, 31), no doubt the second inversion of chromosome 17 had played an important role in the evolution of *t* haplotypes. It has allowed the accumulation of various *t* haplotype characteristics and prevented them from their disassociation by recombination. The inversion might have occurred in the lineage leading to the present-day "wild-type" chromosomes (32) in *M. musculus* subspecies. The mouse population having the second inversion was probably small, facilitating the fixation of the inversion. We speculate that an ancestor

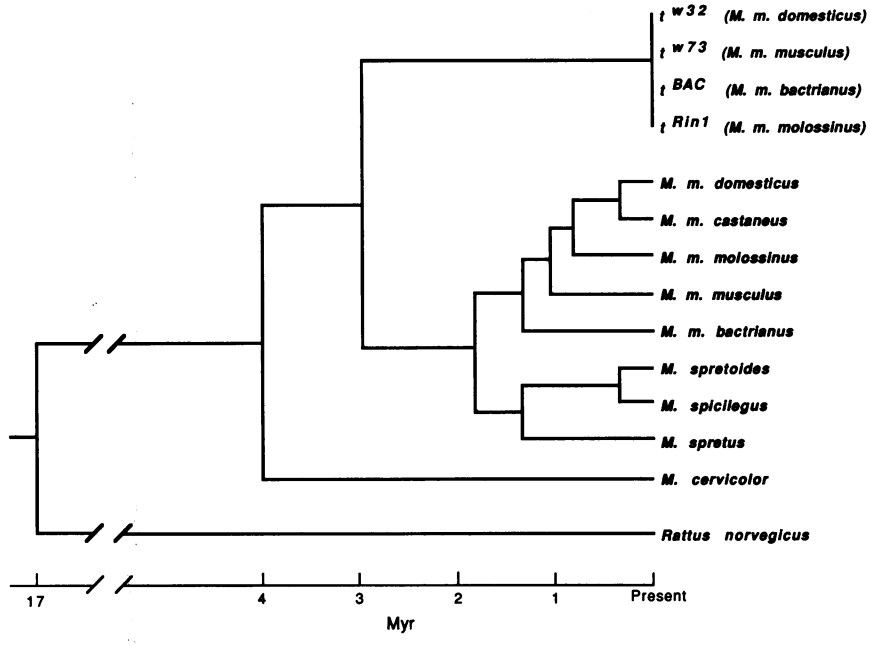

FIG. 3. Evolutionary tree for 13 *Mus* species and a *Rattus* species based upon their *Tcp-1* intron DNA sequences. The tree is constructed by UPGMA (unweighted pair-group clustering) (26) with multiple hit correction by Jukes and Cantor's method (17). Provided that the species divergence time between mouse and rat is 17 mya (22), a common ancestor of four independent *t* haplotypes was estimated to have existed 2.9 ± 0.7 mya.

of *t* haplotype mice originated 2.9 ± 0.7 mya. Around this time, the ancestral *t* haplotype mice might have survived as homozygotes, provided that recessive lethality factors had not yet accumulated. A high transmission ratio (>95%) and homozygote sterility [e.g., *Tcd-1* (33) of *t* haplotype] might be simultaneously acquired relatively recently. We suggest that the acquisition was no earlier than 0.8 mya and around the same time the *t* haplotype had started to spread over various populations of *M. musculus* subspecies. During this phase, the original population consisting exclusively of *t* haplotypes might have become extinct due to the homozygote sterility at the expense of a high transmission ratio (33), whereas the *t* haplotype in introgressing populations has survived as heterozygotes and been driven by a transmission ratio distortion. Evolutionary characterization of inversions other than the second will further elucidate the evolutionary process of the *t* haplotype in mice.

We thank Drs. K. Moriwaki and T. Shiroishi (National Institute of Genetics, Mishima, Japan) and F. Bonhomme (Université Montepellier, France) for their kind gifts of mouse strains. This work is supported by a grant from the Japanese Ministry of Education, Science and Culture of Japan (no. 01618508).

1. Silver, L. M. (1985) *Annu. Rev. Genet.* **19,** 179–208.
2. Klein, J. (1986) *Natural History of the Major Histocompatibility Complex* (Wiley, New York).
3. Wu, C.-I. & Hammer, M. F. (1991) *Evolution at the Molecular Level* (Sinaur, Sunderland, MA), pp. 177–203.
4. Lyon, M. (1984) *Cell* **37,** 621–628.
5. Hammer, M. F., Schimenti, J. & Silver, L. M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 3261–3265.
6. Nizetic, D., Figueroa, F. & Klein, J. (1984) *Immunogenetics* **19,** 311–320.
7. Willison, K. R., Dudley, K. & Potter, J. (1986) *Cell* **44,** 727–738.
8. Silver, L. M., Artzt, K. & Bennett, D. (1979) *Cell* **17,** 275–284.
9. Erhart, M. A., Phillips, S. J. & Nadeau, J. H. (1988) *Curr. Top. Microbiol. Immunol.* **137,** 70–76.
10. Artzt, K., Barlow, K., Dove, W. F., Lindahl, K. F., Klein, J., Lyon, M. F. & Silver, L. (1991) *Mamm. Genome* **1,** 5–29.
11. Gupta, R. S. (1990) *Biochem. Int.* **20,** 833–841.
12. Ursic, D. & Culbertson, M. (1991) *Mol. Cell. Biol.* **11,** 2629–2640.
13. Kubota, H., Morita, T., Nagata, T., Takemoto, Y., Nozaki, M., Gachelin, G. & Matsushiro, A. (1991) *Gene* **105,** 269–273.
14. Bonhomme, F. (1986) *Curr. Top. Microbiol. Immunol.* **127,** 19–34.
15. Delarbre, C., Kashi, Y., Boursot, P., Beckmann, J. S., Kourilsky, P., Bonhomme, F. & Gachelin, G. (1988) *Mol. Biol. Evol.* **5,** 120–133.
16. Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A. & Georgiev, G. P. (1980) *Nucleic Acids Res.* **8,** 1201–1215.
17. Jukes, T. H. & Cantor, C. H. (1969) *Mammalian Protein Metabolism* (Academic, New York), pp. 21–132.
18. Takahata, N. & Tajima, F. (1991) *Mol. Biol. Evol.* **8,** 494–502.
19. Kubota, H., Morita, T., Satta, Y., Nozaki, M. & Matsushiro, A. (1992) *Mamm. Genome* **2,** 246–251.
20. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46,** 573–639.
21. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
22. Miyata, T., Hayashida, H., Kikuna, R., Hasegawa, M., Kobayashi, M. & Koike, K. (1982) *J. Mol. Evol.* **19,** 28–35.
23. Li, W.-H., Luo, C.-C. & Wu, C.-I. (1985) *Molecular Evolutionary Genetics* (Plenum, New York), pp. 1–94.
24. Ferris, S. D., Sage, R. D., Prager, E. M., Ritte, U. & Wilson, A. C. (1984) *Genetics* **105,** 681–721.
25. Bonhomme, F., Catalan, J., Davidian, J.-B., Chapman, V. M., Moriwaki, K., Nevo, E. & Thaler, L. (1984) *Biochem. Genet.* **22,** 275–303.
26. Sakal, R. R. & Sneath, P. H. A. (1963) *Principles of Numerical Taxonomy* (Freeman, New York).
27. Ruvinsky, A., Polyakov, A., Agulnik, A., Tichy, H., Figueroa, F. & Klein, J. (1991) *Genetics* **127,** 161–168.
28. Klein, J., Sipos, P. & Figueroa, F. (1984) *Genet. Res.* **44,** 39–46.
29. Sokal, R. R., Oden, N. L. & Wilson, C. (1991) *Nature (London)* **351,** 143–145.
30. Klein, J., Golubic, M., Budimir, O., Schopfer, R., Kasahara, M. & Figueroa, F. (1986) *Curr. Top. Microbiol. Immunol.* **127,** 239–246.
31. Silver, L. M., Hammer, M., Fox, H., Garrels, J., Bucan, M., Herrmann, B., Frischauf, A.-M., Lehrach, H., Winking, H., Figueroa, F. & Klein, J. (1987) *Mol. Biol. Evol.* **4,** 473–482.
32. Bullard, D. C. & Schimenti, J. C. (1990) *Genetics* **124,** 957–966.
33. Lyon, M. F. (1986) *Cell* **44,** 357–363.
34. Moriwaki, K., Miyashita, N., Suzuki, H., Kurihara, Y. & Yonekawa, H. (1986) *Curr. Top. Microbiol. Immunol.* **127,** 55–61.