Genet. Res., Camb. (1996), **67**, *pp.* 1–10 *With 6 text-figures*

Recent evolution of mouse *t* haplotypes at polymorphic microsatellites associated with the *t* complex responder (*Tcr*) locus

KRISTIN G. ARDLIE[1] AND LEE M. SILVER

Department of Ecology and Evolutionary Biology and the Department of Molecular Biology, Princeton University, Princeton, NJ 08544

(*Received 10 July 1995 and in revised form 3 October 1995*)

Summary

Microsatellites closely associated with each member of the *Tcp10* gene family were amplified simultaneously from *t* haplotype and wild-type forms of mouse chromosome 17, by PCR. The *t* complex responder (*Tcr*) locus, which plays a central role in transmission ratio distortion, maps within the *Tcp10* cluster on the *t* haplotype. Thus the amplified set of microsatellite loci (referred to collectively as Tcp10ms) provides a direct marker for this central component of the meiotic drive system associated with all naturally occurring *t* haplotypes. A unique Tcp10ms pattern of microsatellite alleles was obtained for a number of independent, laboratory-maintained complete and partial *t* haplotypes. Independent *t* chromosomes found in wild mice from US populations also had unique patterns, even when they were classified within the same lethal complementation group. Wild and laboratory chromosomes in the t^{w5} group showed similarly-sized but non-identical Tcp10ms patterns, suggesting they share a recent common ancestor. These chromosomes are likely to have derived from an ancestral chromosome within the founding population of North American house mice. The Tcp10ms pattern was also shown to be useful in field studies for distinguishing among independent *t* haplotypes, when more than one is present within a single population.

1. Introduction

t haplotypes are geographically widespread variants of the proximal third of chromosome 17 that occur in natural populations of house mice (Silver, 1985). Complete *t* haplotypes are typified by their capacity for meiotic drive, or Transmission Ratio Distortion (TRD). Although segregation in females is normal, males heterozygous for a wild type (+) and a *t* haplotype (*t*) form of chromosome 17 will transmit the *t* haplotype to greater than 95% of their offspring in a clear departure from Mendelian segregation (Dunn, 1957).

The genetic basis for *t*-specific transmission ratio distortion is complex, and results from the interactions of a number of *cis* and *trans*-acting linked distorter loci (*Tcd*'s) and a haploid-acting responder locus (*Tcr*) (Lyon, 1984; Lyon, 1986; Silver & Remis, 1987). The structural integrity of these loci is maintained through the suppression of recombination along the length of *t* haplotypes, by a series of 4 major inversions (Herrmann, Barlow & Lehrach, 1987; Hammer, Schimenti & Silver, 1989). Collectively these span almost the entire *t* DNA and result in a 50–100-fold reduction in normal levels of recombination. Each of the inverted regions carries one or more of the loci required for maximal expression of TRD (Hammer, Schimenti & Silver, 1989).

A candidate gene for *Tcr*, *Tcp10b*^{t}, has been characterized in the *Tcp10* gene family. Members of this gene family are present in 2 to 4 copies on both homologues of chromosome 17 from different strains of mice (Schimenti *et al.* 1987; Schimenti *et al.* 1988; Bullard & Schimenti, 1990). *Tcp10b*^{t} produces a unique, haploid-specific, alternatively spliced mRNA in elongating spermatids, consistent with it's putative role as *Tcr*^{t} (Cebra-Thomas *et al.* 1991).

The high transmission bias in favour of *t* haplotypes is expected to lead to their rapid fixation in mouse populations; however, TRD is counterbalanced both by the complete sterility of *t*/*t* homozygous males, and by the recessive lethal mutations carried by most complete *t* haplotypes. Thus *t* haplotypes are maintained as a polymorphism in natural populations (Petras, 1967; Figueroa *et al.* 1988; Lenington, Franks & Williams, 1988). *t* haplotypes express no obvious phenotype in wild mice, and only with the recent

[1] Current address, and address for reprint requests: Kristin Ardlie, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts, 02138.

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

availability of molecular probes that identify restriction fragment length polymorphisms (RFLPs) between *t* and wild-type DNA, has it become possible to definitively characterize the genotypes of large numbers of wild mice.

Phylogenetic analyses of wild-type and *t* haplotype forms of chromosome 17 indicate that all contemporary *t* haplotypes can be traced back to a common ancestral chromosome that may have existed as recently as 10000 years before present (B.P.), and not more than 100000 years B.P. (Willison, Dudley & Potter, 1986; Morita *et al.* 1992; Hammer & Silver, 1993). This recent common ancestry is further supported by the observation of identical *t*-specific protein polymorphisms and by a scarcity of RFLPs among different *t* haplotypes (Silver *et al.* 1983; Fox *et al.* 1985; Schimenti *et al.* 1987; Howard *et al.* 1990; Horiuchi *et al.* 1992). Thus, although many molecular markers have been characterized that detect differences between wild-type and *t* chromosomes, in nearly all cases, these markers do not provide a means for distinguishing among the different *t* haplotypes.

One exception to this rule is provided by the Bb40 marker which detects two different *t*-specific RFLP patterns based on the presence or absence of a Tcp10 pseudogene named *Tcp10ps* (Schimenti *et al.* 1987; Horiuchi *et al.* 1992; Pilder *et al.* 1992). Another more extensive set of inter-*t* polymorphisms is detected by the TSE marker (Uehara *et al.* 1990). Unfortunately, the Bb40 polymorphism is bi-allelic and only divides the complete set of *t* haplotypes into two broad groups. Although the TSE polymorphism uniquely identifies individual *t* haplotypes, its usefulness is limited by the requirement for very high molecular weight DNA that must be analysed by pulsed field gel electrophoresis, and it is rarely feasible to use this technique for samples collected in the field, in particular. Thus the different *t* haplotypes have remained largely distinguishable only on the basis of their traditional lethal complementation groups.

Microsatellites are regions of DNA composed of short simple sequences repeated in tandem, such as dinucleotide repeats (e.g. $(CA)_n$) (Tautz, 1993). They are abundant and widespread in eukaryotic DNA (Hamada, Petrino & Kakunaga, 1982; Tautz & Renz, 1984), and show a high degree of polymorphism in length variation of repeat units. Because microsatellite polymorphisms can be detected rapidly by PCR, they provide excellent markers for genetic analysis (Weber & May, 1989; Love *et al.* 1990; Dietrich *et al.* 1992). They are also extremely useful in studies of variation at the level of populations, even in species where other genetic techniques reveal little to no detectable genetic polymorphisms (Bruford & Wayne, 1993; Hughes & Queller, 1993).

Dinucleotide repeats have also proven useful in finding polymorphisms among *t* haplotypes. In contrast with the lack of polymorphism between *t* haplotypes for most RFLP probes, many of the dinucleotide repeats that have been studied have shown some polymorphism (Uehara *et al.* 1990; Ebersole, Lai & Artzt, 1992; Lai & Artzt, 1992). They not only detect differences between *t* haplotype and wild type chromosomes, but have also been shown to differentiate among some of the different *t* haplotypes (Lai & Artzt, 1992).

To find polymorphic differences that distinguish among *t* haplotypes, primers were designed around a short microsatellite repeat in an intron of the *Tcp10b*^{t} gene. Because this gene is a candidate for the *t* complex responder (*Tcr*), this provides an important means for directly assessing inter-*t* variation at this locus, which is of central importance to the meiotic drive phenotype expressed by *t* haplotypes.

2. Materials and methods

(i) *Mice and DNA*

DNAs were all prepared from mice that had been maintained in the animal colony at Princeton University, or were gifts to M. F. Hammer as part of an earlier study (Hammer & Silver, 1993). Cosmid clones containing the *Tcp10a*^{t}, *Tcp10b*^{t}, and *Tcp10c*^{t} alleles were derived from a homozygous t^{w5} cell line cosmid library provided by Dr John Schimenti (Rosen *et al.* 1990).

(ii) *Wild caught animals*

Wild mice were live-trapped in various locations around North America. Tail biopsies were collected from wild caught animals and preserved in ethanol until being brought back to the laboratory. DNA was then isolated using a standard tail DNA preparation protocol (Hogan *et al.* 1994).

A number of *t* haplotypes from wild-caught mice were bred to 129/SvJ animals to control the contribution to the PCR pattern from the wild-type chromosome. The wild-caught *t* haplotype chromosomes tested here are: BM1 (a semi-lethal *t* haplotype trapped in New Jersey); MV1, MV3, and MV12 (three semi-lethals from a single population trapped elsewhere in New Jersey); PV1 (a semi-lethal *t* haplotype trapped in Southern Illinois); CF109 (a lethal *t* haplotype trapped in Ithaca, NY); and T57, T59, T67, and T68 (four lethal *t* haplotypes collected from a population in Tennessee). All *t* haplotype-bearing mice were also confirmed independently, both by breeding assays, and by Southern blot analysis using the markers Bb40 and Tu119 which distinguish RFLPs between *t* haplotype and wild type forms of chromosome 17 (Herrmann *et al.* 1986; Schimenti *et al.* 1987).

In breeding analyses, CF109, and all of the Tennessee chromosomes failed to produce viable homozygotes, indicating that all carried recessive lethal alleles (data not shown). Furthermore, all except T68 failed to complement the t^{w5} chromosome,

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

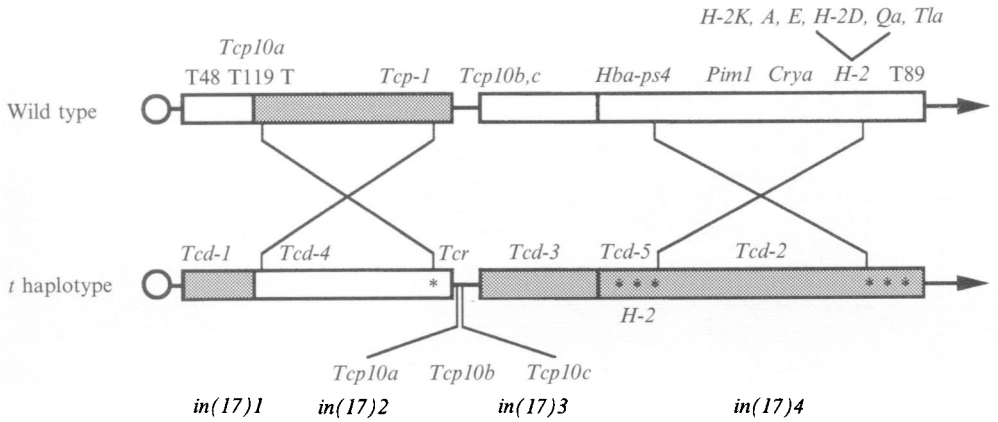

Fig. 1. Schematic representation of *t* haplotype and wild type forms of chromosome 17. Shaded boxes represent the four *t*-associated inversions, *in(17)1* through *in(17)4*. The loci that play a role in the TRD phenotype are the 5 distorter loci, *Tcd-1* through *Tcd-5*, and the responder locus, *Tcr*, which has been mapped to the *Tcp10b* gene on the *t* haplotype. A number of DNA markers and loci are shown in their relative positions on the wild type chromosome. The *t*-associated lethal mutations are indicated by the stars (Silver, 1993).

and therefore carry the t^{w5} lethal mutation. T68 did complement with both a number of the other chromosomes, and t^{w5}, indicating that it contains a different undetermined lethal allele.

(iii) *PCR analysis*

PCR primers were designed in unique sequence flanking a $(CA)_{34}$ microsatellite located in the intron following the first amino acid coding exon of the *Tcp10b*^{t} gene from t^{w5} (Bullard & Schimenti, 1991). The complete amplification pattern obtained with this primer pair is referred to as Tcp10ms (for *Tcp10* family microsatellites). Unpublished sequence for this region was kindly supplied by Dr John Schimenti. The sequence of the forward and reverse primers are,

Tcp10ms F:
(5′-GCGTGCCCCTTGGACAGGG-3′)

and

Tcp10ms R:
(5′-GCTGTACTGTAACCTTGCTTAG-3′).

PCR reactions were performed both with and without radioactive end-labelling of one of the primers. The products were visualized initially on non-denaturing acrylamide gels and subsequently, for better resolution of alleles, radioactively-labelled, on denaturing sequencing gels. Approximately 50 ng of genomic DNA was amplified in a 10 μl PCR reaction using 0·25 units of *Taq* polymerase. 2·5 μM each of the two unlabelled primers were used, with an additional 0·75 μM of one of the primers end-labelled. The reactions were amplified for 25 cycles; at 95 °C for 25 s, 60 °C for 1 min, and 72 °C for 30 s. PCR products were electrophoresed on 7% denaturing polyacrylamide gels (SequaGel, National Diagnostics). Gels were dried and exposed to film at room temperature for 4–16 h. The amplified products ranged in size between ~ 250–300 bp. A sequencing reaction was used as a size marker to determine relative allele sizes.

(iv) *Sequence analysis*

Direct sequencing of the *Mus macedonicus* PCR product, to determine its microsatellite repeat length, was performed by reamplifying the initial product using one kinased primer and digesting with lambda exonuclease (Higuchi & Ochman, 1989). Primers were removed from the exonucleased reaction (Kreitman & Landweber, 1989), and the ssDNA was sequenced in both directions using the Sequenase Kit (USB). Reactions were electrophoresed on a 6% polyacrylamide gel. Gels were soaked in a 10% methanol/10% acetic acid bath, dried onto 3 mm Watman paper and autoradiographed.

3. Results

(i) *Simultaneous detection of all members of the* Tcp10 *gene family*

The primer pair defining the Tcp10 microsatellite was used to amplify products from the various DNA samples described in the Materials and Methods. The amplification pattern obtained with each whole genome sample (referred to as TCP10ms) was complex. This is because the primer pair simultaneously amplifies microsatellite alleles from all members of the *Tcp10* gene family, both on the *t* haplotype and the wild type forms of chromosome 17 (Fig. 1). Individuals may vary both in the number of *Tcp10* genes they contain, and in the sizes of the allelic products amplified from each of these genes.

To determine the correspondence between individual components of the complete TCP10ms pattern and particular *Tcp10* genes from the *t* form of chromosome 17, we amplified single alleles from individual cosmid clones containing the *Tcp10a*^{t} (CosA), the *Tcp10b*^{t} (CosB), or the *Tcp10c*^{t} (CosC) genes. The results obtained are shown beside the complete pattern obtained from the wild-type 129/SvJ

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

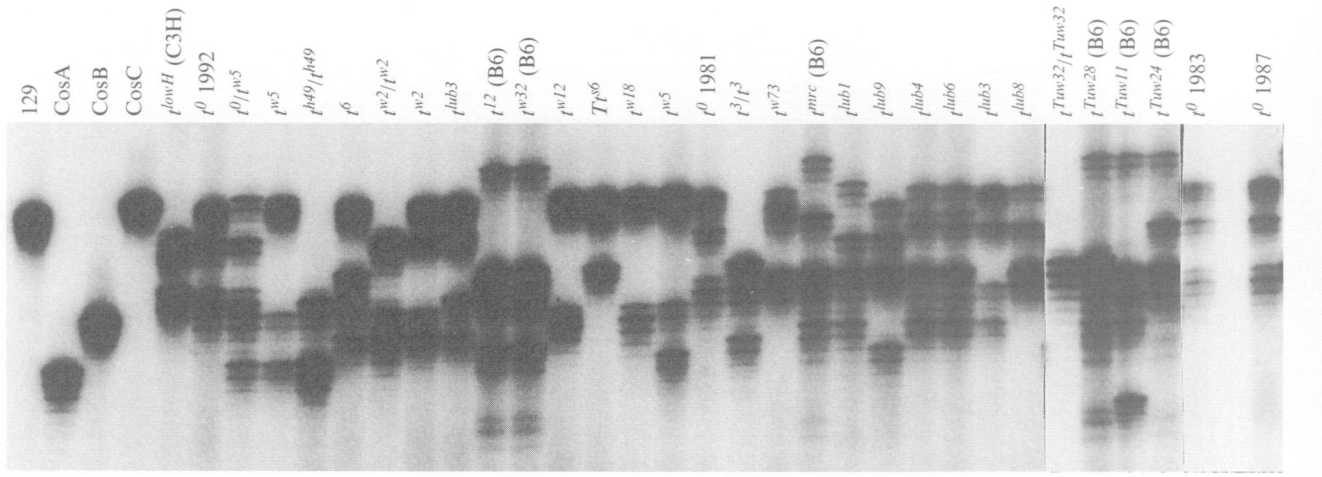

Fig. 2. Autoradiograph of PCR products obtained for various complete and partial *t* haplotypes. The independent *Tcp10* alleles of t^{w5} are shown in the cosmid clones A, B and C, beside the 129/SvJ wild type bands, and the 129/t^{w5} composite heterozygote animal is shown nearby for comparison. Most *t* haplotypes are shown against a 129 wild type chromosome unless otherwise indicated. The following chromosomes are partial *t* haplotypes: t^{lowH}, t^6, Tt^{s6}, t^3, t^{Tuw32}, t^{w18} (Committee for Mouse Chromosome 17, 1991). All others are complete *t* haplotypes. For allele sizes see Table 1.

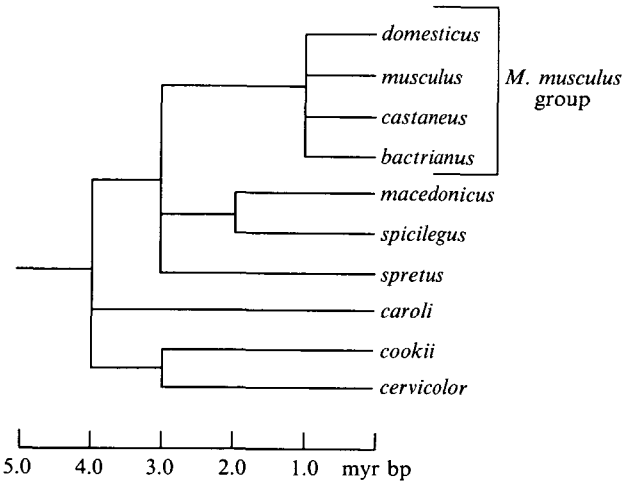

Fig. 3. Phylogenetic tree of the *Mus* species related to the *Mus musculus* group. *t* haplotypes are present in natural populations of all four subspecies of the *M. musculus* group, *domesticus*, *musculus*, *castaneus* and *bactrianus*. Modified from fig. 2.2 (Silver, 1995).

chromosome (Fig. 2). The cosmid clones were derived from the t^{w5} chromosome, and DNA from a 129/t^{w5} heterozygote mouse is shown near these for comparison. This individual shows the complex, multiple banding pattern of a single +/*t* heterozygote. With the exception of the *Tcp10c*^{t} associated band, the *t* haplotype *Tcp10* alleles are typically smaller in size (i.e. have fewer $(CA)_n$ repeats) than the wild type *Tcp10* alleles. This generally holds true for all *t* haplotypes collected from wild mice as well. *t*-bearing mice can typically, but not reliably, be identified as such by this characteristic.

(ii) *Species differentiation within the* Mus *genus*

In addition to amplifying alleles from the wild-type *M. m. domesticus* and *t* haplotype chromosomes, the Tcp10 primers also amplify products from other

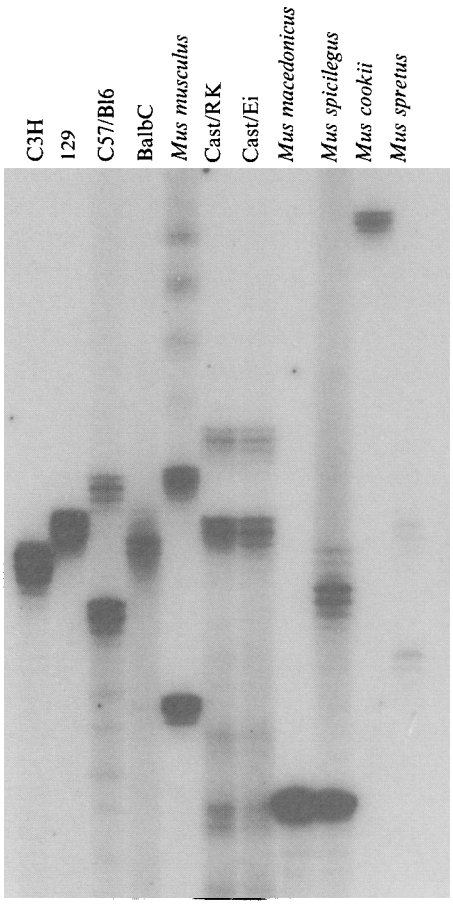

Fig. 4. Autoradiograph of PCR products of wild type *Tcp10* alleles obtained for four inbred strains of *M. m. domesticus* and a number of related species of *Mus*. The 2 weakly amplifying bands from *Mus spretus* can just be seen.

species and sub-species within the *Mus* species group (Fig. 3). In both *M. m. musculus*, and *M. m. castaneus*, at least two predominant PCR products are observed (Figs 4, 5) suggesting that *Tcp10* may have more than

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

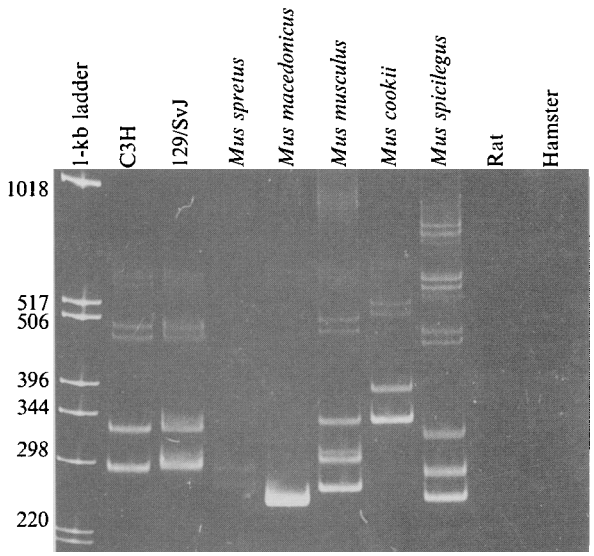

Fig. 5. Ethidium bromide stained, non-denaturing acrylamide gel of *Tcp10* PCR products, amplified from two inbred strains of *M. m. domesticus*, and a number of related species of *Mus*. The two very faint bands amplifying from *Mus spretus* can just be seen. *Mus macedonicus* clearly has the smallest PCR product, and is also the only PCR product without associated 'echo', or other larger, PCR artifact bands.

one family member in these subspecies as previously shown for *M. m. domesticus*. The Tcp10 primers also amplify products in some of the more distantly related *Mus* species. A single band was amplified from *Mus macedonicus*, and the same sized primary band, as well as some additional fainter bands, were amplified from *Mus spicilegus*. These 2 species have been shown by a number of molecular phylogenies to be sister groups (Lundrigan & Tucker, 1994). At 60 °C, two very faint bands can be amplified from *Mus spretus*. Nothing can be amplified from *Mus caroli* at this temperature, but a single band can be amplified from *Mus cookii*, which is more distantly related to the *Mus musculus* sub-group than is *Mus spretus*. No products could be amplified from either rat or hamster (Fig. 5), nor human DNA (data not shown). Reducing the annealing temperature to 50–55 °C didn't resolve any of these, and resulted in many non-specific bands. Both the weak bands of *M. spretus*, and the failure to amplify these other microsatellite alleles, may be due to base variation in the region of primer annealing (Callen *et al.* 1993).

Individual PCR amplified microsatellites visualized on agarose, or non-denaturing acrylamide gels, frequently show a single main product and an artifactual ladder of additional bands (Fig. 5). Some of these, particularly those near or below the size of the desired band, have been attributed to polymerase slippage (Litt & Lucy, 1989; Weber & May, 1989). This apparent strand slippage during PCR amplification is most common for dinucleotide microsatellite repeats (Beckman & Weber, 1992). Higher molecular weight artifactual bands, however, may be due to secondary structures formed within the sample, as these can usually be eliminated by denaturing the PCR products and running them on sequencing gels (compare Figs 4 and 5). When the PCR products here were initially run on non-denaturing gels, *Mus macedonicus* alone did not shown any echo or artifact bands, and had the smallest PCR product (see Fig. 5). Direct sequencing of this PCR product, to determine whether it's small size was due to fewer dinucleotide repeats, revealed that the number of dinucleotide repeats was in fact reduced from $(CA)_{34}$ to $(CA)_5$. This suggests that the 'echo' bands seen on these non-denaturing gels are internally derived and are a function of both the presence, and the number, of such dinucleotide repeats, and are probably conformational.

(iii) *Different complete* t *haplotypes display characteristic patterns*

A survey was performed on the Tcp10 microsatellite family contained within different laboratory-maintained complete and partial *t* haplotypes as shown in Fig. 2. This survey demonstrated distinct patterns of *Tcp10* microsatellite alleles for each independent *t* haplotype. Different *t* haplotypes are polymorphic with respect to the band sizes of any given *Tcp-10* microsatellite allele, as well as to the number of visible bands (Table 1). Thus, each *t* haplotype has a unique set of microsatellite bands, and it is possible to distinguish between almost all of the different *t* haplotypes tested.

Complete *t* haplotypes carry three functional *Tcp-10* genes, and some may carry a fourth copy as a pseudogene, while wild type chromosome 17 homologues may carry between two to three different functional *Tcp10* genes (Bullard & Schimenti, 1990; Pilder *et al.* 1992). Thus a +/*t* animal can potentially have up to 6 or 7 different Tcp10 amplification products. Rarely are so many bands visible, however. Presumably some alleles are not visible because they overlap in band size with other alleles. An example of this can be seen in the t^0/t^{w5} compound heterozygote (Fig. 2), where the *Tcp-10c*^{t} allele from t^{w5} can be seen clearly. This allele is only slightly larger than the *Tcp-10* alleles of the 129 chromosome, and is therefore obscured in the 129-t^{w5} heterozygote, which thus appears to have only two *t*-associated alleles.

There is also a certain amount of size polymorphism associated with any given gene in the *t* haplotype complex of *Tcp10*. The cosmid-amplified *Tcp10b*^{t} allele is derived from the t^{w5} chromosome, and can be seen clearly in the 129-t^{w5} individual. The partial *t* haplotype t^{lowH} (C3H background) is a partial *t* haplotype containing only the *t* form of the *Tcr*, or *Tcp10b*^{t} gene, and it is not the same size as the allele from t^{w5}. Similarly, the t^{Tuw32} chromosome, which is a naturally occurring partial *t* haplotype from Israel

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

Table 1. *Relative allele sizes and Tcp10ms band patterns of laboratory-maintained and wild-caught* t *haplotype chromosomes*

Genotype	Origin*	Allele pattern and size (bp)†
129/SvJ	Inbred	293, 292
C3H	Inbred	288
B6	Inbred	298, 278
129/t^{w5}‡	NY, USA	129/SvJ, 294 (*Tcp10c*), 274 (*Tcp10b*), 266 (*Tcp10a*)
129/t^{0}	Paris lab non-inbred	129/SvJ, 286, 278, 276
t^{w2}/t^{w2}	NY/PA, USA	286, 274, 270
B6/t^{12}	USA	B6, 280, 268
B6/t^{w32}	USA	B6, 280, 268
129/t^{w12}	USA	129/SvJ, 294, 274, 273
129/t^{w73}	Denmark	129/SvJ, 290, 278
129/t^{w18}	USA	129/SvJ, 294, 274, 272
129/t^{lub3}	Italy	129/SvJ, 286, 276, 270
129/t^{lub1}	Italy	129/SvJ, 284, 278, 270
129/t^{lub4}	Italy	129/SvJ, 288, 278, 270
129/t^{lub6}	Italy	129/SvJ, 288, 278, 270
129/t^{lub8}	Italy	129/SvJ, 286, 278
B6/t^{Tuw11}	Chile	B6, 274, 270, 258
B6/t^{Tuw24}	Germany	B6, 286, 278
t^{Tuw32}/t^{Tuw32}	Israel	280
129/T58	TN, USA	129/SvJ, 294, 278, 272
129/T5	TN, USA	129/SvJ, 294, 278, 272
129/T67	TN, USA	129/SvJ, 294, 278, 272
129/T68	TN, USA	129/SvJ, 294, 274, 272
129/CF109	NY, USA	129/SvJ, 294, 274, 272
BM1/BM1	NJ, USA	280, 276
PV1/PV1	IL, USA	280, 274, 264
MV1/MV1	NJ, USA	286, 274

* Data on origins of laboratory *t* haplotypes taken from Silver *et al.* (1987), and described in materials and methods for wild-caught chromosomes.
† May not represent all alleles, as identical or similar-sized alleles will appear as a single allele. Only chromosomes with distinct patterns in Fig. 2 are represented here.
‡ The exact size and identity of alleles is known for t^{w5} only.

(Silver *et al.* 1987), only has the *Tcp-10a*^{t} allele, and it too differs in size from the cosmid *Tcp10a*^{t} allele from t^{w5}.

A few *t* haplotypes are indistinguishable from one another. t^{lub4} and t^{lub6} produce identical PCR patterns. Both of these are complete *t* haplotypes originally derived from Italy (Silver *et al.* 1987), and it is therefore possible that they are related chromosomes. The chromosomes t^{w32} and t^{12} are also identical at all bands, further confirming that they are probably the same laboratory-derived *t* chromosome, as suggested previously by Artzt *et al.* (1985) and Uehara *et al.* (1990).

These primers additionally allow some fine scale mapping of some of the recombination breakpoints of partial *t* haplotypes. t^{h49} is a partial *t* haplotype derived from a recombination event near the *Tcr* locus, between the complete *t* haplotype, t^{w5} and a partial *t* haplotype (t^{lowH}) carrying the responder locus from the t^{6} haplotype (Lyon, 1984; Lyon & Zenthon, 1987). From the data presented here (Fig. 2), it can be seen that t^{h49}, has the *Tcp10a*^{t} allele from t^{w5} but has the *Tcp10b*^{t} allele (i.e. the responder form of the allele) from t^{6}. This is consistent with other recent findings showing that the crossover event occurred between these two members of the *Tcp10* gene family, which are adjacent to one another on the *t* haplotype chromosome (Bullard, Ticknor & Schimenti, 1992).

(iv) *Polymorphism among wild-type and* t *haplotype chromosomes in natural populations*

Wild-caught animals that are heterozygous for *t* haplotypes have uncharacterized, and potentially highly polymorphic, wild-type forms of chromosome 17. This variation was found to be sufficiently great that the resulting overall PCR pattern was too confusing to interpret directly in these animals. Although *t* haplotype-specific alleles are, on the whole, smaller in size than those from wild-type chromosomes, this distinction between *t* and wild-type chromosomes is less clear in wild animals. To overcome this problem, *t* haplotypes derived from a number of wild mice were bred opposite to the inbred 129/SvJ form of chromosome 17. With a defined

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

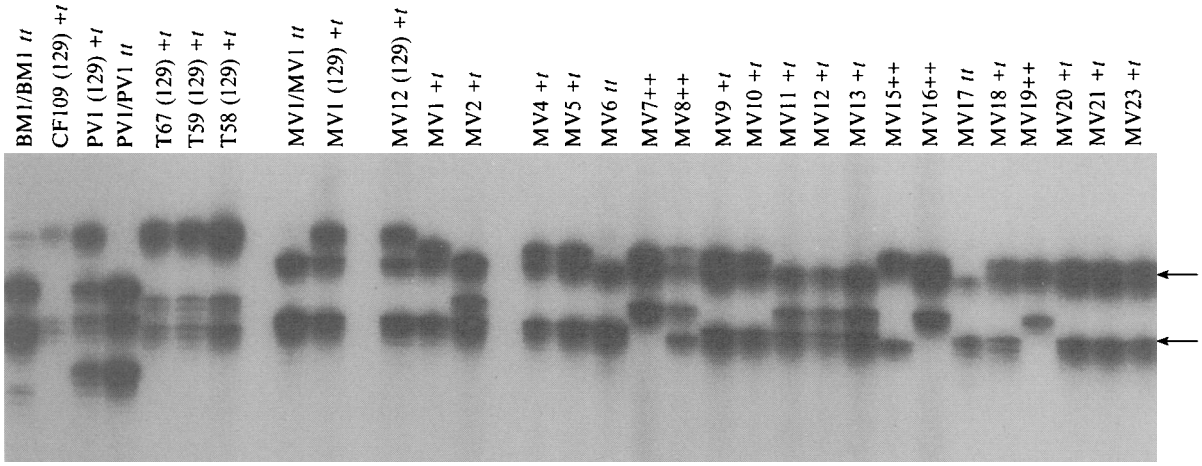

Fig. 6. Autoradiograph of PCR products obtained for a number of wild-caught *t* haplotypes. BM1 is a semi-lethal trapped in NJ, PV1 is a semi-lethal trapped in Southern IL, and MV1 through MV23 are mice trapped from a separate population in NJ, that contained the semi-lethal MV *t* haplotype. The band pattern of this *t* haplotype can be seen in the MV1/MV1 homozygote, and all mice carrying a *t* haplotype in this population can be seen to have these same two bands. The genotypes shown were determined independently by Southern blot analysis. CF109, T58, T67, and T59 are all lethal alleles in the t^{w5} complementation group, trapped in NY and TN. They are all shown here against a 129 wild-type chromosome.

wild-type contribution, it was then possible to distinguish those PCR products that derived from wild-caught *t* haplotypes.

None of the PCR patterns associated with wild-caught *t* haplotypes were identical to any of the laboratory-maintained *t* haplotypes tested (Figs. 2 and 6, and Table 1). Among the wild-derived lethal haplotypes, both the CF109 and three of the four Tennessee chromosomes (T58, T59, T67) were classified as belonging to the t^{w5} lethal complementation group. All of these haplotypes show a general similarity in banding pattern with the t^{w5} chromosome, however, none are identical to it. The alleles that appear to correspond to the *Tcp10a*^{t} and/or *Tcp10b*^{t} alleles differ in a number of repeat units from those in the t^{w5} chromosome. T58, T59 and T67 have an *Tcp10a*^{t} allele that is 6 bp larger, and a *Tcp10b*^{t} allele that is 4 bp larger than the corresponding products in t^{w5} respectively. The CF109 chromosome was trapped in Ithaca, NY, from the same region from which t^{w5} was originally collected. It has the same sized *Tcp10a*^{t} allele as the other T-chromosomes, but a *Tcp10b*^{t} allele that is 4 bp smaller and corresponds in size to the same allele in t^{w5}. Overall, these chromosomes are all relatively similar to one another and to the t^{w18} haplotype, which also came from an independent and now extinct chromosome carrying the t^{w5} lethal (Silver, 1983).

The recently trapped semi-lethal haplotypes (BM1, PV1, and MV1 through MV23), can also be seen to bear only a slight resemblance to the previously known semi-lethal haplotypes shown here (t^{w2}, t^{tub8} and t^{Tuw28}). Because of their phenotype of incomplete lethality, the semi-lethal alleles cannot be classified genetically into separate complementation groups, and are therefore considered as a single class.

Semi-lethals collected from different populations show quite different allele patterns. Semi-lethals derived from the same population, in contrast, have identical allele patterns (Fig. 6). Mice trapped from the MV population in New Jersey were known from breeding studies to contain a semi-lethal allele, however, it was impossible to know from breeding studies alone, whether they contained one or two such haplotypes. DNA's from all mice trapped from the entire population were then PCR amplified, to determine whether the population contained one, or more than one, semi-lethal and/or lethal *t* haplotype. All mice that carried a *t* haplotype in this population were found to be carrying the same one. The presence of any two independent *t* haplotypes could not have been discerned previously in the absence of breeding studies.

4. Discussion

All *t* haplotypes are descendants from a recent ancestral chromosome (Willison *et al.* 1986; Morita *et al.* 1992; Hammer & Silver, 1993). As a consequence of their close relationship to each other, there is minimal inter-*t* variation at the molecular level. This has made it difficult to distinguish different *t* haplotypes, particularly in wild populations, with the use of molecular markers alone. Here we show that a single pair of PCR primers designed around a microsatellite locus in the *Tcp-10b*^{t} gene can provide a means for distinguishing among nearly all well-characterized *t* haplotypes as well as among undefined *t* haplotypes from independent wild populations.

The TCP10ms primers allow the simultaneous amplification of a dinucleotide repeat unit found in all members of the *Tcp10* gene family which is present in

https://doi.org/10.1017/S0016672300033425 Published online by Cambridge University Press

two to four copies on both the wild type and *t* haplotype homologues of chromosome 17. The various members of the gene family are highly conserved relative to one another, which seems to be the result of the concerted evolution of an ancient gene family (Pilder *et al.* 1992). Despite the high level of identity of genes in this gene family (Davies & Willison, 1991), they display considerable variation in repeat length of the microsatellite allele among the different members. All of the inbred strains tested show characteristic size differences in their alleles, and there is a similarly high level of variation in the alleles contained on the *t* haplotype chromosomes.

The ability to uniquely distinguish all classically-defined *t* haplotypes has been demonstrated by only one other means, through clusters of *t*-haplotype specific elements (TSEs) unique to the *Tla* region of *t* haplotypes (Uehara *et al.* 1990). The variation observed at both the TSE and Tcp10 clusters uniquely identifies individual *t* haplotypes to an even greater extent than the only previously available method for distinguishing among *t* haplotypes, which is based on complementation analysis.

Uehara *et al.* (1990) showed that the extensive diversity in TSEs in independent *t* haplotypes was due to different numbers of clusters and different numbers of repeats in each cluster, which may reflect a pronounced instability of TSE clusters on an evolutionary time scale. Microsatellite alleles are similarly unstable and show relatively high mutation rates (e.g. 10^{-3}–10^{-5} (Weber & Wong, 1993; Ellegren, 1995), thus none of the diversity found in either of these highly mutable clusters is at odds with the finding that all *t* haplotypes share a recent common ancestor (Morita *et al.* 1992; Hammer & Silver, 1993).

The conservation of microsatellites between taxa has been demonstrated in a number of studies. Dinucleotide repeats in particular, are conserved by chromosomal location between closely related species (e.g. mouse and rat, and sheep and cattle) (Moore *et al.* 1991; Hino *et al.* 1993; Stallings, 1995), but not between more distantly related species (e.g. humans and rodents) (Stallings *et al.* 1991; Stallings, 1994). The microsatellite repeat investigated here, is conserved in a number of closely, and more distantly, related species of *Mus*. However, it amplifies only weakly from *Mus spretus*, and not at all from the rat, hamster, or human, even though humans have been demonstrated to have a homologue of the *Tcp10* gene family (Islam *et al.* 1993). This non-amplification could result either from the absence of the microsatellite in these species, or from DNA sequence divergence in the region of primer annealing, which can produce null alleles (Callen *et al.* 1993; Pemberton *et al.* 1995).

Despite the relatively high mutation rate associated with microsatellite loci over evolutionary time scales, our results show that microsatellite loci can be sufficiently stable over shorter time periods to be used as markers to follow individual *t* chromosomes through multiple generations. The Tcp10ms patterns observed for both the t^0 haplotype and the inbred 129 wild-type chromosome were found to have remained unchanged over, at least, the 12 years that these chromosomes have been bred in this laboratory (a number of these are shown in Fig. 2). Similarly, the identical patterns of t^{12} and t^{w32} supports previous suggestions that these two chromosomes are descendants from the same chromosome. These two chromosomes have been maintained in separate stocks since 1957 (Silver, 1983) during which time their Tcp10ms patterns have not diverged (Fig. 2).

In contrast, while newly identified wild-derived chromosomes from the t^{w5} complementation group (CF109, T57, T59 and T67) share similar TCP10ms patterns relative to one another and to the laboratory defined t^{w5} haplotype, they are all distinguishable on the basis of small allele size differences of 2–6 bp at the *Tcp10a* and *Tcp10b* loci. Both empirical and theoretical research on the mutational generation of new microsatellite alleles (Tautz & Renz, 1984; Schlötterer & Tautz, 1992; Valdes, Slatkin & Freimer, 1993; Stephan & Cho, 1994), suggest that new alleles form predominantly through processes such as slipped strand mispairing, so that the size of a new allele depends on the size of the allele that mutated. Thus the small allele size differences seen here among alleles of the t^{w5} complementation group may well reflect a recent common ancestry of these alleles.

It seems likely that house mice probably invaded North America for the first time with their human counterparts from Western Europe sometime during the early colonial period ~ 1620–1650. While a large number of different *t* haplotype complementation groups are found in Europe (Klein, Sipos & Figueroa, 1984) only a relatively small number of *t* complementation groups are found in North America, and a single lethal, t^{w5}, predominates (Bennett, 1975). Thus all current North American *t* haplotypes are probably descendants from a few original founders that existed no more than 375 years ago. Among the independent members of the t^{w5} group investigated here, all of the observed changes in Tcp10ms patterns are small, and are therefore likely to have occurred during this time period.

In summary, the Tcp10ms pattern of PCR products provides a powerful tool for the analysis of *t* haplotypes in both laboratory and field studies. In the laboratory, the Tcp10ms pattern can be used as a means to test and control for the purity of different *t* bearing stocks that are not easily distinguished except through complementation testing. In the field, the Tcp10ms pattern can be used for population genetic studies where identifying particular alleles is not of concern (Bruford & Wayne, 1993; Slatkin, 1995). Moreover, the pattern can be used to demonstrate the presence of a single, as opposed to multiple, *t* haplotypes founders within a given population.

K.A. wishes to thank D. Stern and H. Himmelbauer for help at the outset of this project, and S. Pilder for comments on the manuscript. This research was funded by a grant from the NIH (HD20275).

References

Artz, K., Shin, H. S., Bennett, D. & Dimeo-Talento, A. (1985). Analysis of major histocompatability complex haplotypes of *t*-chromosomes reveals that the majority of diversity is generated by recombination. *Journal of Experimental Medicine* **162**, 93–104.

Beckman, J. S. & Weber, J. L. (1992). Survey of human and rat microsatellites. *Genomics* **12**, 627–631.

Bennett, D. (1975). The T-locus of the mouse. *Cell* **6**, 441–454.

Bruford, M. W. & Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. *Current Opinions in Genetics and Development* **3**, 939–943.

Bullard, D. C. & Schimenti, J. C. (1990). Molecular cloning and genetic mapping of the *t* complex responder candidate gene family. *Genetics* **124**, 957–966.

Bullard, D. C. & Schimenti, J. C. (1991). Molecular structure of Tcp-10 genes from the *t* complex responder locus. *Mammalian Genome* **1**, 228–234.

Bullard, D. C., Ticknor, C. & Schimenti, J. C. (1992). Functional analysis of a *t* complex responder locus transgene. *Mammalian Genome* **3**, 588–596.

Callen, D. F., Thompson, A. D., Shen, Y., Phillips, H. A., Richards, R. I., Mulley, J. C. & Sutherland, G. R. (1993). Incidence and origin of 'null' alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics* **52**, 922–927.

Cebra-Thomas, J. A., Decker, C., Snyder, L. C., Pilder, S. H. & Silver, L. M. (1991). Allele- and haploid-specific product generated by alternative splicing from a mouse *t* complex responder locus candidate. *Nature* **349**, 239–241.

Committee for Mouse Chromosome 17 (1991). Maps of mouse chromosome 17: First report. *Mammalian Genome* **1**, 5–29.

Davies, P. O. & Willison, K. R. (1991). Sequence of the *t*-complex *Tcp-10a*^{t} gene and examination of the *Tcp-10*^{t} gene family. *Mammalian Genome* **1**, 235–241.

Dietrich, W., Katz, H., Lincoln, S., Shin, H.-S., Friedman, J., Dracopoli, N. C. & Lander, E. (1992). A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**, 423–447.

Dunn, L. C. (1957). Evidence of evolutionary forces leading to the spread of lethal genes in wild populations of house mice. *Proceedings of the National Academy of Sciences, USA* **43**, 158–163.

Ebersole, T., Lai, F. & Artzt, K. (1992). New molecular markers for the distal end of the *t*-complex and their relationships to mutations affecting mouse development. *Genetics* **131**, 175–182.

Ellegren, H. (1995). Mutation rates at porcine microsatellite loci. *Mammalian Genome* **6**, 376–377.

Figueroa, F., Neufeld, E., Ritte, U. & Klein, J. (1988). *t*-specific DNA polymorphisms among wild mice from Israel and Spain. *Genetics* **119**, 157–60.

Fox, H. S., Martin, G. R., Lyon, M. F., Herrmann, B., Frischauf, A.-M., Lehrach, H. & Silver, L. M. (1985). Molecular probes define different regions of the mouse *t* complex. *Cell* **40**, 63–69.

Hamada, H., Petrino, M. G. & Kakunaga, T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proceedings of the National Academy of Sciences, USA* **79**, 6465–6469.

Hammer, M. F., Schimenti, J. & Silver, L. M. (1989). Evolution of mouse chromosome 17 and the origin of inversions association with *t* haplotypes. *Proceedings of the National Academy of Sciences, USA* **86**, 3261–3265.

Hammer, M. F. & Silver, L. M. (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. *Molecular Biology and Evolution* **10**, 971–1001.

Herrmann, B., Búcan, M., Mains, P. E., Frischauf, A.-M., Silver, L. M. & Lehrach, H. (1986). Genetic analysis of the proximal portion of the mouse *t* complex: evidence for a second inversion within *t* haplotypes. *Cell* **44**, 469–476.

Herrmann, B. G., Barlow, D. P. & Lehrach, H. (1987). A large inverted duplication allows homologous recombination between chromosomes heterozygous for the proximal *t* complex inversion. *Cell* **48**, 813–825.

Higuchi, R. G. & Ochman, H. (1989). Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. *Nucleic Acids Research* **17**, 5865.

Hino, O., Testa, J. R., Buetow, K. H., Taguchi, T., Zhou, J.-Y., Bremer, M., Bruzel, A., Yeung, R., Levan, G., Levan, K. K., Knudson, A. G. & Tartof, K. D. (1993). Universal mapping probes and the origin of human chromosome 3. *Proceedings of the National Academy of Sciences, USA* **90**, 730–734.

Hogan, B., Beddington, R., Costantini, F. & Lacy, E. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual, Second Edition.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Horiuchi, Y., Agulnik, A., Figueroa, F., Tichy, H. & Klein, J. (1992). Polymorphisms distinguishing different mouse species and *t* haplotypes. *Genetical Research* **60**, 43–52.

Howard, C. A., Gummere, G. R., Lyon, M. F., Bennett, D. & Artzt, K. (1990). Genetic Molecular analysis of the proximal region of the mouse *t*-complex using new molecular probes and partial *t*-haplotypes. *Genetics* **126**, 1103–1114.

Hughes, C. R & Queller, D. C. (1993). Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology* **2**, 131–137.

Islam, S. D., Pilder, S. H., Decker, C. L., Cebra-Thomas, J. A. & Silver, L. M. (1993). The human homolog of a candidate mouse *t* complex responder gene: conserved motifs and evolution with punctuated equilibria. *Human Molecular Genetics* **2**, 2075–2079.

Klein, J., Sipos, P. & Figueroa, F. (1984). Polymorphism of *t*-complex genes in European wild mice. *Genetical Research* **44**, 39–46.

Kreitman, M. & Landweber, L. F. (1989). A strategy for producing single stranded DNA in the polymerase chain reaction: a direct method for genomic sequencing. *Gene Analysis Techniques* **6**, 84–88.

Lai, F. & Artzt, K. (1992). Map positions of four dinucleotide repeats in the mouse *t* complex. *Mammalian Genome* **3**, 476–477.

Lenington, S., Franks, P. & Williams, J. (1988). Distribution of *t* haplotypes in natural populations of wild house mice. *Journal of Mammalogy* **69**, 489–499.

Litt, M. & Lucy, J. A. (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**, 397–401.

Love, J. M., Knight, A. M., McAleer, M. A. & Todd, J. A. (1990). Towards construction of a high resolution map of the mouse genome using PCR-analyzed microsatellites. *Nucleic Acids Research* **18**, 4123–4130.

Lundrigan, B. L. & Tucker, P. K. (1994). Tracing paternal

ancestry in mice, using the Y-linked, sex-determining locus. *Sry*. *Molecular Biology and Evolution* **11**, 483–492.
Lyon, M. F. (1984). Transmission ratio distortion in mouse *t*-haplotypes is due to multiple distorter genes acting on a responder locus. *Cell* **37**, 621–628.
Lyon, M. F. (1986). Male sterility of the mouse *t*-complex is due to homozygosity of the distorter genes. *Cell* **44**, 357–363.
Lyon, M. F. & Zenthon, J. (1987). Differences in or near the responder region of complete and partial mouse *t*-haplotypes. *Genetical Research* **50**, 29–34.
Moore, S. S., Sargeant, L. L., King, T. J., Mattick, J. S., Georges, M. & Hetzel, D. J. S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics* **10**, 654–660.
Morita, T., Kubota, H., Murata, K., Nozaki, M., Delarbre, C., Willison, K., Satta, Y., Sakaizumi, M., Takahata, N., Gachelin, G. & Matsushiro, A. (1992). Evolution of the mouse *t* haplotype: Recent and worldwide introgression to *Mus musculus*. *Proceedings of the National Academy of Sciences, USA* **89**, 6851–6855.
Pemberton, J. M., Slate, J., Bancroft, D. R. & Barrett, J. A. (1995). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**, 249–252.
Petras, M. L. (1967). Studies of natural populations of *Mus*. II. Polymorphism at the *t* locus. *Evolution* **21**, 466–478.
Pilder, S. H., Decker, C. L., Islam, S., Buck, C., Cebra-Thomas, J. A. & Silver, L. M. (1992). Concerted evolution of the mouse Tcp-10 gene family: implications for the functional basis of *t* haplotype transmission ratio distortion. *Genomics* **12**, 35–41.
Rosen, L. L., Bullard, D. C., Silver, L. M. & Schimenti, J. C. (1990). Molecular cloning of the *t* complex responder genetic locus. *Genomics* **8**, 134–140.
Schimenti, J., Cebra-Thomas, J. A., Decker, C., Islam, S., Pilder, S. H. & Silver, L. M. (1988). A candidate gene family for the mouse *t complex responder* (*Tcr*) locus responsible for haploid effects on sperm function. *Cell* **55**, 71–78.
Schimenti, J., Vold, L., Socolow, D. & Silver, L. M. (1987). An unstable family of large DNA elements in the center of the mouse *t* complex. *Journal of Molecular Biology* **194**, 583–594.
Schlötterer, C. & Tautz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* **20**, 211–215.
Silver, L. M. (1983). Reevaluation of the evidence for the generation of new lethal *t* haplotypes by mutation. *Immunogenetics* **18**, 91–96.
Silver, L. M. (1985). Mouse *t* haplotypes. *Annual Reviews of Genetics* **19**, 179–208.
Silver, L. M. (1993). The peculiar journey of a selfish chromosome: mouse *t* haplotypes and meiotic drive. *Trends in Genetics* **9**, 250–254.
Silver, L. M. (1995). *Mouse Genetics: Concepts and Applications*, Oxford University Press, New York.
Silver, L. M., Hammer, M., Fox, H., Garrels, J., Búcan, M., Herrmann, B., Frischauf, A. M., Lehrach, H., Winking, H., Figueroa, F. & Klein, J. (1987). Molecular evidence for the rapid propagation of mouse *t* haplotypes from a single, recent, ancestral chromosome. *Molecular Biology and Evolution* **4**, 473–482.
Silver, L. M. & Remis, D. (1987). Five of the nine genetically defined regions of mouse *t* haplotypes are involved in transmission ratio distortion. *Genetical Research* **49**, 51–56.
Silver, L. M., Uman, J., Danska, J. & Garrels, J. I. (1983). A diversified set of testicular cell proteins specified by genes within the mouse *t* complex. *Cell* **35**, 35–45.
Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457–462.
Stallings, R. L. (1994). Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequences: Implications for human genetic diseases. *Genomics* **21**, 116–121.
Stallings, R. L. (1995). Conservation and evolution of (CT)n/(GA)n microsatellite sequences at orthologous positions in diverse mammalian genomes. *Genomics* **25**, 107–113.
Stallings, R. L., Ford, A. F., Nelson, D., Torney, D. C., Hildebrand, C. E. & Moyzis, R. K. (1991). Evolution and distribution of (GT)n repetitive sequences in mammalian genomes. *Genomics* **10**, 807–815.
Stephan, W. & Cho, S. (1994). Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. *Genetics* **136**, 333–341.
Tautz, D. (1993). Notes on the defunction and nomenclature of tandemly repetitive DNA sequences. In *DNA Fingerprinting: State of the science*, (ed. S. D. J. Pena, R. Chakraborty, J. T. Epplen and A. J. Jeffreys; pp. 21–28, Basel: Birkhauser Verlag.
Tautz, D. & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* **12**, 4127–4138.
Uehara, H., Ebersole, T., Bennett, D. & Artzt, K. (1990). Submegabase clusters of unstable tandem repeats unique to the *Tla* region of mouse *t* haplotypes. *Genetics* **126**, 1093–1102.
Valdes, A. M., Slatkin, M. & Freimer, N. B. (1993). Allele frequencies at microsatellite loci: The stepwise mutation model revisited. *Genetics* **133**, 737–749.
Weber, J. L. & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388–396.
Weber, J. L. & Wong, C. (1993). Mutation of human short tandem repeats. *Human Molecular Genetics* **2**, 1123–1128.
Willison, K. R., Dudley, K. & Potter, J. (1986). Molecular cloning and sequence analysis of a haploid expressed gene encoding *t* complex polypeptide-1. *Cell* **44**, 727–738.