*Genetica* **119:** 201–218, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*


# Genetic variation of microsatellite loci in the major histocompatibility complex (MHC) region in the southeast Asian house mouse (*Mus musculus castaneus*)

Shiao-Wei Huang & Hon-Tsen Yu*
*Department of Zoology, National Taiwan University, Taipei 106, Taiwan, ROC; *Author for correspondence (Phone: +886-2-23630231, ext. 2127; Fax: +886-2-23636837; E-mail: ayu@ccms.ntu.edu.tw)*

Received 27 February 2001 Accepted 15 April 2003

*Key words:* balancing selection, major histocompatibility complex, microsatellites, *M. m. castaneus*, *Mus musculus*, polymorphism, wild mouse

## Abstract

Major histocompatibility complex (MHC) genes are the most polymorphic loci known for vertebrates. Here we employed five microsatellite loci closely linked to the MHC region in an attempt to study the amount of genetic variation in 19 populations of the southeast Asian house mouse (*Mus musculus castaneus*) in Taiwan. The overall polymorphism at the five loci was high ($\bar{H}_e = 0.713$), and the level of polymorphism varied from locus to locus. Furthermore, in order to investigate if selection is operating on MHC genes in natural mouse populations, we compared the extent and pattern of genetic variation for the MHC-linked microsatellite loci (the MHC loci) with those for the microsatellite loci located outside the MHC region (the non-MHC loci). The number of alleles and the logarithm of variance in repeat number were significantly higher for the MHC loci than for the non-MHC loci, presumably reflecting linkage to a locus under balancing selection. Although three statistical tests used do not provide support for selection, their lack of support may be due to low statistical power of the tests, to weakness of selection, or to a profound effect of genetic drift reducing the signature of balancing selection. Our results also suggested that the populations in the central and the southwestern regions of Taiwan might be one part of a metapopulation structure.

## Introduction

The major histocompatibility complex (MHC) is a multigene family encoding cell-surface glycoproteins that mediate both humoral and cell-mediated immune responses. The MHC region of the mouse spans a region containing 2000–4000 kb of DNA on chromosome 17 and comprises genes grouped as class I, class II, and class III (Hood, Steinmetz & Malissen, 1983). Some of the class II as well as the class I genes are the most polymorphic genes known to date. The characteristics of some of these highly polymorphic MHC genes – long persistence of allelic lineages, prevalence of non-synonymous over synonymous substitutions in the peptide-binding region (PBR), and departures from the level of homozygosity expected under neutrality – indicate that some sort of selection is acting to maintain their extreme diversity (Nei & Hughes, 1991; Satta et al., 1994; Paterson, 1998).

The mechanism maintaining polymorphism at the MHC loci has been widely debated. Due to their critical role in immune response, it is generally assumed that the selective pressures affecting MHC diversity arise from infectious disease, but whether that selection takes the form of heterozygote advantage (over-dominance selection) or negative frequency-dependent selection, remains controversial (Takahata & Nei, 1990; Takahata, Satta & Klein, 1992; Potts & Wakeland, 1993). On the other hand, in addition to parasite-mediated selection, reproductive selection through MHC-based mating preferences and/or selective abortion, has also been contributed to the evolution of MHC diversity in mice (Potts & Wakeland, 1993).

In order to understand the mechanism for maintaining genetic variation, it is important to study the amount of genetic variation in natural populations. However, most surveys of natural population polymorphism in the extensive MHC region so far have been conducted on humans (Klitz, Thomson & Baur, 1986), with a few limited studies on other mammals (e.g., bighorn sheep, Boyce et al., 1997; domestic sheep, Paterson, 1998), due to lack of a simple and inexpensive MHC genotyping system for any animal species. The recent description of over 7000 microsatellites densely distributed across the house mouse genome (Dietrich et al., 1996) now makes it possible to perform large-scale microsatellite-based MHC typing for this species. Although microsatellite loci are generally thought to be selectively neutral, a few theoretical studies have already shown that the statistics of microsatellite allele distribution would be influenced by an adjacent locus under selection, suggesting that neutral microsatellite loci can be indicators of selective processes at closely linked loci (Slatkin, 1995; Schlötterer & Wiehe, 1999). Several studies have used microsatellites to infer selection indirectly, for example, Schlötterer, Vogl and Tautz (1997), Paterson (1998), Huttley et al. (1999) and Kohn, Pelz and Wayne (2000).

Furthermore, comparing patterns of genetic polymorphism among populations at different types of loci is an increasingly popular approach to assess the role of selection in determining allelic variation (Lewontin & Krakauer, 1973; Karl & Avise, 1992; Spitze, 1993; Pogson et al., 1995; Lynch et al., 1999). As gene flow and genetic drift affect all loci equally whereas selection is more likely to be locus-specific (Lewontin & Krakauer, 1973), discordance between potentially selected and neutral loci may be taken as evidence of selection (Spitze, 1993; Lynch et al., 1999). This can be achieved by comparing the patterns of genetic differentiation at the microsatellite loci within the MHC (MHC-microsatellites) with those outside the MHC region (non-MHC microsatellites) among populations. If the selective factors acting on MHC have been of significantly greater magnitude than non-selective factors, the patterns of variation for MHC and non-MHC microsatellite loci may differ greatly. On the contrary, the hypothesis of neutrality (no selection or weak selection) shaping MHC genetic structure would in turn be supported by the absence of statistical differences between the patterns of population structure at MHC- and at non-MHC microsatellites.

In the present study, we use five microsatellite loci to investigate MHC polymorphism in 19 populations of the southeast Asian house mouse (*Mus musculus castaneus*) in Taiwan. We describe the population genetics of these loci by using standard genetic parameters and estimate the relationship among the populations. Further, as the pattern and level of polymorphism across the genome are useful for identifying genes or genomic regions subject to natural selection (Satta, Li & Takahata, 1998), we compare the extent and pattern of genetic variation for MHC- and non-MHC microsatellite loci, in order to discover if selection is, or has been, acting on the MHC region.

## Materials and methods

### *Mouse samples*

Since June 1995, we have collected a large number of house mice from 19 populations inhabiting centralized rice granaries in townships in lowland Taiwan (Chou et al., 1998). Mice trapped from the same township were considered as one population. We grouped all populations into five geographical regions on the island of Taiwan (Figure 1, Table 2 and Appendix A). In addition, mice from Jinmen were treated as a separate population on an offshore isle.

### *Microsatellite loci*

Five microsatellite loci (*D17Mit28*, *D17Mit21*, *D17Mit33*, *D17Mit233*, and *D17Mit124*) closely linked to the MHC genes were employed in this study. These microsatellites lie between 18.00 and 22.50 cM from the centromere of Chromosome 17 (Figure 2). All are simple $(CA)_n$ repeat loci and are described in detail on the website of the Whitehead Institute for Biomedical Research at http://www.genome.wi.mit.edu. Of the five loci, three (*D17Mit28*, *D17Mit21*, and *D17Mit33*) are located within the conventional *H-2* region (bounded by the *K* and *L* genes) where most of the highly polymorphic antigen-presenting MHC loci are found, whereas the other two (*D17Mit233* and *D17Mit124*) are located close to the boundary of the MHC region and near the *M* region. General information about these MHC microsatellite loci is summarized in Table 1.

In an earlier attempt to investigate population differentiation in these mice, we employed six unlinked autosomal microsatellite loci (*D6Mit138*, *D10Mit20*, *D15Mit16*, *34*, *105*, and *150*) (Yu & Peng, 2002).

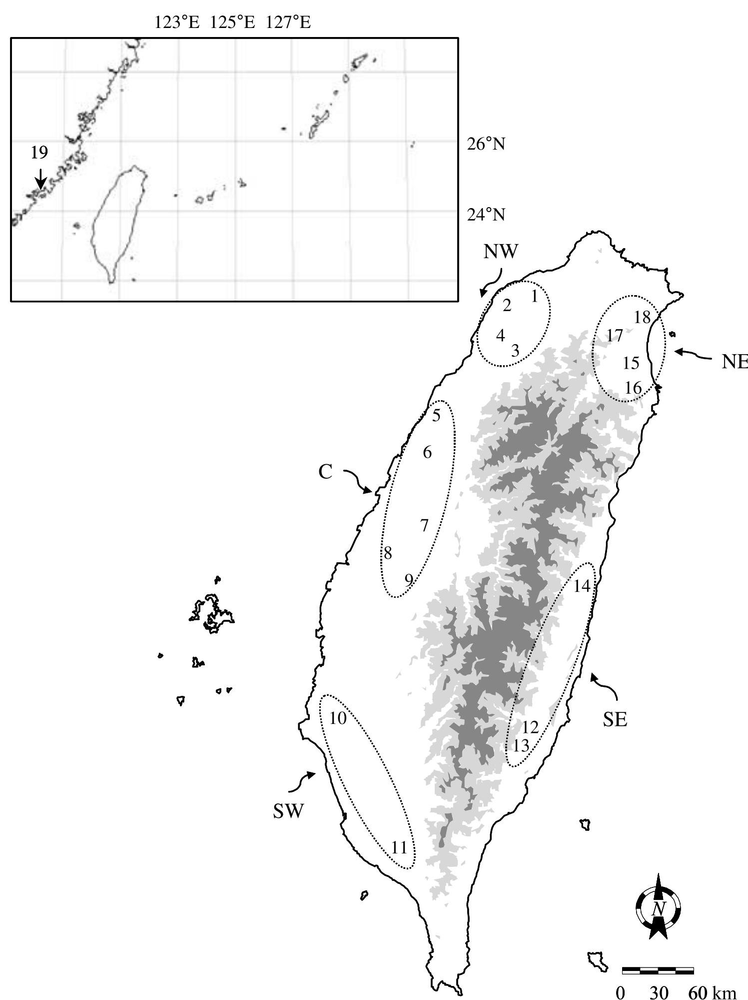


*Figure 1.* Map indicating the location of the 19 mouse populations grouped into six geographic regions in our survey. Northwestern region (NW) includes: 1, Tayuan; 2, Hsinwu; 3, Guanshi; 4, Hsinpu. Central region (C) includes: 5, Miaoli; 6, Taichung; 7, Tsaotun; 8, Shijou; 9, Linnei. Southwestern region (SW) includes: 10, Tainan; 11, Kaoping. Southeastern region (SE) includes: 12, Chrshang; 13, Guanshan; 14, Huanlien. Northeastern region (NE) includes: 15, Wuchiai; 16, Tungshan; 17, Ilan City; 18, Toucheng. The relative position of an offshore island, Jinmen (19), is shown in the inset panel. Light gray areas, 1000–2000 m; dark gray area, above 2000 m in elevation.

The data from that study are compared with those reported here to reveal any difference in the levels of polymorphism between loci located within and outside the MHC region. These six microsatellite loci are assumed to be neutral. The repeat motifs of the six microsatellite loci are as follows (Yu & Peng, 2002): *150* [$(ATT)_n$], *34* [$(GAAG)_n$], *D10Mit20* [$(TAGA)_n$], *D15Mit16* [$(TAGA)_n$], *105* [$(ATTTT)_n$], and *D6Mit138* [$(GA)_n$ $(GAAA)_m$].

*Genotyping*

Genomic DNA extraction from mice (from liver, spleen, etc.) followed the standard phenol–chloroform procedures (Ausubel et al., 1995). PCR was carried out in a thermal cycler (AG9600) in 10 μl reactions overlaid with mineral oil containing: 200 ng of genomic DNA; 1× PCR buffer; 0.5 mM $MgCl_2$; 0.25 mM dNTP; 0.05 U *Taq* polymerase (Promega); 0.3 μM of each primer, one of which was end-labeled with [$\gamma$-$P^{33}$]-ATP (2000 Ci/mmol). Standard cycling conditions were 94°C for 3 min followed by 21–25 cycles of 94°C for 1 min, 59°C (for *D17Mit33*, *D17Mit21* and *D17Mit28*) or 61°C (for *D17Mit233* and *D17Mit124*) for 1 min, and 72°C for 1 min. A final 10 min extension step was added to complete the thermal profile. PCR products were then separated on 8 or 10% denaturing sequencing gels and visualized by exposure

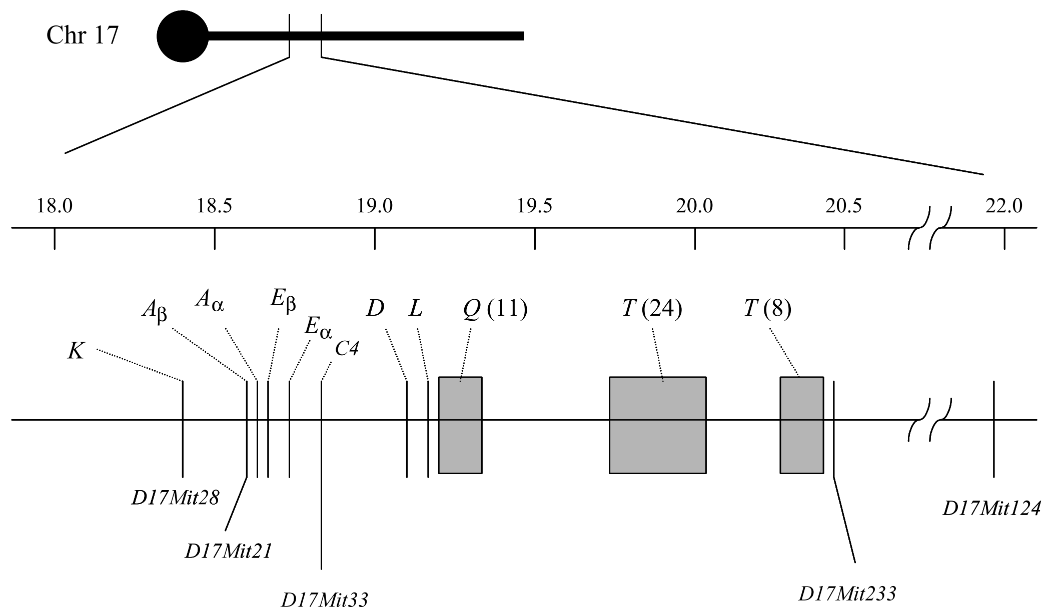


*Figure 2.* Linkage map of the five microsatellite loci and selected MHC genes on the 17th chromosome of the house mouse. Traditional *H-2* regions include *K*, $A_\beta$, $A_\alpha$, $E_\beta$, $E_\alpha$, *D*, and *L* genes. Scales are linkage distances (cM) from centromere. The MHC loci are based on a linkage map in Meagher and Potts (1997).

*Table 1.* Number of alleles, variance in repeat number corrected for length difference ($V_{\text{length-corrected}}$), and other information for the five microsatellite loci in 19 mouse populations in Taiwan

| Locus | Size range (bp) | No. of alleles | $V_{\text{length-corrected}}$ | Locus location[a] | Classification |
|---|---|---|---|---|---|
| *D17Mit33* | 158–212 | 15 | 1.658 | *C4/Slp hybrid 3* gene, non-transcriptional region | Class III |
| *D17Mit21* | 106–204 | 27 | 0.855 | *H2-A-beta-2* segment, intron 3 | Class II |
| *D17Mit28* | 86–130 | 19 | 1.060 | *H2-K* (*bm1*) gene, promoter region | Class I |
| *D17Mit124* | 147–169 | 18 | 0.114 | *H2-M* region | Class I |
| *D17Mit233* | 102–136 | 17 | 0.557 | *H2-M* region | Class I |

[a] From http: www.ncbi.nlm.nih.gov.

to Kodak X-OMAT Blue film for 1–2 days. DNA sequences of plasmids (pUC 18 or pUC 19) were run along with the samples as markers to determine the allele sizes. Mice of two inbred strains (B6 and CBA/J) and their $F_1$ hybrid (purchased from the Lab Animal Center, National Taiwan University) were used as controls to recognize the bands (each representing an allele) for homozygotes (inbred strains) and for heterozygotes ($F_1$ hybrids). This procedure helps remove ambiguity in scoring alleles.

### *Data analysis*

Both observed heterozygosity ($H_o$) and unbiased expected heterozygosity ($H_e$; Nei, 1978) were calculated to estimate the genetic variability in the 19 mouse populations. Hardy–Weinberg expectation for each locus was tested by the Markov-chain method in GENEPOP 3.1 (Raymond & Rousset, 1995; http://wbiomed.curtin.edu.au/genepop), which implements Fisher's exact tests for multiple alleles (Guo & Thompson, 1992).

Variance in the number of repeats was calculated as another measure of variability. To account for length dependence of microsatellite mutation rates at each locus, variance in repeat number ($V_i$) was divided by the maximal number of repeat ($X_i$) as follows:

$$V_{\text{length-corrected}} = \frac{1}{k}\sum_{i=1}^{k}\frac{V_i n_i}{X_i(n_i - 1)}$$

where $k$ is the number of population, and $n_i$ is the number of chromosomes analyzed in the $i$th population.

Two genetic distance measures were calculated: Nei's (1972) standard distance, $D_s$, assuming an infinite allele model (IAM) (Nei, 1972), and Goldstein et al.'s (1995) distance $(\delta\mu)^2$, assuming the stepwise mutation model (SMM). Calculations were implemented in MICROSAT 1.4 (Minch et al., 1995).

Population differentiation at different geographic scales was examined by determining $F_{ST}$ values (Weir & Cockerham, 1984) between all possible pairs of 18 populations, or among populations within geographic regions (northwestern, central, southwestern, southeastern, and northeastern) in Taiwan. The calculations were implemented in GENEPOP 3.1 (Raymond & Rousset, 1995). Significant departures from zero of the $F_{ST}$ values were tested using permutations (see Dallas et al., 1995) and were implemented in FSTAT (Goudet, 1995).

Isolation by distance was tested by the Mantel's test (Mantel, 1967), which examines the overall relationship between population differentiation ($F_{ST}/(1 - F_{ST})$) and the logarithm of the geographic distance separating populations (Rousset, 1997). The Mantel's test was implemented in GENEPOP 3.1 (Raymond & Rousset, 1995).

Balancing selection was tested using Watterson's (1978) homozygosity test. The expected proportion of homozygotes under Hardy–Weinberg equilibrium, $F_{obs} = (\sum p_i^2)$, in a sample of size $2n$ genes containing $k$ alleles, is used as a measure of the allele frequency distribution and is compared to the homozygosity ($F_{exp}$) expected in a sample drawn from a population in mutation-drift equilibrium under neutrality (Ewens, 1972). The distribution of the homozygosity statistic under the finite-alleles neutral model, for combination of $2n$ and $k$ up to 500 and 40, has been obtained from computer simulations and tabulated extensively at http://allele5.biol.berkeley.edu/homozygosity/homozygosity.html. Significantly low $p$-values reject the null hypothesis of neutrality and suggest the presence of selection.

## Results

### *Genetic variability*

Genetic variability for the five MHC microsatellite loci in 19 mouse populations was generally high. For all populations combined, the numbers of alleles for each locus ranged from 15 to 27, and the variance in the repeat number corrected for length differences at each locus ($V_{\text{length-corrected}}$) ranged from 0.855 to 1.658 (Table 1).

For individual populations, mean allele numbers per locus ranged from 3.6 to 9.8 (Appendix A). The observed ($H_o$) and the expected heterozygosity ($H_e$) for individual populations ranged from 0.219 to 0.768 ($\bar{H}_o = 0.601$) and from 0.330 to 0.835 ($\bar{H}_e = 0.713$), respectively (Table 2).

The level of polymorphism at each locus seems to be related to its position (Tables 1 and 2). The highest level of polymorphism was observed for *D17Mit21* located in the class II region, with 27 alleles and an expected heterozygosity ($H_e$) equal to 0.801. The lowest level of polymorphism was observed for *D17Mit33* located in the class III region, with 15 alleles and an expected heterozygosity ($H_e$) equal to 0.659. The three loci (*D17Mit28*, *D17Mit124* and *D17Mit233*) in the class I region showed an intermediate level of polymorphism, with 17–19 alleles and averaged expected heterozygosities ($H_e$) 0.684–0.717.

The levels of polymorphism varied among individual populations as measured by the expected heterozygosity (Table 2). Guanshi was clearly the least polymorphic, with $H_e = 0.330$. On the other hand, four populations (Jinmen, Tainan, Taichung, and Kaoping) showed high expected heterozygosity, ranging from 0.800 to 0.835.

Private alleles were found in over half of the populations (10/19) for the five loci. The number of private alleles was highest in Jinmen ($n = 5$) (Appendix A). While most of the private alleles had frequencies less than 0.1, the *D17Mit233* allele*162 in Miaoli had a frequency as high as 0.3 (Appendix A).

For the 19 populations at these five loci, 28 out of the 95 cases involved a single most common allele with frequency exceeding 0.5 (= 'predominant' allele) (Appendix A). For individual populations, Guanshi had the largest number of loci with a 'predominant' allele, that is, each of the five loci had its own predominant allele. For each individual locus, while there were 8, 6, 7, and 6 populations containing a 'predominant' allele at *D17Mit33*, *D17Mit28*, *D17Mit124*, and *D17Mit233*, respectively, there was only one (Guanshi) at *D17Mit21*. Otherwise, most populations contained more than four alleles, with roughly equal frequencies, at each locus.

The proportions of microsatellite genotypes observed in each population sample were compared with Hardy–Weinberg expectations (HWE) using Fisher's exact test (Guo & Thompson, 1992). Significant departures ($P < 0.05$) were found in 28 out of the 95

*Table 2.* Observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), and sample size for the five MHC microsatellite loci in 19 Taiwanese house mouse populations

| Region | Population | Observed heterozygosity ($H_o$) | | | | | | Expected heterozygosity ($H_e$) | | | | | | Sample size | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *D17Mit 33* | *D17Mit 21* | *D17Mit 28* | *D17Mit 124* | *D17Mit 233* | Mean | *D17Mit 33* | *D17Mit 21* | *D17Mit 28* | *D17Mit 124* | *D17Mit 233* | Mean | *D17Mit 33* | *D17Mit 21* | *D17Mit 28* | *D17Mit 124* | *D17Mit 233* |
| Northwestern | Tayuan | 0.389 | 0.579 | 0.684 | 0.421 | 0.684 | 0.551 | 0.492 | 0.824 | 0.670 | 0.551 | 0.799 | 0.667 | 18 | 19 | 19 | 19 | 19 |
| | Hsinwu | 0.533 | 0.667 | 0.400 | 0.933 | 0.933 | 0.693 | 0.632 | 0.798 | 0.492 | 0.772 | 0.825 | 0.704 | 15 | 15 | 15 | 15 | 15 |
| | Guanshi | 0.143 | 0.143 | 0.095 | 0.571 | 0.143 | 0.219 | 0.138 | 0.307 | 0.304 | 0.558 | 0.343 | 0.330 | 21 | 21 | 21 | 21 | 21 |
| | Hsinpu | 0.600 | 0.800 | 0.200 | 0.300 | 0.800 | 0.520 | 0.668 | 0.790 | 0.442 | 0.337 | 0.737 | 0.595 | 10 | 10 | 10 | 10 | 10 |
| Central | Miaoli | 0.200 | 0.400 | 0.600 | 0.400 | 0.800 | 0.440 | 0.200 | 0.844 | 0.711 | 0.600 | 0.867 | 0.644 | 5 | 5 | 5 | 5 | 5 |
| | Taichung | 0.875 | 0.500 | 0.875 | 0.625 | 0.750 | 0.725 | 0.817 | 0.892 | 0.850 | 0.750 | 0.792 | 0.820 | 8 | 8 | 8 | 8 | 8 |
| | Tsaotun | 0.556 | 0.778 | 0.778 | 0.667 | 0.500 | 0.656 | 0.532 | 0.730 | 0.787 | 0.687 | 0.522 | 0.652 | 18 | 18 | 18 | 18 | 18 |
| | Shijou | 0.569 | 0.778 | 0.585 | 0.769 | 0.569 | 0.654 | 0.689 | 0.878 | 0.814 | 0.828 | 0.660 | 0.774 | 65 | 63 | 65 | 65 | 65 |
| | Linnei | 0.556 | 0.657 | 0.580 | 0.720 | 0.590 | 0.620 | 0.676 | 0.878 | 0.687 | 0.792 | 0.607 | 0.728 | 99 | 99 | 100 | 100 | 100 |
| Southwestern | Tainan | 0.667 | 0.741 | 0.556 | 0.593 | 0.778 | 0.667 | 0.783 | 0.894 | 0.827 | 0.773 | 0.804 | 0.816 | 27 | 27 | 27 | 27 | 27 |
| | Kaoping | 1.000 | 0.778 | 0.667 | 0.333 | 0.889 | 0.733 | 0.808 | 0.895 | 0.752 | 0.686 | 0.856 | 0.800 | 8 | 9 | 9 | 9 | 9 |
| Southeastern | Chrshang | 0.611 | 0.333 | 0.444 | 0.556 | 0.722 | 0.533 | 0.798 | 0.767 | 0.687 | 0.703 | 0.757 | 0.743 | 18 | 18 | 18 | 18 | 18 |
| | Guanshan | 0.462 | 0.333 | 0.538 | 0.385 | 0.462 | 0.436 | 0.563 | 0.815 | 0.726 | 0.748 | 0.532 | 0.677 | 13 | 12 | 13 | 13 | 13 |
| | Hualien | 0.842 | 0.868 | 0.632 | 0.216 | 0.921 | 0.696 | 0.809 | 0.838 | 0.745 | 0.202 | 0.760 | 0.671 | 38 | 38 | 38 | 37 | 38 |
| Northeastern | Wuchiai | 0.500 | 0.545 | 0.417 | 0.250 | 0.700 | 0.482 | 0.529 | 0.887 | 0.833 | 0.377 | 0.760 | 0.677 | 12 | 11 | 12 | 12 | 10 |
| | Tungshan | 0.375 | 0.267 | 0.533 | 0.750 | 0.714 | 0.528 | 0.548 | 0.352 | 0.637 | 0.790 | 0.878 | 0.641 | 16 | 15 | 15 | 12 | 14 |
| | Ilan City | 0.595 | 0.515 | 0.343 | 0.700 | 0.719 | 0.580 | 0.720 | 0.803 | 0.781 | 0.806 | 0.835 | 0.789 | 37 | 33 | 35 | 30 | 32 |
| | Toucheng | 0.667 | 0.786 | 0.235 | 0.333 | 0.462 | 0.497 | 0.651 | 0.743 | 0.745 | 0.595 | 0.406 | 0.625 | 18 | 14 | 17 | 9 | 13 |
| Offshore | Jinmen | 0.767 | 0.895 | 0.737 | 0.579 | 0.895 | 0.768 | 0.838 | 0.809 | 0.878 | 0.727 | 0.925 | 0.835 | 19 | 19 | 19 | 19 | 19 |
| Mean (over population) | | 0.580 | 0.641 | 0.530 | 0.597 | 0.659 | 0.601 | 0.659 | 0.801 | 0.717 | 0.684 | 0.692 | 0.713 | – | – | – | – | – |

*Table 3.* Measures of genetic distance for all pairwise combinations of house mouse populations in Taiwan[a]

| Region | Population | Northwestern | | | | Central | | | | | Southwestern | | Southeastern | | | Northeastern | | | | Offshore, JM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | TY | SW | GS | SP | ML | TC | TT | SJ | LN | TN | KP | CS | GA | HL | WC | TS | IC | TE | |
| Northwestern | TY | | 0.440 | 0.871 | 0.320 | 0.541 | 0.420 | 0.456 | 0.483 | 0.461 | 0.361 | 0.379 | 0.721 | 0.396 | 0.341 | 0.231 | 0.601 | 0.335 | 0.806 | 0.421 |
| | SW | 12.11 | | 0.893 | 0.588 | 0.250 | 0.343 | 0.733 | 0.566 | 0.398 | 0.375 | 0.201 | 0.673 | 0.572 | 0.587 | 0.850 | **1.279** | 0.602 | 0.523 | 0.775 |
| | GS | **54.24** | 18.47 | | 0.772 | 0.720 | 0.878 | **1.321** | 0.989 | 0.905 | **1.348** | 0.635 | **1.166** | **1.849** | 0.827 | **1.480** | **1.454** | **1.452** | **1.028** | 0.910 |
| | SP | 3.52 | 13.02 | 39.95 | | 0.889 | 0.760 | 0.611 | 0.475 | 0.427 | 0.434 | 0.448 | 0.658 | 0.786 | 0.547 | 0.450 | 0.966 | 0.480 | 0.565 | 0.598 |
| Central | ML | **40.24** | 11.97 | 26.86 | **43.50** | | 0.458 | **1.123** | 0.836 | 0.597 | 0.620 | 0.447 | 0.743 | 0.894 | 0.463 | 0.852 | **1.042** | 0.916 | **1.160** | **1.043** |
| | TC | 4.82 | 4.90 | 31.37 | 2.50 | 25.04 | | 0.177 | 0.179 | 0.261 | 0.327 | 0.265 | 0.555 | 0.327 | 0.460 | 0.626 | 0.763 | 0.645 | 0.489 | 0.310 |
| | TT | 5.75 | 25.37 | **62.69** | 2.56 | **63.25** | 6.67 | | 0.181 | 0.333 | 0.680 | 0.535 | 0.838 | 0.278 | 0.570 | 0.638 | 0.895 | 0.558 | 0.516 | 0.435 |
| | SJ | 14.77 | 17.86 | **42.02** | 8.08 | 39.17 | 1.59 | 8.37 | | 0.149 | 0.519 | 0.496 | 0.559 | 0.271 | 0.616 | 0.664 | 0.655 | 0.542 | 0.512 | 0.456 |
| | LN | 11.89 | 5.32 | 28.44 | 9.37 | 22.74 | −0.31 | 14.35 | 4.04 | | 0.402 | 0.320 | 0.482 | 0.338 | 0.482 | 0.707 | 0.873 | 0.530 | 0.460 | 0.595 |
| Southwestern | TN | 6.19 | 9.69 | **44.67** | 7.26 | 28.13 | −0.57 | 8.71 | 3.17 | 1.97 | | 0.245 | 0.589 | 0.516 | 0.498 | 0.426 | 0.691 | 0.542 | 0.421 | 0.362 |
| | KP | 4.32 | 0.32 | 22.16 | 2.72 | 20.74 | 0.05 | 12.00 | 10.22 | 3.54 | 5.17 | | 0.627 | 0.646 | 0.515 | 0.615 | **1.203** | 0.592 | 0.557 | 0.512 |
| Southeastern | CS | 12.92 | 8.46 | 21.80 | 4.47 | 25.78 | 0.97 | 12.33 | 4.41 | 4.25 | 6.85 | 2.42 | | 0.384 | 0.656 | 0.656 | 0.929 | 0.592 | 0.596 | 0.867 |
| | GA | 11.50 | 18.61 | 35.19 | 1.33 | **54.97** | 6.93 | 5.12 | 10.48 | 13.38 | 14.60 | 6.79 | 5.40 | | 0.564 | 0.597 | 0.865 | 0.514 | 0.614 | 0.802 |
| | HL | 12.87 | 2.95 | 18.18 | 8.29 | 27.50 | 2.97 | 16.77 | 11.16 | 3.39 | 8.97 | 1.14 | 5.40 | 8.88 | | 0.446 | 0.926 | 0.677 | 0.578 | 0.484 |
| Northeastern | WC | 0.24 | 13.87 | **48.75** | 0.83 | **40.33** | 6.10 | 5.89 | 15.79 | 15.46 | 9.85 | 3.37 | 9.39 | 7.24 | 14.71 | | 0.332 | 0.486 | 0.763 | 0.338 |
| | TS | 13.26 | 25.98 | **79.23** | 21.12 | **41.29** | 10.39 | 17.02 | 12.71 | 13.69 | 4.64 | 21.14 | 23.87 | 34.38 | 28.13 | 20.64 | | 0.568 | **1.174** | 0.414 |
| | IC | 0.66 | 17.51 | **65.50** | 6.53 | **50.91** | 8.02 | 5.43 | 17.44 | 14.92 | 8.31 | 9.28 | 18.74 | 14.51 | 16.45 | 4.43 | 13.13 | | 0.587 | 0.613 |
| | TE | 5.02 | 6.89 | **41.28** | 6.61 | 31.38 | −0.36 | 8.15 | 5.56 | 1.31 | 0.32 | 3.50 | 8.22 | 12.39 | 4.73 | 10.59 | 8.37 | 5.65 | | 0.557 |
| Offshore | JM | 18.59 | 31.34 | **72.45** | 17.92 | **45.36** | 9.29 | 13.91 | 5.70 | 13.59 | 5.94 | 22.65 | 16.03 | 26.74 | 29.15 | 20.97 | 4.43 | 20.28 | 12.57 | |

[a] Both $D_s$ (above diagonal) and $(\delta\mu)^2$ (below diagonal) are given. $D_s$ and $(\delta\mu)^2$ values more than 1 and 40 are shown bold-faced. TY: Tayuan; SW: Hsinwu; GS: Guanshi; SP: Hsinpu; ML: Miaoli; TC: Taichung; TT: Tsaotun; SJ: Shijou; LN: Linnei; TN: Tainan; KP: Kaoping; CS: Chrshang; GA: Guanshan; HL: Haulien; WC: Wuchiai; TS: Tungshan; IC: Ilan City; TE: Toucheng; JM: Jinmen.

*Table 4.* $F_{ST}$ values for the five MHC microsatellite loci in 19 Taiwanese house mouse populations grouped into different regions (NW: northwestern; C: central; SW: southwestern; SE: southeastern; NE: northeastern)

| Locus | NW | C | SW | SE | NE | All[a] | All |
|---|---|---|---|---|---|---|---|
| *D17Mit33* | 0.4476*** | 0.1226*** | 0.0968** | 0.1005*** | 0.1823*** | 0.1767*** | 0.1728*** |
| *D17Mit21* | 0.2210*** | 0.0417*** | 0.0135* | 0.1249*** | 0.2390*** | 0.1412*** | 0.1403*** |
| *D17Mit28* | 0.3365*** | 0.0539*** | 0.0589*** | 0.2025*** | 0.0901*** | 0.1539*** | 0.1482*** |
| *D17Mit124* | 0.1570*** | 0.0779*** | 0.0394* | 0.2609*** | 0.1365*** | 0.1285*** | 0.1261*** |
| *D17Mit233* | 0.2718*** | 0.0514*** | 0.0064 | 0.1383*** | 0.1537*** | 0.1664*** | 0.1631*** |
| Overall | 0.3209*** | 0.0791*** | 0.0495 | 0.1773*** | 0.2029*** | 0.1764*** | 0.1722*** |

[a] All populations were combined except offshore island Jinmen.
* $P < 0.05$ from permutation tests in FSTAT program.
** $P < 0.01$ from permutation tests in FSTAT program.
*** $P < 0.001$ from permutation tests in FSTAT program.

cases, most of which showed heterozygote deficit ($P < 0.05$) except one (Hwalien at *D17Mit28*). The distribution of the 28 cases was somewhat clustered by three populations and two loci. For individual populations, Guanshi, Shijou, and Linnei had four loci that showed heterozygote deficiencies. For individual loci, *D17Mit21* and *D17Mit28* showed significant deviations from HWE in 9 and 8 out of the 19 populations, respectively. If the data from the five loci and the 19 populations were combined, the overall genotype frequencies deviated significantly from HWE ($P < 0.001$).

### *Genetic distances and differentiation among regions*

Two genetic distance measures, allele-frequency-based genetic distance, $D_s$, and repeat-size-based genetic distance $(\delta\mu)^2$ were estimated (Table 3). Two aspects are noteworthy. First, the high values of genetic distance for Guanshi and Miaoli from other populations indicate that they were quite different from the other mouse populations in Taiwan. The $(\delta\mu)^2$ values between Guanshi and other populations averaged more than 40 ($\bar{X} = 40.05$, $SD = 17.68$; $n = 17$), while the pairwise average between all 18 populations was 15.20 ($SD = 15.11$; $n = 153$). The averaged $(\delta\mu)^2$ between Miaoli and others was 34.93 ($SD = 13.43$; $n = 17$). The same pattern was observed for $D_s$. Second, the population of the offshore island, Jinmen, was particularly similar to those in central and southwestern Taiwan (Taichung, Tsaotun, Shijou, Linnei, Tainan, and Kaoping; Figure 1), except Miaoli. The mean values of $(\delta\mu)^2$ (11.85) and $D_s$ (0.445) between Jinmen and these six populations were significantly less than those between Jinmen and the remaining 12 Taiwanese mouse populations ($(\delta\mu)^2 = 26.32$ and $D_s = 0.652$) ($P = 0.024$ for $(\delta\mu)^2$, and $P = 0.016$ for $D_s$).

Wright's $F_{ST}$ (Weir & Cockerham, 1984) for pairs of populations was calculated as a measure of population differentiation. We found no significant correlation between population differentiation and geographic distance ($P = 0.365$), using Mantel's (Mantel, 1967).

Moreover, we examined inter-population genetic differentiation within different geographic regions (Table 4). For the five MHC loci, all $F_{ST}$ values for each region were highly significant ($P < 0.001$) except for southwestern populations, indicating that there was substantial isolation among populations in each region. The northwestern region had the most differentiated populations ($F_{ST} = 0.3209$), followed by

*Table 5.* Summarized result of Watterson's neutrality test applying to the five MHC microsatellite loci in 19 Taiwanese mouse populations

| Locus | Neutrality rejected (%)[a] | $F_{obs}^{b,c}$ | $F_{exp}^{c,d}$ | Ratio[c] |
|---|---|---|---|---|
| *D17Mit33* | 31.6 (6/19) | 0.3938 | 0.4410 | 0.8708 |
| *D17Mit21* | 31.6 (6/19) | 0.2509 | 0.2801 | 0.8675 |
| *D17Mit28* | 15.8 (3/19) | 0.3198 | 0.3604 | 0.8846 |
| *D17Mit124* | 31.6 (6/19) | 0.3770 | 0.4086 | 0.8799 |
| *D17Mit233* | 10.5 (2/19) | 0.3068 | 0.3443 | 0.8722 |

[a] Number of populations in which neutrality was rejected ($P < 0.1$).
[b] $F_{obs}$: sum of the squares of all allele frequencies.
[c] $F_{obs}$, $F_{exp}$, and ratio ($= F_{obs}/F_{exp}$) were weighted values averaged over all populations.
[d] $F_{exp}$: expected $F$-values under neutrality obtained by computer simulations.

*Table 6.* Comparisons of allele number, variance in repeat number, heterozygosity, $F_{ST}$ value, and proportion of populations in which neutrality was rejected between the MHC and the non-MHC microsatellite loci

| | MHC[a] | | | | | | Non-MHC[a] | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *D17Mit33* | *D17Mit21* | *D17Mit28* | *D17Mit124* | *D17Mit233* | Mean | *D6Mit138* | *D10Mit20* | *D15Mit16* | *34* | *105* | *150* | Mean |
| Total no. of allele | 14 | 21 | 17 | 13 | 15 | 16.0 | 12 | 9 | 12 | 9 | 7 | 12 | 10.2 |
| Average no. of allele per population[b] | 6.0 | 11.4 | 8.9 | 5.9 | 7.6 | 8.0 | 7.6 | 6.5 | 6.8 | 6.8 | 4.1 | 6.9 | 6.5 |
| Logarithm of variance in repeat number | 3.513 | 3.256 | 3.180 | 0.771 | 2.352 | 2.614 | 0.642 | 1.164 | 1.225 | 0.832 | 0.633 | 1.349 | 0.974 |
| Variance in repeat number ($V_{length\text{-}corrected}$) | 1.771 | 0.739 | 0.972 | 0.106 | 0.443 | 0.806 | 0.088 | 0.165 | 0.161 | 0.122 | 0.211 | 0.146 | 0.149 |
| Observed heterozygosity ($H_o$)[b] | 0.615 | 0.686 | 0.540 | 0.561 | 0.674 | 0.615 | 0.579 | 0.705 | 0.711 | 0.590 | 0.366 | 0.467 | 0.570 |
| Expected heterozygosity ($H_e$)[b] | 0.679 | 0.810 | 0.686 | 0.640 | 0.675 | 0.698 | 0.690 | 0.753 | 0.739 | 0.664 | 0.490 | 0.546 | 0.647 |
| $F_{ST}$ value | 0.199 | 0.171 | 0.183 | 0.175 | 0.215 | 0.193 | 0.137 | 0.132 | 0.147 | 0.216 | 0.198 | 0.269 | 0.201 |
| % neutrality rejected | 57.1 | 42.9 | 14.3 | 28.6 | 14.3 | 31.4 | 14.3 | 57.1 | 28.6 | 14.3 | 0 | 14.3 | 21.4 |

[a] All values are averaged for the seven populations that have been surveyed for both of the MHC and non-MHC loci.
[b] Values were weighted by population size.

the northeastern region ($F_{ST} = 0.2029$), whereas the southwestern and the central regions were the least differentiated ($F_{ST} = 0.0495$ and 0.0791, respectively) (Table 4).

*Test of neutrality*

Although MHC-linked microsatellite loci are not directly involved in the function of MHC molecules, strong selection acting on the MHC coding regions may have effects extended to these microsatellite loci that are otherwise evolving neutrally. Since an even distribution in allele frequency characterizes the highly polymorphic MHC loci, we applied the homozygosity test (Watterson, 1978) to assess whether the neutral MHC-linked microsatellite loci give similar evidence of non-neutrality, serving as markers for adjacent sites under selection. The results showed heterogeneity as to rejection of neutrality across populations (data not shown) and across loci (Table 5). The proportion of populations for which neutrality was rejected at the 5 or 10% significance level varied among loci. The loci at which the largest proportion of populations (31.6%) rejected neutrality were *D17Mit33*, *D17Mit21*, and *D17Mit124*. At all five loci, the average ratios of observed to expected homozygosity ($F_{obs}/F_{exp}$) across all populations were less than 1 (i.e., $F_{obs}$ smaller than $F_{exp}$), and similar to each other (0.8675–0.8846).

*Comparison of genetic variation and $F_{ST}$ values between MHC and non-MHC microsatellite loci*

To evaluate the neutral and selective forces influencing the MHC-linked microsatellite polymorphism, we further compared its extent and pattern with that of six non-MHC-linked microsatellite loci, in seven populations that have been surveyed for both categories of markers (Table 6). The total number of alleles for the MHC loci was significantly greater than that for the non-MHC loci (*t*-test; $P = 0.010$), ranging from 13 to 21 for the MHC loci ($\bar{X} = 16.0$, SD = 3.2), and from 7 to 12 for the non-MHC loci ($\bar{X} = 10.2$, SD = 2.1).

Variance in the number of repeats, another measure of diversity, was ln-transformed such that it was normally distributed. For the MHC-linked microsatellite loci, the natural logarithms of variance in repeat number were significantly higher than for the non-MHC microsatellite loci (*t*-test; $P = 0.03$) (Table 6). Length-corrected variances in repeat number ($V_{length\text{-}corrected}$) were also compared between two classes of loci, as microsatellite variability critically depends on the repeat number at a locus. They were still higher for the MHC (0.106–1.771) than non-MHC loci (0.088–0.211).

However, the other estimators, such as the averaged number of allele per population, the heterozygosity, and the $F_{ST}$ value, were similar and not significantly different between both categories of markers (Table 6). Finally, from the result of Watterson's neutrality test, the proportions of populations for which neutrality was rejected at the five MHC loci ranged from 14.3 to 57.1% ($\bar{X} = 31.4\%$); whereas those at the six non-MHC loci ranged from 0 to 57.1% ($\bar{X} = 21.4\%$) (Table 6).

## Discussion

*Level of genetic variation in MHC*

MHC polymorphism in humans (known as *HLA*) is one of the few cases of adaptive evolution well-documented at the molecular level. In the present study, we employed five MHC-linked microsatellite loci to assay genetic variation in the MHC region of southeast Asian house mouse, *M. m. castaneus*, and found there were some differences in levels of allelic variation among the five microsatellite loci, as measured by the total number of alleles and the expected heterozygosity ($H_e$). Of those five microsatellites, *D17Mit21*, located within the intron 3 of the *A-beta-2* gene, and *D17Mit28*, in the promoter region of the *K* gene, exhibited higher degrees of polymorphism than the others. This result was in agreement with the fact that the $A_\beta$ gene in the class II family and the *K* gene in the class I family are among the most polymorphic MHC genes (Hood, Steinmetz & Malissen, 1983). By contrast, *D17Mit33*, located within the class III *C4/Slp hybrid 3* gene, showed less polymorphism than the others. It was consistent with the DNA sequence and allozyme data from previous surveys showing low genetic variation at the class III genes, in spite of their being closely linked to highly polymorphic class I and II genes (Hood, Steinmetz & Malissen, 1983; Klitz, Thomson & Baur, 1986; Satta, Li & Takahata, 1998). The fact that some genes located within the MHC region exhibit extensive polymorphism, whereas other do not, is an interesting feature that may reflect functional requirements for diversity in the polymorphic loci (Hood, Steinmetz & Malissen, 1983).

*Comparing genetic variation between MHC and non-MHC microsatellite loci*

We compared the genetic variability of the MHC-linked with that of the non-MHC-linked microsatellite loci, attempting to assess the influence of strong linkage disequilibrium extending across the MHC region. Here we presumed the six non-coding, non-MHC-linked microsatellite loci to be selectively neutral. Although there are a few notable exceptions (e.g., trinucleotide expansions) (Thornton et al., 1997; Koob et al., 1999; Rubinsztein, 1999), most microsatellites appear to conform to the expectations for neutral loci (Watkins, Bamshad & Jorde, 1995; Hambuch & Lacey, 2000). Variation at these loci should be independent of selection, but instead should reflect demographic parameters such as effective population size. Therefore, analyses of the non-MHC microsatellites provide a potentially valuable complement to studies of the microsatellites closely linked to functional MHC loci.

The greater genetic variation at the five MHC-linked than the six non-MHC-linked microsatellite loci (Table 6) was revealed by the total number of alleles and the variance in repeat number. It was consistent with the result from computer simulations that showed that for a microsatellite locus closely linked to a locus under overdominant selection, their variance in the number of repeats increased (Slatkin, 1995), reflecting the longer-than-expected coalescence times (Hudson & Kaplan, 1988; Kaplan, Darden & Hudson, 1988).

Higher diversity at the MHC-linked microsatellite loci also agrees with allozyme data, showing increased variation at MHC-linked allozyme genes (Nadeau, Collins & Klein, 1982). This might be attributed to the 'linkage' of neutral microsatellite alleles to selectively maintained MHC alleles (Maynard Smith & Haigh, 1974; Kaplan, Hudson & Langley, 1989; Ellegren, Davies & Andersson, 1993). In sheep, the length variation of one microsatellite immediately adjacent to a class II gene *DRB* was found highly correlated with the sequence polymorphism at *DRB* (Paterson, 1998). In humans, it has been shown that balancing selection operating at a class *IA* locus can strongly influence polymorphism within some 100 kb (Satta & Takahata, 2000). The extent of polymorphism in a linked neutral sub-region is sensitive to its recombination rate with the nearest selected locus (Satta, 1997). Other similar examples associated with close linkage to a locus under balancing selection include the increased silent variation in the *S* alleles of plants (Clark & Kao, 1991; Richman, Uyenoyama & Kohn, 1996) and at the MHC loci in humans and mice (Hughes & Nei, 1989; Grimsley, Mather & Ober, 1998; O'hUigin et al., 2000).

A high degree of microsatellite variation in the MHC region may also be caused by high underlying mutation rates. Although we cannot exclude this possibility completely, it was not supported by the finding of Melvold, Wang and Kohn (1997) that the spontaneous mutation rate of *H2* class I genes, except for a single allele at one locus ($H2\text{-}K^b$), appears to be equivalent to that found for non-*H2* histocompatibility genes and comparable to rates reported for a variety of other mouse genes. Also, studies by Satta et al. (1993) on nucleotide substitutions in primates indicated that the mutation rate at the MHC loci is no higher than that of other non-MHC loci despite their extraordinary polymorphism.

Based on the assumption that the extent and patterns of variation for the MHC and the non-MHC loci would differ greatly if the selective forces acting on MHC are of significantly greater magnitude than the non-selective forces, we further compared the allele frequency distributions against neutral expectations for these two categories of microsatellite loci. This result, however, did not provide strong evidence for selection acting on the MHC region (Table 6). For the five MHC-linked and the six non-MHC-linked loci, the proportions of populations for which neutrality was rejected did not differ significantly ($P = 0.4099$).

Considering the low power of the homozygosity test, we applied a modified Lewontin–Krakauer (L–K) test comparing the mean $F_{ST}$ values for two classes of loci with various types of polymorphism (Lewontin & Krakauer, 1973; Pogson, Mesa & Boutilier, 1995; Barker et al., 1997), to the seven mouse populations that have been surveyed for both categories of loci. A simulation study showed that, if there is population subdivision, balancing selection leads to decreased expected $F_{ST}$ values for neutral sites linked to the selected locus (Schierup, Charlesworth & Vekemans, 2000b). With balancing selection, two causes can lead to lower degree of genetic differentiation among proximate populations. First, since alleles are kept in more equal frequencies than for a neutral locus, balancing selection increases within-deme diversity $H_S$, relative to total diversity, thus decreasing the numerator in the expression for $F_{ST}(= (H_T - H_S)/H_T)$. Second, an incoming migrant allele new to a deme will be selected for and thereby maintained, increasing its chance of

invasion compared with that of a neutral allele, that is, the effective migration rate $m_e$ is higher (Schierup, Vekemans & Charlesworth, 2000a). Both contribute to reduction in $F_{ST}$ values at overdominant loci among populations. A comparison of the $F_{ST}$ values of overdominant loci with those of neutral loci should provide evidence of balancing selection.

The result of the L–K test, however, did not show a significant indication of selection ($F_{ST}(\text{MHC})/F_{ST}(\text{non-MHC}) = 0.193/0.201$). In addition to low power of the statistic test, there are two possibilities that might help explain failure to detect balancing selection. First, although a locus closely linked to a locus under balancing selection will show higher polymorphism than a neutral locus (Nei & Li, 1980; Strobeck, 1983), the polymorphism seen at such a linked locus will be a function of the strength of linkage to the selected locus (Hughes & Yeager, 1998). In addition, even though the microsatellite loci we used are within selected genes, they mostly reside in the non-coding region or the introns of genes, inasmuch as their mutation would disrupt any open reading frame (Schlötterer, Amos & Tautz, 1991). However, balancing selection acts primarily on exons, particularly those encoding the PBR, rather than on introns. Data from human class I loci *HLA-A*, *-B*, and *-C* have shown that nucleotide diversity is generally lower in introns than in exons (Hughes & Yeager, 1998; Hughes, 2000). In addition, despite being linked to exon sequences to some extent, intron sequences are subject to homogenization by interallelic recombination, subsequent genetic drift, and loss of polymorphism (Hughes & Yeager, 1998; Meyer & Blasczyk, 2000).

Secondly, other evolutionary forces, such as genetic drift, may have a profound effect in shaping MHC diversity in the populations we surveyed. If that is the case, the extent and pattern of variation at the two types of microsatellite markers may not differ much; the overall $F_{ST}$ values estimated from MHC- and non-MHC-linked microsatellite loci did not differ significantly. The similar population structure estimated from the two classes of loci might indicate that drift and limited migration have been more important than selection in shaping inter-population differentiation at the MHC-linked microsatellite loci. Comparable results have been found in bighorn sheep (Boyce et al., 1997) and in Atlantic salmon (Landry & Bernatchez, 2001), in which MHC genes and microsatellites outside the MHC region gave similar $F_{ST}$ estimates across regions. In addition, recent discoveries of MHC variation in isolated human populations have suggested that the effects of balancing selection are not so strong as to override significant effects of genetic drift (Belich et al., 1992; Watkins et al., 1992; Titus-Trachtenberg, Bugawan & Erlich, 1994). As these mouse populations inhabiting granaries were typically ephemeral and unstable due to regular turnover of grain within a 2–3-year period and occasional applications of poisons (Chou et al., 1998), the influence of genetic drift may have a major effect although our results also indicate that selection continues to play a role in shaping the pattern of MHC variation.

There may be other problems with using microsatellite loci in this context. For example, the mutation rate and/or mutation process may actually vary among microsatellite loci (Harr et al., 1998). Several parameters including the number of repeats, repeat motif length, and repeat motif composition, have been considered to influence microsatellite variability (Schlötterer & Tautz, 1992; Schlötterer et al., 1998; Schug et al., 1998; Estoup & Cornuet, 1999; Bachtrog et al., 2000). Furthermore, the mutation process of microsatellites might not be as simple as first assumed for microsatellites and may follow a two-phase model. The prevalence of different mechanisms varies according to the structure of the repeat unit itself (Shriver et al., 1993; Estoup & Cornuet, 1999), confounding the interpretation of the data. However, even if the microsatellites cannot provide conclusive evidence of the effects of linked selected loci, their abundance in genomes and ease of survey are still believed to complement and test hypotheses based on comparisons of DNA sequences.

### *Genetic differentiation among populations and regions*

Our results suggest that the seven populations in the central and southwestern regions of Taiwan might be part of a metapopulation structure. Within this region (central–southwestern), the mean $F_{ST}$ value was relatively small both for the MHC (0.091) and for non-MHC microsatellite loci (0.104) (data not shown), suggesting that there may be high gene flow between these populations. By contrast, the northwestern region appeared to be very different from these two regions. From Tables 3 and 4, it appears that the sharpness of the differentiation between the central–southwestern and the northwestern regions was attributable to two populations (Guanshi and Miaoli).

If these two populations were excluded, the mean $F_{ST}$ value of the western part of Taiwan (northwestern + central + southwestern regions) would be lowered from 0.161 to 0.108 for the five MHC microsatellite loci. The frequent gene exchange, implied by the low genetic differentiation among the populations in the western part of Taiwan, may occur through some pockets of feral populations (Chou et al., 1998), or through long-distance dispersal associated with human activities (Yu & Peng, 2002).

Among the 19 populations, there are some lines of evidence suggesting the Guanshi might be a population established by a few founders, or subject to extinction–recolonization events. First, it was the least polymorphic population of all, with the lowest heterozygosity (Table 2) and effective number of alleles ($n_e$) (data not shown). Second, this population was genetically distinct from most other populations, and in particular, it was quite different from its neighboring population (10.5 km apart only) of Shinpu (Table 3). This gave northwestern Taiwan the greatest inter-population differentiation (Table 4). Third, most of the homozygosity values for Guanshi were higher than the expected neutral value (i.e., $F_{obs}/F_{exp} > 1$), either at the five MHC loci ($F_{obs}/F_{exp} = 1.4055$, 1.9524, 1.6686, and 1.5796 for *D17Mit33*, *D17Mit21*, *D17Mit28*, and *D17Mit233*), or at the six non-MHC loci ($F_{obs}/F_{exp} = 1.4948$, 1.1544, 1.2377, and 1.1827 for *D6Mit138*, *D15Mit16*, *34*, and *150*). In general, a value of homozygosity higher than the expected neutral value could reflect founder effects and population bottlenecks, genetic drift, or positive directional selection for an allele (Mack et al., 2000). As a result, the bottleneck/founder effect appeared to be the explanation most consistent with the low genetic variation found in Guanshi.

On the other hand, the genetic diversity of Jinmen's mouse population was the greatest among all the populations studied (Table 2, Appendix A), even though this offshore island has an area of only 132 km$^2$. Furthermore, the population of Tainan, where the early settlement of Han Chinese people began, was the one most similar to that of Jinmen (Table 3) in addition to Taichung. Interestingly, Tainan was also the only population with a genetic diversity comparable to Jinmen (Table 2, Appendix A). Therefore, we presume that the mice from Jinmen, only 1–2 km away from Mainland China, are representative of mouse genomes from Southeast China whence certain founders of modern Taiwanese mouse populations might have originated. And the close genetic relatedness between Jinmen's and Tainan's populations, separated by the Taiwan Strait, indicates that there must have been human-mediated gene flow in *M. m. castaneus* populations due to historical agriculture expansion and recent frequent traffic (Yu & Peng, 2002).

## Acknowledgements

We thank Hwei-Yu Chang, Wen-Hsiung Li, Manyuan Long, Janice Spofford, Christian Schlötterer and four anonymous reviewers for their comments on earlier drafts of the manuscript. The comments help to improve the manuscript. This research was supported by the National Science Council of the Republic of China.

## References


Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith & K. Struhl, 1995. Short Protocols in Molecular Biology – A Compendium of Methods from Current Protocols in Molecular Biology, Wiley, Canada, 3rd edn.

Bachtrog, D., M. Agis, M. Imhof & C. Schlötterer, 2000. Microsatellite variability differs between dinucleotide repeat motifs-evidence from *Drosophila melanogaster*. Mol. Biol. Evol. 17: 1277–1285.

Barker, J.S.F., S.S. Moore, D.J.S. Hetzel, D. Evans, S.G. Tan & K. Byrne, 1997. Genetic diversity of Asian water buffalo (*Bubalus bubalis*): microsatellite variation and a comparison with protein-coding loci. Anim. Genet. 28: 103–115.

Belich, M.P., J.A. Madrigal, W.H. Hildebrand, J. Zemmour, R.C. Williams, R. Luz, M.L. Petzl-Erler & P. Parham, 1992. Unusual *HLA-B* alleles in two tribes of Brazilian Indians. Nature 257: 326–329.

Boyce, W.M., P.W. Hedrick, N.E. Muggli-Cockett, S. Kalinowski, M.C.T. Penedo & R.R. Ramey, 1997. Genetic variation of major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. Genetics 145: 421–433.

Chou, C.-W., P.-F. Lee, K.-H. Lu & H.-T. Yu, 1998. A population study of house mice (*Mus musculus castaneus*) inhabiting rice granaries in Taiwan. Zool. Stud. 37(3): 201–212.

Clark, A. & T. Kao, 1991. Excess nonsynonymous substitutions at shared polymorphic sites among self-incompatibility alleles of Solanaceae. Proc. Natl. Acad. Sci. USA 88: 9823–9827.

Dallas, J.F., B. Dod, P. Boursot, E.M. Prager & F. Bonhomme, 1995. Population subdivision and gene flow in Danish house mice. Mol. Ecol. 4: 311–320.

Dietrich, W.F., J. Miller, R. Steen, M.A. Merchant, D. Damron-Boles, Z. Husain, R. Dredge, M.J. Daly, K.A. Ingalls, T.J. O'Connor, C.A. Evans, M.M. DeAngelis, D.M. Levinson, L. Kruglyak, N. Goodman, N.G. Copeland, N.A. Jenkins, T.L. Hawkins, L. Stein, D.C. Page & E.S. Lander, 1996. A comprehensive genetic map of the mouse genome. Nature 380: 149–152.

Ellegren, H., C.J. Davies & L. Andersson, 1993. Strong association between polymorphisms in an intronic microsatellite and in the coding sequence of the *BoLA-DRB3* gene: implications for

microsatellites stability and PCR-based *DRB3* typing. Anim. Genet. 24: 269–275.

Estoup, A. & J.-M. Cornuet, 1999. Microsatellite evolution: inferences from population data, pp. 49–65 in Microsatellites: Evolution and Application, edited by D.B. Goldstein & C. Schlötterer. Oxford University Press, New York.

Ewens, W.J., 1972. The sampling theory of selectively neutral alleles. Theor. Pop. Biol. 3: 87–112.

Goldstein, D.B., A.R. Linares, L.L. Cavalli-Sforza & M.W. Feldman, 1995. Genetic absolute dating based on microsatellites and the origin of modem humans. Proc. Natl. Acad. Sci. USA 92: 6723–6727.

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate *F*-statistics. J. Hered. 86: 485–486.

Grimsley, C., K.A. Mather & C. Ober, 1998. *HLA-H*: a pseudogene with increased variation due to balancing selection at neighboring loci. Mol. Biol. Evol. 15(12): 1581–1588.

Guo, S.W. & E.A. Thompson, 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48: 361–372.

Hambuch, T.M. & E.A. Lacey, 2000. Contrasting patterns of *MHC* and microsatellite diversity in social and solitary tuco-tucos (Rodentia: Ctenomyidae), pp. 542–554 in Major Histocompatibility Complex: Evolution, Structure, and Function, edited by M. Kasahara. Springer, Tokyo.

Harr, B., B. Zangerl, G. Brem & C. Schlötterer, 1998. Conservation of locus specific microsatellite variability across species: a comparison of two *Drosophila* sibling species *D. melanogaster* and *D. simulans*. Mol. Biol. Evol. 15: 176–184.

Hood, L., M. Steinmetz & B. Malissen, 1983. Genes of the major histocompatibility complex of the mouse. Ann. Rev. Immunol. 1: 529–568.

Hudson, R.R. & N.L. Kaplan, 1988. The coalescent process in models with selection and recombination. Genetics 120: 831–840.

Hughes, A.L., 2000. Balancing selection: the major histocompatibility complex, pp. 54–89 in Adaptive Evolution of Genes and Genomes. Oxford University Press, New York.

Hughes, A.L. & M. Nei, 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc. Natl. Acad. Sci. USA 86: 958–962.

Hughes, A.L. & M. Yeager, 1998. Natural selection at major histocompatibility complex loci of vertebrates. Ann. Rev. Genet. 32: 415–435.

Huttley, G.A., M.W. Smith, M. Carrington & S.J. O'Brein, 1999. A scan for linkage disequilibrium across the human genome. Genetics 152: 1711–1722.

Kaplan, N.L., T. Darden & R.R. Hudson, 1988. The coalescent process in models with selection. Genetics 120: 819–829.

Kaplan, N.L., R.R. Hudson & C.H. Langley, 1989. The hitchhiking effect revisited. Genetics 123(4): 887–899.

Karl, S.A. & J.C. Avise, 1993. Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. Science 256: 100–102.

Klitz, W., G. Thomson & M.P. Baur, 1986. Contrasting evolutionary histories among tightly linked *HLA* loci. Am. J. Hum. Genet. 39: 340–349.

Kohn, M.H., H.-J. Pelz & R.K. Wayne, 2000. Natural selection mapping of the warfarin-resistance gene. Proc. Natl. Acad. Sci. USA 97(14): 7911–7915.

Koob, M.D., M.L. Moseley, L.J. Schut, K.A. Benzow, T.D. Bird, J.W. Day & L.P.W. Ranum, 1999. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). Nat. Genet. 21(4): 379–384.

Landry, C. & L. Bernatchez, 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). Mol. Ecol. 10: 2525–2539.

Lewontin, R.C. & J. Krakauer, 1973. Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. Genetics 74: 175–195.

Lynch, M., M. Pfrender, K. Spitze, N. Lehman, J. Hicks, D. Allen, L. Latta, M. Ottene, F. Bogue & J. Colbourne, 1999. The quantitative and molecular genetic architecture of a subdivided species. Evolution 53: 100–110.

Mack, S.J., T.L. Bugawan, M. Stoneking, M. Saha, H.-P. Beck & H.A. Erlich, 2000. *HLA* class I and class II loci in Pacific/Asian populations, pp. 427–445 in Major Histocompatibility Complex: Evolution, Structure, and Function, edited by M. Kasahara. Springer, Tokyo.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209–220.

Maynard Smith, J. & J. Haigh, 1974. The hitch-hiking effect of a favorable gene. Genet. Res. Camb. 23: 23–35.

Meagher, S. & W.K. Potts, 1997. A microsatellite-based *MHC* genotyping system for house mice (*Mus domesticus*). Hereditas 127: 75–82.

Melvold, R.W., K. Wang & H.I. Kohn, 1997. Histocompatibility gene mutation rates in the mouse: a 25-year review. Immunogenetics 47: 44–54.

Meyer, D. & R. Blasczyk, 2000. The effect of mutation, recombination and selection on *HLA* non-coding sequences, pp. 398–411 in Major Histocompatibility Complex: Evolution, Structure, and Function, edited by M. Kasahara. Springer, Tokyo.

Minch, E., A. Ruiz-Linares, D. Goldstein, M. Feldman & L.L. Cavalli-Sforza, 1995. Microsat (version 1.4d): a computer program for calculating various statistics on microsatellite allele data.

Nadeau, J.H., R.L. Collins & J. Klein, 1982. Organization and evolution of the mammalian genome. I. Polymorphism of *H-2* linked loci. Genetics 102: 583–598.

Nei, M., 1972. Genetics distance between populations. Am. Nat. 106: 283–292.

Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Nei, M. & W.-H. Li, 1980. Non-random association between electromorphs and inversion chromosomes in finite populations. Genet. Res. Camb. 35: 65–83.

Nei, M. & A.L. Hughes, 1991. Polymorphism and evolution of the Major histocompatibility complex loci in mammals, pp. 222–247 in Evolution at the Molecular Level, edited by R.K. Selander, A.G. Clark & T.S. Whittam. Sinauer Associates Inc., Sunderland, MA.

O'hUigin, C., Y. Satta, A. Hausmann, R.L. Dawkins & J. Klein, 2000. The implications of intergenic polymorphism for major histocompatibility complex evolution. Genetics 156(2): 867–877.

Paterson, S., 1998. Evidence for balancing selection at the major histocompatibility complex in a free-living ruminant. J. Hered. 89: 289–294.

Pogson, G.H., K.A. Mesa & R.G. Boutilier, 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics 139: 376–385.

Potts, W.K. & E.K. Wakeland, 1993. Evolution of *MHC* genetic diversity: a tale of incest, pestilence and sexual preference. Trends Genet. 9(12): 408–412.

Raymond, M. & F. Rousset, 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Heredity 86: 248–249.
Richman, R., M. Uyenoyama & J. Kohn, 1996. Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. Science 273: 1212–1216.
Rousset, F., 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145: 1219–1228.
Rubinsztein, D.C., 1999. Trinucleotide expansion mutation cause diseases which do not conform to classical Mendelian expectation, pp. 80–97 in Microsatellites – Evolution and Applications, edited by D.B. Goldstein & C. Schlötterer. Oxford University Press, New York.
Satta, Y., 1997. Effects of intra-locus recombination of *HLA* polymorphism. Hereditas 127: 105–112.
Satta, Y., Y.-J. Li & N. Takahata, 1998. The neutral theory and natural selection in the *HLA* region. Frontier Biosci. 3: d459–d467.
Satta, Y. & N. Takahata, 2000. Polymorphism in the *HLA* class I region, pp. 398–411 in Major Histocompatibility Complex: Evolution, Structure, and Function, edited by M. Kasahara. Springer, Tokyo.
Satta, Y., C. O'hUigin, N. Takahata & J. Klein, 1993. The synonymous substitution rate of the major histocompatibility complex loci in primates. Proc. Natl. Acad. Sci. USA 90: 7480–7484.
Satta, Y., C. O'hUigin, N. Takahata & J. Klein, 1994. Intensity of natural selection at the major histocompatibility complex loci. Proc. Natl. Acad. Sci. USA 91: 7184–7188.
Schierup, M.H., X. Vekemans & D. Charlesworth, 2000a. The effect of subdivision on variation at multi-allelic loci under balancing selection. Genet. Res. Camb. 76: 51–62.
Schierup, M.H., D. Charlesworth & X. Vekemans, 2000b. The effect of hitch-hiking on genes linked to a balanced polymorphism in a subdivided population. Genet. Res. Camb. 76: 63–73.
Schlötterer, C. & D. Tautz, 1992. Slippage synthesis of simple sequence DNA. Nucl. Acid Res. 20: 211–215.
Schlötterer, C. & T. Wiehe, 1999. Microsatellites, a neutral marker to infer selective sweeps, in Microsatellites: Evolution and Application, edited by D.B. Goldstein & C. Schlötterer. Oxford University Press, New York.
Schlötterer, C., B. Amos & D. Tautz, 1991. Conservation of polymorphic simple sequence loci in cetacean species. Nature 354: 63–65.
Schlötterer, C., C. Vogl & D. Tautz, 1997. Polymorphism and locus-specific effects on polymorphism at microsatellite loci in natural *Drosophila melanogaster* populations. Genetics 146: 309–320.
Schlötterer, C., R. Ritter, B. Harr & G. Brem, 1998. High mutation rates of a long microsatellite allele in *Drosophila melanogaster* provide evidence for allele-specific mutation rates. Mol. Biol. Evol. 15: 1269–1274.
Schug, M.D., C.M. Hutter, K.A. Wetterstrand, M.S. Gaudette, T.F. Mackay & C.F. Aquadro, 1998. The mutation rates of di-, tri-, and tetranucleotide repeats in *Drosophila melanogaster*. Mol. Biol. Evol. 15: 1751–1760.
Shriver, M.D., L. Jin, R. Chakraborty & E. Boerwinkle, 1993. VNTR allele frequency distributions under the SMM: A computer simulation approach. Genetics 134: 983–993.
Slatkin, M., 1995. Hitchhiking and associative overdominance at a microsatellite locus. Mol. Biol. Evol. 12(3): 473–480.
Spitze, K., 1993. Population structure in *Daphnia obtuse*: quantitative genetic and allozymic variation. Genetics 135: 367–374.
Strobeck, C., 1983. Expected linkage disequilibrium for a neutral locus linked to a chromosomal arrangement. Genetics 103: 545–555.
Takahata, N. & M. Nei, 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124: 967–978.
Takahata, N., Y. Satta & J. Klein, 1992. Polymorphism and balancing selection at major histocompatibility complex loci. Genetics 130: 925–938.
Thornton, C.A., J.P. Wymer, Z. Simmons, C. McClain & R.T. Moxley III, 1997. Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking *DMAHP* gene. Nat. Genet. 16: 407–409.
Titus-Trachtenberg, E.A., T.L. Bugawan & H.A. Erlich, 1994. Identification of a novel *DQB1* allele (*0609) segregating in an Ashkenazi Jewish family: implications for *DQB1* typing systems. Tissue Antigens 44(2): 120–124.
Watkins, D.I., S.N. McAdam, X. Liu, C.R. Strang, E.L. Milford, C.G. Levine, T.L. Garber, A.L. Dogon, C.I. Lord, S.H. Ghim, G.M. Troup, A.L. Hughes & N.L. Letvin, 1992. New recombinant *HLA-B* alleles in a tribe of South American Amerindians indicate rapid evolution of *MHC* class I loci. Nature 357: 329–333.
Watkins, W.S., M. Bamshad & L.B. Jorde, 1995. Population genetics of trinucleotide repeat polymorphism. Hum. Mol. Genet. 4(9): 1485–1491.
Watterson, G.A., 1978. The homozygosity test of neutrality. Genetics 88: 405–417.
Weir, B.S. & C.C. Cockerham, 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38: 1358–1370.
Yu, H.T. & Y.H. Peng, 2002. Population differentiation and gene flow revealed by microsatellite DNA markers in the house mouse (*Mus musculus castaneus*) in Taiwan. Zool. Sci. 19(3): 475–483.

## Appendix A.

Allele frequencies at the five MHC microsatellite loci in 19 Southeast Asian mouse populations grouped into six geographic regions in Taiwan. Populations were grouped into six geographical regions by locations. Alleles are in sizes (bp). Private alleles in each population are shown bold-faced. See Table 2 for sample sizes and Table 3 for population abbreviations.

| Locus | Allele | Northwestern | | | | Central | | | | | Southwestern | | Southeastern | | | Northeastern | | | | Offshore, JM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | TY | SW | GS | SP | ML | TC | TT | SJ | LN | TN | KP | CS | GA | HL | WC | TS | IC | TE | |
| *D17Mit33* | 158 | – | – | – | – | – | – | – | – | – | **0.037** | – | – | – | – | – | – | – | – | – |
| | 160 | – | – | – | – | – | – | – | 0.008 | 0.005 | – | – | – | – | – | – | – | – | – | – |
| | 172 | 0.139 | 0.467 | 0.929 | – | 0.900 | 0.250 | 0.028 | 0.239 | 0.379 | 0.037 | 0.313 | 0.278 | 0.115 | 0.224 | 0.042 | 0.094 | – | – | – |
| | 174 | – | – | – | – | – | – | – | – | – | – | **0.063** | – | – | – | – | – | – | – | – |
| | 176 | – | – | 0.048 | – | – | 0.063 | – | – | – | – | – | – | – | 0.197 | – | – | – | – | 0.105 |
| | 192 | – | – | – | – | – | – | – | – | – | 0.074 | – | 0.056 | 0.077 | 0.092 | – | – | – | – | – |
| | 194 | 0.028 | 0.400 | 0.028 | 0.200 | – | 0.188 | – | 0.069 | 0.005 | 0.315 | 0.250 | 0.167 | 0.077 | 0.013 | 0.208 | 0.063 | 0.027 | 0.528 | 0.132 |
| | 196 | – | – | – | – | 0.100 | – | – | – | – | – | – | 0.028 | – | 0.184 | 0.042 | – | 0.014 | 0.222 | 0.132 |
| | 198 | – | – | – | 0.450 | – | – | – | 0.100 | 0.303 | 0.296 | – | 0.278 | 0.077 | 0.026 | 0.042 | 0.156 | 0.230 | 0.194 | – |
| | 200 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.028** | – |
| | 202 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.014 | – | 0.053 |
| | 206 | 0.694 | 0.100 | – | 0.350 | – | 0.313 | 0.556 | 0.485 | 0.303 | 0.185 | 0.125 | 0.194 | 0.654 | 0.263 | 0.667 | 0.656 | 0.446 | 0.028 | 0.316 |
| | 208 | – | – | – | – | – | – | – | – | – | 0.037 | – | – | – | – | – | 0.031 | 0.135 | – | – |
| | 210 | 0.139 | 0.033 | – | – | – | 0.188 | 0.417 | 0.100 | 0.005 | 0.019 | 0.250 | – | – | – | – | – | 0.135 | – | 0.158 |
| | 212 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.105** |
| *D17Mit21* | 106 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.026** |
| | 110 | – | 0.033 | – | – | – | 0.063 | 0.167 | 0.262 | 0.187 | 0.185 | 0.056 | – | – | – | – | – | – | 0.321 | 0.395 |
| | 112 | – | 0.033 | – | – | – | – | – | 0.056 | 0.025 | – | – | – | – | – | – | – | 0.015 | – | – |
| | 114 | – | – | – | – | – | – | – | 0.008 | 0.005 | 0.130 | – | – | – | – | – | – | – | – | – |
| | 120 | – | – | – | – | 0.300 | – | – | 0.071 | 0.035 | 0.093 | – | 0.056 | – | 0.066 | 0.182 | 0.800 | 0.030 | – | 0.132 |
| | 122 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.105** |
| | 124 | – | – | – | – | 0.100 | 0.125 | 0.028 | 0.032 | – | – | – | – | – | – | – | – | – | – | – |
| | 126 | – | – | – | – | – | – | – | – | – | 0.019 | – | – | – | – | – | – | – | – | 0.053 |
| | 128 | – | – | – | – | – | 0.063 | – | – | – | – | 0.056 | 0.333 | – | – | – | – | – | – | – |
| | 130 | 0.026 | – | 0.048 | – | – | 0.063 | – | – | – | – | – | – | – | 0.197 | 0.182 | 0.100 | 0.076 | – | 0.105 |
| | 132 | 0.079 | – | – | – | – | 0.188 | – | 0.024 | 0.096 | – | – | – | – | – | 0.227 | – | 0.015 | – | – |
| | 134 | 0.026 | – | – | 0.050 | – | – | – | 0.103 | 0.202 | – | – | 0.167 | – | 0.105 | 0.046 | – | – | 0.036 | 0.026 |
| | 136 | 0.316 | 0.133 | – | 0.150 | 0.100 | 0.063 | 0.083 | 0.119 | 0.106 | 0.167 | 0.167 | – | 0.375 | – | 0.091 | 0.100 | 0.242 | 0.107 | 0.026 |
| | 138 | 0.105 | 0.400 | – | 0.400 | 0.300 | – | 0.111 | 0.071 | 0.005 | 0.074 | – | – | – | 0.013 | – | – | 0.318 | – | – |
| | 140 | 0.026 | 0.033 | 0.024 | 0.150 | – | 0.188 | – | 0.024 | 0.081 | 0.148 | 0.222 | – | – | 0.105 | – | – | 0.015 | – | 0.105 |
| | 142 | 0.211 | 0.033 | 0.048 | – | 0.200 | 0.250 | 0.139 | 0.048 | 0.071 | 0.056 | 0.056 | 0.306 | 0.125 | 0.224 | – | – | – | – | – |
| | 144 | 0.026 | – | 0.024 | – | – | – | 0.472 | 0.127 | 0.126 | 0.037 | 0.111 | – | 0.208 | – | 0.091 | – | 0.015 | 0.036 | 0.026 |
| | 146 | 0.184 | 0.167 | 0.833 | 0.200 | – | – | – | 0.016 | 0.015 | 0.019 | 0.167 | – | 0.042 | 0.013 | – | – | 0.182 | 0.393 | – |
| | 148 | – | 0.133 | 0.024 | – | – | – | – | – | – | – | – | 0.139 | 0.042 | – | – | – | 0.091 | 0.107 | – |
| | 150 | – | – | – | 0.050 | – | – | – | 0.032 | 0.025 | – | – | – | 0.083 | 0.026 | – | – | – | – | – |

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 152 | – | – | – | – | – | – | – | – | 0.020 | 0.037 | 0.167 | – | 0.083 | 0.237 | – | – | – | – | – |
| | 154 | – | – | – | – | – | – | – | – | – | – | – | – | **0.042** | – | – | – | – | – | – |
| | 156 | – | – | – | – | – | – | – | **0.008** | – | – | – | – | – | – | – | – | – | – | – |
| | 158 | – | – | – | – | – | – | – | – | – | 0.037 | – | – | – | 0.013 | – | – | – | – | – |
| | 162 | – | **0.033** | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| | 180 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.091** | – | – | – | – |
| | 204 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.091** | – | – | – | – |
| *D17Mit28* | 86 | 0.079 | – | – | – | – | | 0.028 | 0.008 | 0.030 | 0.093 | 0.111 | 0.278 | – | 0.013 | 0.167 | 0.267 | 0.371 | 0.412 | 0.026 |
| | 90 | 0.553 | 0.700 | 0.071 | 0.300 | 0.500 | 0.313 | 0.306 | 0.192 | 0.460 | 0.333 | 0.389 | – | 0.308 | 0.368 | 0.125 | 0.167 | 0.214 | 0.265 | 0.184 |
| | 92 | 0.105 | – | – | – | – | – | – | 0.008 | 0.005 | 0.204 | – | – | – | – | – | – | – | – | 0.053 |
| | 94 | – | – | – | – | 0.300 | 0.125 | – | – | – | – | – | – | – | – | – | – | 0.171 | – | |
| | 96 | – | – | – | – | – | – | – | 0.062 | 0.030 | 0.037 | – | – | – | 0.026 | 0.292 | 0.533 | 0.043 | – | 0.184 |
| | 102 | 0.026 | 0.033 | – | – | – | – | – | – | – | – | – | – | – | 0.132 | – | – | – | – | – |
| | 104 | – | – | – | – | – | 0.063 | – | 0.062 | 0.030 | 0.111 | – | – | – | – | – | 0.033 | – | – | – |
| | 106 | – | – | – | – | 0.100 | 0.063 | 0.028 | 0.139 | 0.005 | – | – | – | – | 0.316 | – | – | 0.114 | – | 0.026 |
| | 108 | 0.079 | 0.167 | 0.024 | – | – | 0.250 | 0.167 | 0.077 | 0.085 | 0.074 | 0.111 | – | – | – | 0.250 | – | – | 0.029 | 0.184 |
| | 110 | 0.132 | 0.067 | 0.833 | 0.700 | – | 0.125 | 0.250 | 0.339 | 0.305 | – | 0.333 | 0.139 | – | 0.053 | – | – | 0.014 | 0.147 | 0.158 |
| | 112 | – | – | 0.024 | – | – | – | 0.222 | 0.008 | – | – | – | – | 0.115 | – | – | – | 0.014 | – | 0.026 |
| | 114 | – | – | – | – | 0.100 | 0.063 | – | – | – | 0.019 | 0.056 | – | – | – | – | – | 0.043 | 0.118 | 0.053 |
| | 116 | – | 0.033 | 0.048 | – | – | – | – | – | 0.035 | – | – | 0.111 | 0.115 | 0.092 | – | – | – | – | – |
| | 118 | – | – | – | – | – | – | – | – | – | 0.037 | – | 0.472 | 0.423 | – | 0.083 | – | 0.014 | – | 0.105 |
| | 122 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.042** | – | – | – | – |
| | 124 | – | – | – | – | – | – | – | – | – | 0.074 | – | – | – | – | 0.042 | – | – | 0.029 | – |
| | 126 | – | – | – | – | – | – | – | – | **0.005** | – | – | – | – | – | – | – | – | – | – |
| | 128 | – | – | – | – | – | – | – | 0.085 | – | 0.019 | – | – | – | – | – | – | – | – | – |
| | 130 | 0.026 | – | – | – | – | – | – | 0.023 | 0.010 | – | – | – | 0.039 | – | – | – | – | – | – |
| *D17Mit124* | 147 | – | 0.133 | 0.024 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.053 |
| | 149 | – | – | – | – | – | 0.063 | 0.111 | 0.092 | 0.235 | 0.019 | 0.222 | 0.083 | 0.115 | – | – | – | 0.200 | – | – |
| | 150 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.056** | – |
| | 151 | 0.079 | 0.033 | – | – | 0.100 | – | – | – | – | – | – | – | – | – | – | – | 0.217 | 0.111 | 0.053 |
| | 152 | – | – | – | – | – | – | – | – | – | 0.037 | – | 0.028 | – | – | – | – | – | – | – |
| | 153 | 0.105 | 0.267 | 0.024 | – | 0.300 | 0.188 | – | 0.185 | 0.065 | 0.111 | – | 0.361 | 0.308 | 0.027 | – | 0.125 | 0.083 | – | – |
| | 154 | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.014 | – | 0.042 | – | – | – |
| | 155 | 0.658 | 0.367 | 0.500 | 0.800 | 0.600 | 0.250 | 0.472 | 0.200 | 0.315 | 0.426 | 0.389 | 0.417 | 0.346 | 0.892 | 0.792 | 0.250 | 0.300 | 0.611 | 0.395 |
| | 156 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.053** |
| | 157 | – | – | 0.452 | – | – | 0.438 | 0.278 | 0.154 | 0.060 | 0.148 | – | – | – | – | – | 0.375 | 0.017 | 0.222 | 0.342 |
| | 158 | 0.053 | – | – | 0.200 | – | – | 0.139 | 0.254 | 0.055 | – | – | – | – | – | 0.042 | 0.125 | 0.033 | – | – |
| | 159 | – | – | – | – | – | – | – | 0.008 | 0.205 | 0.074 | – | – | – | 0.068 | 0.042 | 0.083 | 0.150 | – | 0.105 |
| | 160 | – | – | – | – | – | – | – | – | – | – | – | 0.028 | 0.231 | – | – | – | – | – | – |

(*continued*)

Appendix A (*continued*)

| Locus | Allele | Northwestern | | | | Central | | | | | Southwestern | | Southeastern | | | Northeastern | | | | Offshore, JM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | TY | SW | GS | SP | ML | TC | TT | SJ | LN | TN | KP | CS | GA | HL | WC | TS | IC | TE | |
| | 161 | – | 0.167 | – | – | – | 0.063 | – | 0.008 | 0.065 | 0.167 | 0.389 | 0.056 | – | – | – | – | – | – | – |
| | 162 | 0.105 | 0.033 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| | 166 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.083** | – | – | – | – |
| | 167 | – | – | – | – | – | – | – | 0.085 | – | 0.019 | – | 0.028 | – | – | 0.042 | – | – | – | – |
| | 169 | – | – | – | – | – | – | – | **0.015** | – | – | – | – | – | – | – | – | – | – | – |
| *D17Mit233* | 102 | – | – | 0.048 | – | – | 0.063 | – | 0.069 | 0.035 | 0.074 | 0.056 | – | – | 0.197 | – | 0.071 | 0.063 | – | 0.105 |
| | 104 | – | 0.033 | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.036 | 0.016 | – | – |
| | 106 | 0.026 | 0.300 | – | 0.150 | | 0.375 | 0.639 | 0.546 | 0.585 | 0.148 | 0.111 | 0.361 | 0.654 | 0.329 | – | 0.071 | 0.203 | 0.769 | 0.053 |
| | 108 | – | 0.100 | 0.024 | 0.100 | – | – | – | 0.008 | 0.010 | – | 0.056 | – | – | – | – | – | 0.016 | – | – |
| | 110 | 0.237 | – | 0.810 | – | – | – | – | – | – | 0.093 | – | – | – | 0.303 | 0.050 | 0.179 | 0.078 | – | 0.158 |
| | 112 | – | – | – | – | – | – | – | – | – | 0.056 | 0.056 | – | – | – | – | – | – | – | 0.026 |
| | 114 | 0.026 | – | – | 0.250 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.053 |
| | 116 | 0.132 | 0.167 | | 0.050 | 0.300 | 0.186 | | 0.115 | 0.120 | 0.389 | 0.333 | 0.083 | 0.192 | 0.066 | 0.150 | – | – | – | 0.053 |
| | 118 | – | – | – | – | 0.100 | 0.125 | 0.083 | 0.015 | – | – | 0.056 | 0.028 | – | 0.079 | – | 0.036 | 0.141 | 0.039 | 0.053 |
| | 120 | 0.316 | 0.033 | 0.048 | 0.450 | – | – | – | 0.131 | 0.190 | 0.148 | 0.056 | 0.250 | – | 0.013 | 0.350 | 0.250 | 0.188 | 0.039 | 0.105 |
| | 122 | 0.053 | 0.233 | 0.071 | – | – | – | – | 0.008 | 0.035 | 0.019 | 0.222 | 0.250 | – | – | 0.100 | 0.143 | 0.266 | – | 0.158 |
| | 124 | – | – | – | – | 0.100 | 0.250 | 0.278 | 0.108 | 0.025 | – | – | 0.028 | 0.154 | – | – | – | 0.031 | – | – |
| | 126 | 0.211 | – | – | – | 0.100 | – | – | – | – | 0.037 | – | – | – | 0.013 | – | 0.071 | – | – | 0.053 |
| | 128 | – | – | – | – | 0.100 | – | – | – | – | 0.037 | 0.056 | – | – | – | 0.350 | 0.143 | – | 0.115 | 0.026 |
| | 130 | – | 0.133 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | 0.039 | 0.079 |
| | 132 | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | – | **0.079** |
| | 136 | – | – | – | – | **0.300** | – | – | – | – | – | – | – | – | – | – | – | – | – | – |
| Mean allele no./locus | | 6.4 | 6.2 | 4.6 | 3.6 | 4.0 | 6.0 | 4.4 | 9.8 | 9.0 | 9.4 | 6.0 | 5.6 | 5.0 | 5.0 | 6.0 | 5.4 | 8.4 | 5.2 | 9.2 |
| H–W exact test[a] | | * | ns[b] | *** | ns | ns | ns | ns | *** | *** | *** | ns | *** | *** | * | ** | ns | *** | *** | ns |

[a] Significant levels of H–W probability test.

[b] $P > 0.05$.

* $0.01 < P < 0.05$.

** $0.001 < P < 0.01$.

*** $P < 0.001$.