Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference

WAYNE K. POTTS AND EDWARD K. WAKELAND

Evidence from the house mouse (Mus) suggests that the extreme diversity of genes of the major histocompatibility complex (MHC) results from three different forms of selection involving infectious disease (pestilence), inbreeding (incest) and MHC-based mating (sexual) preferences. MHC-based disassortative mating preferences are presumed to have evolved because they reduce homozygosity throughout the genome, and particularly within loci linked to the MHC. Progeny derived from such disassortative matings would enjoy increased fitness because of both reduced levels of inbreeding depression and increased resistance to infectious disease arising from their increased MHC heterozygosity.

Products of genes within the major histocompatibility complex (MHC) play a critical role during immune recognition, binding self and foreign peptide fragments for presentation to T lymphocytes[1,2]. This crucial immune function has led to the widely held view that the unprecedented genetic diversity of MHC genes results from pathogen-driven selection that favors either MHC heterozygotes (heterozygote advantage or overdominance), or relatively rare MHC genotypes (negative frequency-dependent selection), or a combination of both[3–7]. As discussed below, both types of selection can maintain high levels of genetic diversity[6]. In studies that may appear to suggest an alternative view, mice have been shown to prefer mates whose MHC is different from their own, both under laboratory conditions[8–11] and in seminatural populations[12]. Furthermore, these MHC-based disassortative mating preferences also maintain diversity, and are strong enough to account for most of the genetic diversity observed in *Mus* populations[12,13]. However, there has been resistance to incorporating the consideration of mating patterns into conventional thinking on MHC genetic diversity, primarily because there has seemed to be little connection between MHC-based mating preferences and the acknowledged role of MHC genes in immune recognition. Here, we review the relevant data and theories, and show that both infectious disease and inbreeding depression favor the evolution of MHC-based mating preferences in vertebrate species that, like many mammals, can detect MHC genotype by smell[14,15]. In the context of current knowledge of how MHC genes influence disease resistance, individual odors, mating patterns and inbreeding, we develop a theoretical framework for the evolution of MHC diversity in such species.

Overview

Based primarily on what is known in *Mus*, we advance the following hypothesis for the evolution of MHC genetic diversity. Pathogen-driven selection favors genetic diversity of the MHC through both heterozygote advantage[3,4,16] and frequency-dependent selection[7,17]. This in turns favors the evolution of MHC-based disassortative mating preferences because such matings would preferentially produce MHC-heterozygous progeny that have increased fitness. MHC-based mating preferences are possible because MHC polymorphisms strongly influence individual odors[15]; such mating preferences would further increase MHC genetic diversity[5,12,13], increasing the usefulness of these loci as markers for kin recognition[18–20]. Consequently, the avoidance of inbreeding (incestuous matings) becomes an additional selective force favoring MHC-based disassortative mating. Figure 1 outlines the interactions of these three selective forces that we propose enhance the diversity of MHC genes. Note that only two of the forces, pathogen-driven selection and reproductive (sexual) selection, operate directly on MHC genes. Selection operating through inbreeding depression acts indirectly by favoring MHC-based disassortative mating and, indeed, pathogen-driven selection can also act in this way. The relative importance of these three selective forces in a particular species depends on many variables, such as the strength of the inbreeding load, the strength of the pathogen load (caused by increased susceptibility to infectious disease in animals that are homozygous for MHC loci relative to those that are heterozygous) and the likelihood of inbred matings.

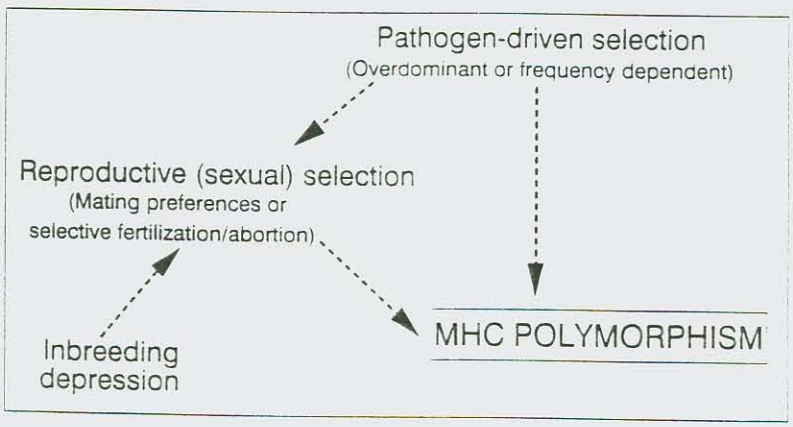


FIG 1

Primary selective forces proposed to maintain MHC genetic diversity in *Mus*. Arrows indicate the direction of selective forces that favor MHC polymorphisms, either directly, or indirectly because they favor MHC-based disassortative mating preferences. Reproductive mechanisms such as selective fertilization or selective abortion are viable alternatives to mating preferences, but the evidence for these post-mating mechanisms is weak[5]. Furthermore, pre-mating mechanisms will generally be more efficient at preferentially producing progeny with favored genotypes, because they avoid the costs associated with mating and aborted zygotes.

Below, we evaluate each of the three selective forces.

Pathogen-driven selection favors MHC genetic diversity

The molecular details of how allelic variants of the MHC influence immune responsiveness are well established[21,22]. In theory, these immune functions of MHC molecules lead to two types of pathogen-dependent selection that increase genetic diversity: heterozygote advantage, or overdominance[3,4,16], and frequency-dependent selection[7,17]. It is proposed that MHC heterozygotes have enhanced resistance to disease because they express twice as many MHC genes as homozygotes, allowing them to bind and present a wider variety of peptide antigens. Selection is thought to favor rare MHC genotypes, since pathogens are more likely to have developed mechanisms to evade the MHC-dependent immunity encoded by common MHC genotypes. The relative importance of these two forms of selection is controversial[7,23], but even proponents of frequency-dependent selection generally accept that heterozygote advantage can also operate. Unfortunately, there is as yet no compelling empirical example of either type of selection operating on MHC genes. Both hypotheses predict associations between MHC genotype and resistance to specific infectious diseases; however, the two main examples, malaria in humans[24] and Marek's disease in chickens[25], offer little support for pathogen-dependent balancing selection, as both strongly favor single MHC haplotypes and therefore reduce diversity. Although many attempts to document MHC–disease associations have failed[26–28], a sufficient number of examples have been identified[5] to maintain the viability of both hypotheses.

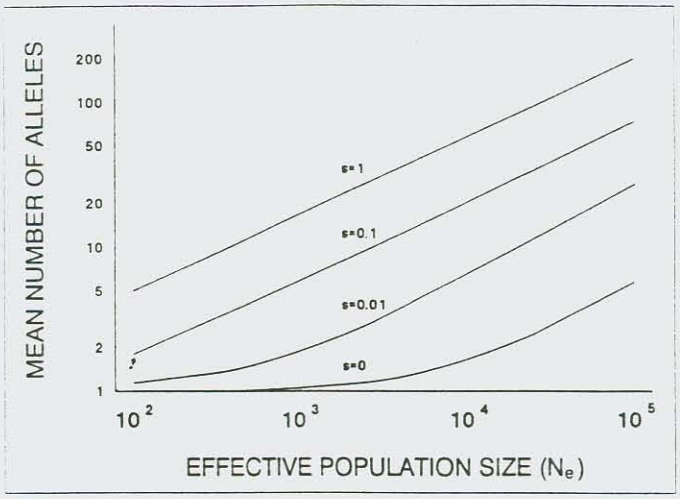


FIG 2

The mean number of alleles maintained at various levels of symmetric overdominant selection (s) and effective population sizes (N_e) at a mutation rate of 10^{-6}. Values are based on an approximation method developed by Yokoyama and Nei[48]. (Note that $s = 0.01$ means, for example, that in a population starting with 100 MHC homozygotes and 100 MHC heterozygotes, homozygotes would suffer a single pre-reproductive death while heterozygotes suffer no mortality.)

Figure 2 shows the number of alleles expected for various population sizes (N_e) and levels of overdominant selection (s), and illustrates that even a level of selection that is too weak to be detected easily can act to maintain a significant level of genetic diversity. For example, a selection coefficient of $s = 0.01$ can maintain ~–29 alleles in population sizes (N_e) of 10^4–10^5, but such a selection coefficient could only be detected by sampling thousands of individuals. If pathogen-dependent selection is this weak, it is not surprising that it has been difficult to demonstrate; this consideration, together with the theoretical persuasiveness of the MHC overdominance hypothesis, supports the idea that pathogen-dependent MHC heterozygote advantage operates in vertebrates.

Sexual preferences reduce both pathogen and inbreeding loads

The pathogen load resulting from MHC heterozygote advantage represents the reduction in fitness of MHC homozygotes caused by decreased resistance to disease. Inbreeding load represents the loss of fitness caused by homozygosity at loci that have either deleterious recessive or overdominant alleles. Progeny from inbred matings will suffer reduced fitness when such loci are present; this is commonly termed inbreeding depression. (Note that MHC-dependent pathogen load is a component of the inbreeding load.) It is generally accepted that MHC-based disassortative mating will reduce both pathogen and inbreeding loads[5,13,29,30]. This is readily illustrated in the case of MHC overdominance: if MHC homozygotes have reduced fitness because of increased susceptibility to disease, then MHC-based disassortative mating preferences will be favored because of the higher proportion of MHC-heterozygous progeny produced relative to random mating expectations.

The case for inbreeding load is similar. Since sharing highly polymorphic genetic markers is predictive of kinship[18,19], avoiding mating with animals that have a similar MHC genotype will reduce the likelihood of matings with relatives (inbreeding)[5,13,29–32]. MHC-based odors may be just one of many kin recognition cues, although it appears that the MHC has a disproportionate influence on individual odors in rodents, constituting an estimated 50% of individual differences as measured by ease of training in olfaction discrimination experiments[15]. Independent evidence that MHC genes influence kin recognition comes from recent studies of communal nesting in mice[33].

If MHC-based mating preferences act by reducing both pathogen and inbreeding loads, which of these two forces is more important? Howard has argued that MHC-based mating preferences function primarily to reduce pathogen load because while they are potentially perfect at preventing MHC homozygosity, they are imperfect at preventing inbreeding[30]. However, if the inbreeding load is sufficiently strong, avoiding inbreeding could be the more important force in maintaining MHC diversity. As discussed above, pathogen-dependent selection acting on the MHC appears to be quite weak. What is the case for inbreeding depression?

Precise measurements of fitness in inbred populations are not available for any vertebrate, but laboratory

Below, we evaluate each of the three selective forces.

Pathogen-driven selection favors MHC genetic diversity

The molecular details of how allelic variants of the MHC influence immune responsiveness are well established[21,22]. In theory, these immune functions of MHC molecules lead to two types of pathogen-dependent selection that increase genetic diversity: heterozygote advantage, or overdominance[3,4,16], and frequency-dependent selection[7,17]. It is proposed that MHC heterozygotes have enhanced resistance to disease because they express twice as many MHC genes as homozygotes, allowing them to bind and present a wider variety of peptide antigens. Selection is thought to favor rare MHC genotypes, since pathogens are more likely to have developed mechanisms to evade the MHC-dependent immunity encoded by common MHC genotypes. The relative importance of these two forms of selection is controversial[7,23], but even proponents of frequency-dependent selection generally accept that heterozygote advantage can also operate. Unfortunately, there is as yet no compelling empirical example of either type of selection operating on MHC genes. Both hypotheses predict associations between MHC genotype and resistance to specific infectious diseases; however, the two main examples, malaria in humans[24] and Marek's disease in chickens[25], offer little support for pathogen-dependent balancing selection, as both strongly favor single MHC haplotypes and therefore reduce diversity. Although many attempts to document MHC–disease associations have failed[26–28], a sufficient number of examples have been identified[5] to maintain the viability of both hypotheses.

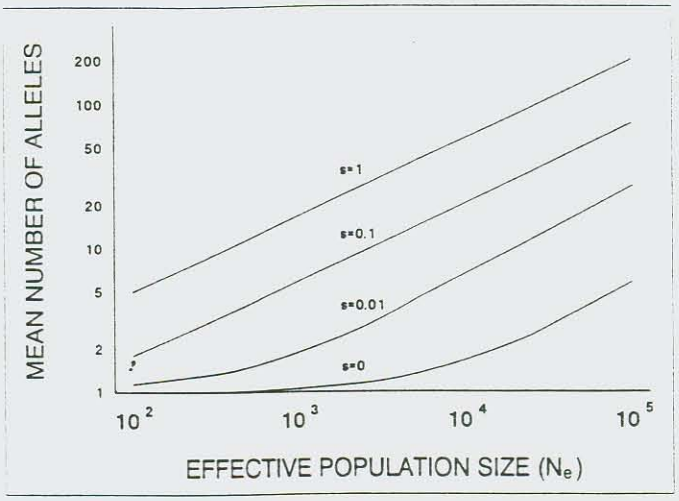


FIG 2

The mean number of alleles maintained at various levels of symmetric overdominant selection (s) and effective population sizes (N_e) at a mutation rate of 10^{-6}. Values are based on an approximation method developed by Yokoyama and Nei[48]. (Note that $s = 0.01$ means, for example, that in a population starting with 100 MHC homozygotes and 100 MHC heterozygotes, homozygotes would suffer a single pre-reproductive death while heterozygotes suffer no mortality.)

Figure 2 shows the number of alleles expected for various population sizes (N_e) and levels of overdominant selection (s), and illustrates that even a level of selection that is too weak to be detected easily can act to maintain a significant level of genetic diversity. For example, a selection coefficient of $s = 0.01$ can maintain 7–29 alleles in population sizes (N_e) of 10^4–10^5, but such a selection coefficient could only be detected by sampling thousands of individuals. If pathogen-dependent selection is this weak, it is not surprising that it has been difficult to demonstrate; this consideration, together with the theoretical persuasiveness of the MHC overdominance hypothesis, supports the idea that pathogen-dependent MHC heterozygote advantage operates in vertebrates.

Sexual preferences reduce both pathogen and inbreeding loads

The pathogen load resulting from MHC heterozygote advantage represents the reduction in fitness of MHC homozygotes caused by decreased resistance to disease. Inbreeding load represents the loss of fitness caused by homozygosity at loci that have either deleterious recessive or overdominant alleles. Progeny from inbred matings will suffer reduced fitness when such loci are present; this is commonly termed inbreeding depression. (Note that MHC-dependent pathogen load is a component of the inbreeding load.) It is generally accepted that MHC-based disassortative mating will reduce both pathogen and inbreeding loads[5,13,29,30]. This is readily illustrated in the case of MHC overdominance, if MHC homozygotes have reduced fitness because of increased susceptibility to disease, then MHC-based disassortative mating preferences will be favored because of the higher proportion of MHC-heterozygous progeny produced relative to random mating expectations.

The case for inbreeding load is similar. Since sharing highly polymorphic genetic markers is predictive of kinship[18,19], avoiding mating with animals that have a similar MHC genotype will reduce the likelihood of matings with relatives (inbreeding)[5,13,29–32]. MHC-based odors may be just one of many kin recognition cues, although it appears that the MHC has a disproportionate influence on individual odors in rodents, constituting an estimated 50% of individual differences as measured by ease of training in olfaction discrimination experiments[15]. Independent evidence that MHC genes influence kin recognition comes from recent studies of communal nesting in mice[33].

If MHC-based mating preferences act by reducing both pathogen and inbreeding loads, which of these two forces is more important? Howard has argued that MHC-based mating preferences function primarily to reduce pathogen load because while they are potentially perfect at preventing MHC homozygosity, they are imperfect at preventing inbreeding[30]. However, if the inbreeding load is sufficiently strong, avoiding inbreeding could be the more important force in maintaining MHC diversity. As discussed above, pathogen-dependent selection acting on the MHC appears to be quite weak. What is the case for inbreeding depression?

Precise measurements of fitness in inbred populations are not available for any vertebrate, but laboratory

TABLE 2. Patterns of genetic diversification of various genes

	Type of gene		
	Self-incompatibility	**MHC**	**'Average'**
Maximum no. of alleles in local populations[a,b]	37[d]	20[e]	1–3
Transpecies evolution[c]	Yes[f]	Yes[g]	No
Maximum % sequence divergence of alleles[h]	58[f]	30[h]	<1
Ratio of non-synonymous: synonymous substitution rates	High[i]	High[j]	Low

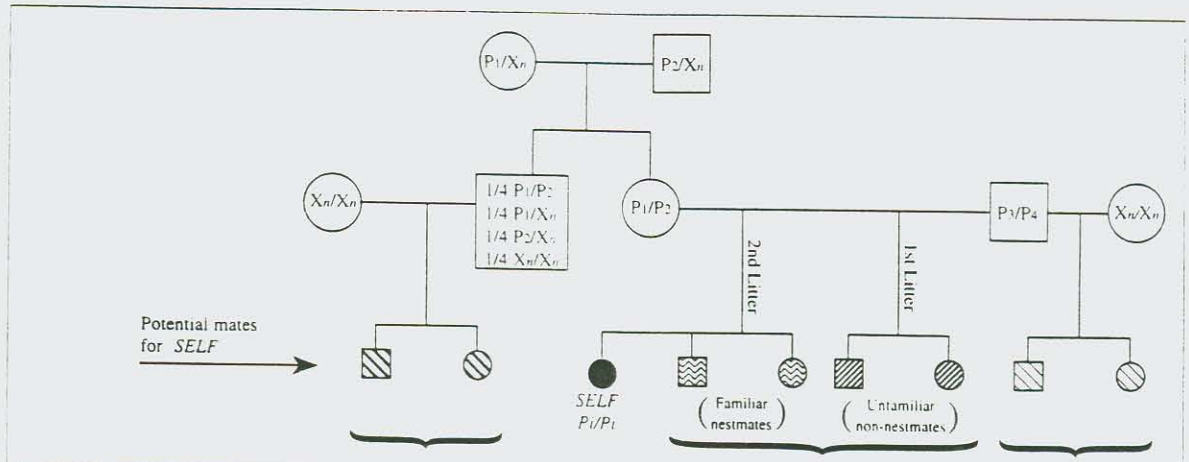
[a] Data taken from studies with approximately equal sample sizes.
[b] Figures are the maximums reported in any study.
[c] Refers to the situation where alleles in one species are more closely related to alleles in other species than to other conspecific alleles.
[d]Ref. 39. [e]Ref. 5. [f]Ref. 38. [g]Ref. 46. [h]Ref. 47. [i]Ref. 40. [j]Refs 4, 16.

genotype-specific. Allele-specific discrimination means that allele *a* can be smelled and identified in the context of both homozygosity (*a a*) and heterozygosity, say with allele *b* (*a b*), while genotype-specific discrimination means that the odor of *a b* heterozygotes is unrelated to that of *a a* and *b b* homozygotes. In Fig. 3, allele-specific discrimination is assumed; a genotype-specific system would be less efficient at discriminating kin from non-kin.

Are MHC-based mating preferences a general trait of vertebrates?

The structure of *Mus* society makes incestuous matings likely. The low levels of migration between *Mus* demes[42] mean that close relatives stay in close proximity, creating a population structure in which a genetically based kin-recognition system would be very useful in avoiding inbreeding. Many other vertebrates have dispersal patterns that make both inbreeding and MHC homozygosity less likely[43], reducing the importance of a genetically based system for avoiding inbreeding. Ascertaining the general importance of MHC-based mating preferences will require further studies of the MHC in a variety of other vertebrates.

Efforts to find MHC-biased patterns of marriage in modern human societies have failed[28] but, because of the high levels of genetic diversity within the MHC, such studies lack sufficient statistical power to allow



Relatedness to *SELF* (r)	Cousins (0.125)		Full Sibs (0.50)	Half Sibs (0.25)
Genotypic frequencies	50% P_i / X_n 50% X_n / X_n		100% P_i / P_i	100% P_i / X_n
Efficiency of inbreeding avoidance if *SELF* avoids matings with P_i / P_i or P_i / X_n	50%		100%	100%

FIG 3

The efficiency of avoiding matings with close kin (r = relatedness) using a single polymorphic locus. P_i and X_n are alleles at a polymorphic locus where P_i (i = 1–4) are the alleles carried by the parents of *SELF* and X_n are other alleles.

even reasonably strong mating patterns to be detected. Modern humans are an unlikely species for odor-based courtship, since we have relatively feeble powers of olfaction, modify our odors with perfume and often bath daily. We also have other systems for avoiding inbreeding; these include both cultural mores (incest taboos)[35] and seemingly innate systems, where children reared together (regardless of their genetic relationship) have an aversion to romance as adults[43,44]. On the other hand, the billion-dollar perfume industry suggests that humans are not indifferent to odors during courtship. Even if modern humans mate randomly as regards their MHC genotype, most human MHC alleles diverged within populations of our primate ancestors[45], so humans may still harbor MHC alleles originally selected through odor-based courtship. If so, this may have immunological consequences; binding sites that are suboptimal for conferring resistance to either infectious or autoimmune disease may have been selected originally for their role during courtship. This hypothesis is consistent with the observation that the frequency of many autoimmune diseases would be dramatically reduced if MHC alleles that predispose to these ailments were eliminated from the population[28].

Conclusions

The ideas and data we present suggest that all of the selective forces proposed (Fig. 1) do indeed operate on MHC genes in *Mus*. In short, both pathogen-driven selection and MHC-based disassortative mating preferences favor MHC genetic diversity. MHC-based mating preferences are favored because they both decrease the incidence of inbreeding and increase resistance to infectious disease, by increasing the proportion of MHC heterozygotes produced. Consequently, all three selective mechanisms contribute to MHC genetic diversity. Indirect evidence suggests that the avoidance of inbreeding may be the more important force in the evolution of MHC-based mating preferences in *Mus*, but determining whether this applies generally among other vertebrates will require further work.

Acknowledgements

We thank S. Edwards, C.J. Manning, and W.T. McCormack for editorial assistance and T.L. Colson for assistance with figures. This work was supported by grants from the NIH and NSF.

References

1 Babbitt, B.P., Allen, P.M. and Matsueda, G. (1985) *Nature* 317, 359–361
2 Bjorkman, P.J. *et al.* (1987) *Nature* 329, 512–518
3 Doherty, P.C. and Zinkernagel, R.M. (1975) *Nature* 256, 50–52
4 Hughes, A.L. and Nei, M. (1988) *Nature* 335, 167–170
5 Potts, W.K. and Wakeland, E.K. (1990) *Trends Ecol. Evol.* 5, 181–187
6 Takahata, N. and Nei, M. (1990) *Genetics* 124, 967–978
7 Slade, R.W. and McCallum, H.I. (1992) *Genetics* 132, 861–862
8 Yamazaki, K. *et al.* (1976) *J. Exp. Med.* 144, 1324–1335
9 Yamazaki, K. *et al.* (1978) *Immunogenetics* 6, 253–259
10 Yamazaki, K. *et al.* (1988) *Science* 240, 1331–1332
11 Egid, K. and Brown, J.L. (1989) *Anim. Behav.* 38, 548–549
12 Potts, W.K., Manning, C.J. and Wakeland, E.K. (1991) *Nature* 352, 619–621
13 Hedrick, P.W. (1992) *Genetics* 132, 575–581
14 Singh, P.B., Brown, R.E. and Roser, B. (1987) *Nature* 327, 161–164
15 Boyse, E.A., Beauchamp, G.K., Bard, J. and Yamazaki, K. (1990) in *Psychoneuroimmunology* (Ader, R., Felten, D.L. and Cohen, N., eds), pp. 831–845, Academic Press
16 Hughes, A.L. and Masatoshi, N. (1989) *Proc. Natl Acad. Sci. USA* 86, 958–962
17 Bodmer, W.F. (1972) *Nature* 237, 139–145
18 Getz, W.M. (1981) *J. Theor. Biol.* 92, 209–226
19 Lacy, R.C. and Sherman, P.W. (1983) *Am. Nat.* 121, 489–512
20 Grafen, A. (1990) *Anim. Behav.* 39, 42–54
21 Rotzschke, O. and Falk, K. (1991) *Immunol. Today* 12, 447–455
22 Matsumura, M., Fremont, D.H., Peterson, P.A. and Wilson, I.A. (1992) *Science* 257, 927–934
23 Hughes, A.L. and Nei, M. (1992) *Genetics* 132, 863–864
24 Hill, A.V.S. *et al.* (1991) *Nature* 352, 595–600
25 Pazderka, F. *et al.* (1975) *Immunogenetics* 2, 93–100
26 Klein, J. (1986) *Natural History of the Major Histocompatibility Complex*, Wiley
27 Klein, J. (1987) *Hum. Immunol.* 19, 155–162
28 Tiwari, J.L. and Terasaki, P.I. (1985) *HLA and Disease Associations*, Springer-Verlag
29 Partridge, L. (1988) *Philos. Trans. R. Soc. London, Ser. B* 319, 525–539
30 Howard, J.C. (1991) *Nature* 352, 565–566
31 Brown, J.L. (1983) in *Ethical Questions in Brain and Behavior* (Pfaff, D.W., ed.), pp. 111–124, Springer-Verlag
32 Uyenoyama, M.K. (1988) in *The Evolution of Sex* (Michod, R.E. and Levin, B.R., eds), pp. 212–232, Sinauer Associates
33 Manning, C.J., Wakeland, E.K. and Potts, W.K. (1992) *Nature* 360, 581–583
34 Green, E. (1966) in *Biology of the Laboratory Mouse* (Green, E., ed.), pp. 11–22, McGraw-Hill
35 May, R.M. (1979) *Nature* 279, 192–194
36 Charlesworth, D. and Charlesworth, B. (1987) *Annu. Rev. Ecol. Syst.* 18, 237–268
37 Allison, A.C. (1964) *Cold Spring Harbor Symp. Quant. Biol.* 29, 137–149
38 Haring, V. *et al.* (1990) *Science* 250, 937–941
39 Wright, S. (1939) *Genetics* 24, 538–552
40 Clark, A.G. and Kao, T-H. (1991) *Proc. Natl Acad. Sci. USA* 88, 9823–9827
41 Karlin, S. and Feldman, M.W. (1968) *Genetics* 59, 117–136
42 Lidicker, W.Z. and Patton, J.L. (1987) in *Mammalian Dispersal Patterns* (Chepko-Sade, B.D. and Halpin, Z.T., eds), pp. 144–161, University of Chicago Press
43 Blouin, S.F. and Blouin, M. (1988) *Trends Ecol. Evol.* 3, 230–233
44 Spiro, M.E. (1958) *Children of the Kibbutz*, Harvard University Press
45 Gyllensten, U.B. and Erlich, H.A. (1989) *Proc. Natl Acad. Sci. USA* 86, 9986–9990
46 McConnell, T.J., Talbot, W.S., McIndoe, R.A. and Wakeland, E.K. (1988) *Nature* 332, 651–654
47 She, J.X. *et al.* (1991) *Proc. Natl Acad. Sci. USA* 88, 453–457
48 Yokoyama, S. and Nei, M. (1979) *Genetics* 91, 609–626

W.K. Potts and E.K. Wakeland are in the Center for Mammalian Genetics and the Department of Pathology and Laboratory Medicine, University of Florida, Gainesville, FL 32610-0275, USA.