8. Mansinha, L. & Smylie, D. E. *Bull. seismol. Soc. Am.* **61,** 1433–1440 (1971).
9. National Earthquake Information Service *Preliminary Determination of Epicenters, December 1989* (Golden, Colorado, 1990).
10. Ebel, J. E., Somerville, P. G. & McIver, J. D. *J. geophys. Res.* **91,** 8231–8247 (1986).
11. North, R. G. *et al. Seismol. Res. Lett.* **60,** 89–93 (1989).
12. Basham, P. W. *Ann. N. Y. Acad. Sci.* **558,** 1–10 (1989).
13. Amick, D. & Gelenas, R. *Science* **251,** 655–658 (1991).
14. Obermeier, S. F. *et al. Science* **251,** 1061–1063 (1991).
15. Doig, R. *Can. J. Earth Sci.* **23,** 930–937 (1986).

ACKNOWLEDGEMENTS. We thank A. C. Johnston, T. Bullard, A. Crone and our colleagues for their help.

Mating patterns in seminatural populations of mice influenced by MHC genotype

Wayne K. Potts*, C. Jo Manning*† & Edward K. Wakeland*

* Laboratory of Molecular Genetics, Department of Pathology and Laboratory Medicine, University of Florida, Gainesville, Florida 32610, USA
† Department of Psychology, University of Washington, Seattle, Washington 98195, USA

BECAUSE of the central role of major histocompatibility complex (MHC) genes in immune recognition[1–3], it is often assumed that parasite-driven selection maintains the unprecendented genetic diversity of these genes[4–7]. But associations between MHC genotype and specific infectious diseases have been difficult to identify[8,9] with a few exceptions such as Marek's disease[10] and malaria[11]. Alternatively, MHC-related reproductive mechanisms such as selective abortion[12–15] and mating preferences[16,17] could be responsible for the diversity. To determine both the nature and strength of selection operating on MHC genes by we have studied components of selection in seminatural populations of mice (*Mus musculus domesticus*). Here we assess MHC-related patterns of reproduction and early (preweaning) mortality by analysing 1,139 progeny born in nine populations, and 662 progeny from laboratory matings. Reproductive mechanisms, primarily mating preferences, result in 27% fewer MHC-homozygous offspring than expected from random mating. MHC genotype had no detectable influence on neonatal (preweaning) mortality. These mating preferences are strong enough to account for most of the MHC genetic diversity found in natural populations of *Mus*.

We[6] and others[18,19] have argued that the artificiality of using inbred strains under laboratory conditions makes it impossible to use such methodologies to assess accurately the evolutionary forces operating on MHC genes in the complexity of natural populations. Our approach to solving these problems was to study populations of outbred *Mus* in enclosures of sufficient size and complexity to allow normal patterns of social competition and reprodutive behaviour under normal pathogen loads. Mice were recently derived from wild-caught individuals, but carried MHC haplotypes (b, d, k and q) from four inbred strains (see Table 1 for breeding details). Founding populations consisted of 24–31 age-matched and MHC-genotyped individuals. Unique ear punch patterns allowed individual identification (with binoculars) permitting documentation of social status, territorial boundaries, home ranges, mating pairs and other social behaviours. Pups were trapped near weaning age (15–21 days), earmarked, and 2-cm tail biopsies taken for DNA samples. After about 150 progeny were born, each population was disbanded and both preweaning stage offspring and embryos from pregnant females were MHC genotyped.

Table 1 summarizes the observed and expected frequencies for MHC haplotypes, heterozygotes and homozygotes. Although there were some departures from expected haplotypic frequencies, the mean change for the b, d, k and q haplotypes was small (0.027, 0.033, −0.026 and −0.036, respectively), indicating that no haplotype showed systematic gains or losses over the nine populations.

The only strong and consistent departure from expected frequencies was the deficiency of MHC homozygous progeny shown by all populations, with a mean deficiency of 27% (Table 1). The combined χ^2 for these homozygote deficiencies was highly significant ($P<0.0001$). As a large proportion of the 1,139 progeny were genotyped near the age of weaning, the deficiency of homozygotes could have resulted from differential neonatal mortality. This hypothesis was excluded because the proportion of homozygotes did not decline between the embryo and nestling stages (Table 2). These results indicate that the observed deficiency of homozygotes must be due to mechanisms operating before birth.

To test for MHC-related selective fertilization or abortion, we conducted informative laboratory matings. Although there was a 6% deficiency of MHC homozygotes relative to expected mendelian ratios, this difference was not statistically significant ($P=0.16$; Table 2). Further, if the genotypic data in Table 1 are adjusted for this potential 6% contribution, the remaining deficiency of MHC homozygotes (mean = 21%) is still significant ($P<0.001$).

The only remaining explanation for the homozygote deficiencies is MHC-based nonrandom mating. To determine if these mating preferences operated through nonrandom settlement patterns associated with MHC genotypes, females were assigned to the male in whose territory her sleeping/nest box was located. When these settlement pairs were used to recalculate the expected homozygosities in Table 1 (originally calculated assuming random mating between all females and all territorial males), the mean deficiency of homozygotes decreased from 27% to 19%, suggesting that settlement patterns do account for a portion of the homozygote deficiency.

The largest portion of the homozygote deficiency is still unaccounted for and must be due to extraterritorial matings, which

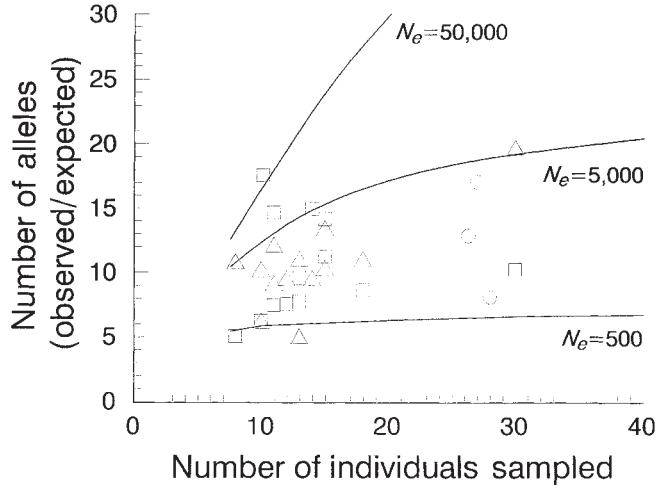

FIG. 1 The number of observed MHC alleles from local *Mus* populations (symbols) plotted with the expected number of alleles (lines) for three effective population sizes (N_e), under levels of disassortative mating that would yield a 27% deficiency of homozygotes. The expected values (corrected for sample size) were derived from an analytical model for symmetric overdominant selection[30], using a selection coefficient of $s=0.27$ and a mutation rate of 10^{-6} (an empirically derived estimate for MHC loci[31]). As N_e for *Mus* is uncertain, the expected number of alleles across the entire estimated range of N_e values (500–50,000) is provided[6]. Data for class I K (□) and D (△) loci are from serological analysis of 13 European populations[32]. The total number of alleles was estimated by assuming that the allelic frequency distributions of blanks and identified alleles were the same. Data for the class II A_β locus (○) are from a restriction fragment length polymorphism study of three North American populations[6].

TABLE 1 Observed and expected frequencies in progeny of MHC haplotypes, heterozygotes and homozygotes

Population	Number of founders ♂	♀	Haplotypic frequency of offspring b obs.	b exp.	d obs.	d exp.	k obs.	k exp.	q obs.	q exp.	MHC heterozygotes obs.	MHC heterozygotes exp.	MHC homozygotes obs.	MHC homozygotes exp.	Homozygote deficiency (%)	P
A	8	16	0.34	0.31	0.41	0.31	0.12	0.30	0.13	0.09	93	87.4	25	30.6	18	0.24
B	8	16	0.26	0.27	0.10	0.06	0.28	0.36	0.36	0.32	88	82.9	24	29.1	18	0.27
C	8	16	0.26	0.15	0.44	0.39	0.04	0.10	0.26	0.37	102	96.7	39	44.3	12	0.39
D	15	15	0.32	0.23	0.46	0.44	0.12	0.08	0.11	0.26	89	82.8	30	36.2	17	0.22
E	7	16	0.29	0.32	0.12	0.12	0.24	0.20	0.34	0.37	121	107.1	24	37.9	37	0.009
F	12	19	0.15	0.10	0.54	0.60	0.04	0.04	0.28	0.26	104	83.1	40	60.9	34	0.0004
G	8	16	0.23	0.21	0.32	0.29	0.13	0.21	0.32	0.29	131	121.6	26	35.4	27	0.07
H	8	16	0.29	0.27	0.31	0.26	0.20	0.18	0.20	0.29	99	90.1	20	28.9	31	0.06
I	8	16	0.07	0.12	0.34	0.26	0.26	0.22	0.34	0.41	72	58.1	12	21.6	44	0.006

Expected genotypic frequencies were calulated as $x_i y_i + x_j y_j$ where x and y are the frequencies in females and territorial males, respectively, of the ith and jth haplotypes. Nonterritorial males were excluded from the calculation of all expected values because this provided a better predictor of allelic frequencies of progeny which is consistent with our behavioural and genetic data indicating nonterritorial males do not breed. Excluding nonterritorial males also controlled for the only differential reproductive success found to be correlated with MHC, the ability of males to gain territories (manuscript in preparation). Regression analysis: when departures from expected haplotypic frequencies were regressed on expected haplotypic frequencies, a significant inverse correlation was found (slope = −0.94, $P < 0.025$; $r^2 - 0.11$, $P < 0.01$), indicating that the frequencies of common haplotypes are decreasing whereas rare haplotypes are increasing. This pattern is predicted by many forms of diversity-maintaining selection. Methods. **Mice**. Animals used in these experiments came from generations three and six of original crosses between wild-caught animals and four inbred strains (C57BL/10J, BALB/c, B10.BR and DBA/1, carrying MHC haplotypes b, d, k and q, respectively). In the F2-generation, only mice homozygous for one of the four inbred-derived MHC haplotypes were used to continue this outbred colony. In the resulting outbred animals, half of the genome is wild-derived and the other half is derived from one or more inbred strains. The number of inbred strains constituting the inbred-derived half of the genome was manipulated in conjunction with experiments designed to test the relative importance of inbreeding at the remainder of the genome (manuscript in preparation). In populations A–C, MHC homozygotes were derived from a single inbred strain, whereas MHC heterozygotes were derived from two inbred strains. This established a systematic correlation between inbreeding at MHC loci and genome-wide inbreeding, a condition that would be expected in natural populations that experience some inbreeding[26], as is the case for *Mus*. In our breeding design this correlation involved 1/4 of the genome and only loci derived from inbred strains. In populations G–I this correlation was eliminated by crosses that systematically introduced uniform contributions of all four inbred strains in each individual. Populations D–F were founded by progeny born in the enclosure to the founders of populations A–C, respectively. Consequently, populations D–F differed from the other six populations in two important respects. First, they were born and reared in a natural social system, in contrast to being cage-reared. Second, some of the male–female combinations were sibs, a condition that was avoided in populations A–C and G–I. To control for the possibility that behavioural mechanisms exist to inhibit matings between sibs, the random mating expectations for populations D–F were calculated with potential sib matings excluded. **MHC genotyping**. Restriction fragment length polymorphisms from *Taq*1-digested genomic DNA (extracted from 2 cm tail biopsies) allowed identification of all 10 MHC genotypes. Southern blots were hybridized with a 5.8-kilobase *Eco*R1 fragment cloned from a *Mus* MHC class II A_β gene. Protocols for DNA extraction, Southern blotting and hybridization are detailed elsewhere[27]. **Enclosures**. The outdoor (but predator free) enclosures (4.9 m × 9.8 m) were subdivided with 46-cm-high screening (0.5-cm grids) into eight equal subsections, each containing an additional 3 m spiral of screening. The screening created spatial complexity thought to be important for normal behaviour[28]. Mice could easily climb the screening, but it was often used as a territorial boundary. Observation periods of 1–2 h at dusk and daytime nest checks were made 5–7 times per week. **Pathogen load**. A variety of factors suggest that the enclosure populations were subjected to relatively normal pathogen loads. We frequently introduce wild-caught mice to our colony, providing a continuous supply of endemic mouse pathogens. The enclosure is surrounded by pasture land allowing normal vectoring of pathogens found in commensal populations of mice. When 55 mice from population C were tested for the presence of antibodies to the endemic pathogen *Mycoplasma pulmonis*, all individuals tested positive.

seem to be controlled by females. We observed 41 matings. All occurred in the territory of the mating male and 13 (32%) of these were extrapair matings, where the female had travelled to a nearby territory. Males were never observed following oestrous females beyond their own territorial boundaries. Paternity analysis of litters where maternity was unambiguous (58 litters, 305 pups), revealed that 52% of litters involved extraterritorial matings (a minimum estimate as genetic exclusions were based solely on the four MHC haplotypes). In the litters involving extraterritorial sires, there were 41% fewer MHC homozygous offspring than expected if the female had mated solely with her territorial male ($P = 0.001$; Table 3), suggesting that females seek extraterritorial matings with males that are relatively more MHC-disparate than their own territorial male.

The precise nature of MHC-based mating preferences awaits further experimentation, but it is generally assumed that they are disassortative[16]. Some forms of disassortative mating have population dynamics similar to overdominance[20], which allows the use of well developed overdominant models to estimate the importance of these mating preferences in maintaining MHC genetic diversity in *Mus*. Figure 1 compares the number of MHC alleles observed in natural populations with the number expected from mating preferences of the magnitude observed in this study. Such preferences can account for most of the MHC polymorphism in natural populations[6,7].

The above results suggest that MHC-based mating preferences are a major force responsible for the maintenance of MHC

TABLE 2 Observed and expected frequencies of MHC heterozygotes and homozygotes for nestlings and embryos (enclosure populations) and from laboratory matings

Enclosure populations	Heterozygotes obs.	Heterozygotes exp.	Homozygotes obs.	Homozygotes exp.	Homozygote deficiency (%)
Nestlings	682	619	186	247	25
Embryos	217	194	54	77	30
Laboratory matings					
♀ ♂					
a/a × a/b	90	85	80	85	6
a/b × a/a	92	81.5	71	81.5	13
a/b × a/b	125	122.5	120	122.5	2
a/b × b/c	63	63	21	21	0
Totals/means	370	352	292	310	6

For the two laboratory mating classes with expected homozygote proportions of 1/4 (a/b × a/c) and 1/2 (a/b × a/b and a/a × a/b), no significant differences in fecundity were found either in mean litter size (6.3 and 6.4, respectively; $P > 0.4$) or in days to first litter (25 and 28, respectively; $P > 0.2$). These results indicate that fecundity was not correlated with the degree of MHC disparity between mates.

TABLE 3 Extraterritorial matings result in more MHC heterozygous offspring than expected if the female had mated with her territorial male

Type of mating	Heterozygotes obs.	Heterozygotes exp.	Homozygotes obs.	Homozygotes exp.	Homozygote deficiency (%)	P*
Within territory matings	114	113.5	28	28.5	2	0.92
Extra-territorial matings	137	118.6	26	44.4	41	0.001

Expected values were calculated on the basis of females mating with the male in whose territory her nest/sleeping box was located. These data are derived from the subset of litters where maternity could be unambiguously assigned. Determination of maternity through observation was often impossible owing to the communal nesting and nursing behaviour of *Mus*[29].

* χ^2 test.

genetic diversity in *Mus*. The laboratory demonstration that mice[21], rats[22] and humans[23] can, through olfaction, discriminate between MHC congenic strains of rodents, provides a mechanistic basis for these MHC-based mating preferences. Two alternative, but not mutually exclusive, functions have been hypothesized for the evolution of MHC-based mating preferences. Progeny from these matings may have increased fitness because either their MHC genotype confers increased disease resistance (for example through heterozygote advantage)[6,16], or MHC serves as a polymorphic marker in a genetic incompatibility system that functions to minimize genome-wide inbreeding[6,18,24,25]. □

Received 22 April; accepted 20 June 1991.

1. Babbit, B. P., Allen, P. M. & Matsueda, G. *Nature* **317,** 359–361 (1985).
2. Roy, S., Scherer, M. T., Briner, T. J., Smith, J. A. & Gefter, M. L. *Science* **244,** 572–575 (1989).
3. Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49,** 273–280 (1987).
4. Doherty, P. C. & Zinkernagel, R. M. *Nature* **256,** 50–52 (1975).
5. Hughes, A. L. & Nei, M. *Nature* **335,** 167–170 (1988).
6. Potts, W. K. & Wakeland, E. K. *Trends Ecol. Evol.* **5,** 181–187 (1990).
7. Takahata, N. & Nei, M. *Genetics* **124,** 967–978 (1990).
8. Tiwari, J. L. & Terasaki, P. I. *HLA and Disease Associations* (Springer, New York, 1985).
9. Klein, J. *Natural History of the Major Histocompatibility Complex* (Wiley, New York, 1986).
10. Pazderka, F., Longenecker, B. M., Law, G. R. J., Stone, H. A. & Ruth, R. F. *Immunogenetics* **2,** 93–100 (1975).
11. Hill, A. V. S. *et al. Nature* (this issue).
12. Clarke, B. & Kirby, D. R. S. *Nature* **211,** 999–1000 (1966).
13. Palm, J. *Cancer Res.* **34,** 2061–2065 (1974).
14. Hamilton, M. & Hellstron, I. *J. Reprod. Immun.* **19,** 267–270 (1978).
15. Thomas, M. L., Harger, J. H., Wagner, D. K., Rabin, B. S. & Gill, T. J. *Am. J. Obstet. Gynec.* **151,** 1053–1058 (1985).
16. Yamazaki, K. *et al. J. exp. Med.* **144,** 1324–1335 (1976).
17. Egid, K. & Brown, J. L. *Anim. Behav.* **38,** 548–549 (1989).
18. Partridge, L. *Phil. Trans. R. Soc. B* **319,** 525–539 (1988).
19. Nei, M. & Hughes, A. L. in *Evolution at the Molecular Level* (eds Selander, R. K., Clark, A. G. & Whittam, T. S.) 222–247 (Sinauer, Sunderland, Massachusetts, 1991).
20. Karlin, S. & Feldman, M. W. *Genetics* **59,** 117–136 (1968).
21. Yamazaki, K. *et al. J. exp. Med.* **150,** 755–760 (1979).
22. Singh, P. B., Brown, R. E. & Roser, B. *Nature* **327,** 161–164 (1987).
23. Gilbert, A. N., Yamazaki, K. & Beauchamp, G. K. *J. comp. Pathol.* **100,** 262–265 (1986).
24. Brown, J. L. in *Ethical Questions in Brain and Behaviour* (ed. Pfaff, D. W.) 111–124 (Springer, New York, 1983).
25. Uyenoyama, M. K. in *The Evolution of Sex* (eds Michod, R. E. & Levin, B. R.) 212–232 (Sinauer, Sunderland, Massachusetts, 1988).
26. Weir, B. S. & Cockerham, C. C. *Genet. Res.* **21,** 247–262 (1973).
27. McConnell, T. J., Talbot, W. S., McIndoe, R. A. & Wakeland, E. K. *Nature* **332,** 651–654 (1988).
28. Bronson, F. H. *Q. Rev. Biol.* **54,** 265–299 (1979).
29. Wilkinson, G. S. & Baker, A. E. M. *Ethology* **77,** 103–114 (1988).
30. Yokoyama, S. & Nei, M. *Genetics* **91,** 609–626 (1979).
31. Melvold, R. W. & Kohn, H. I. in *Transgenic Mice and Mutants in MHC Research* (eds Egorov, I. K. & David, C. S.) (Springer, New York, 1990).
32. Nadeau, J. H., Wakeland, E. K., Gotze, D. & Klein, J. *Genet Res.* **37,** 17–31 (1981).

ACKNOWLEDGEMENTS. We thank M. Brown for *Mycoplasma* screening, M. Uyenoyama for comments on the manuscript, and technicians and students who contributed to the collection of behavioural and molecular data. The work was supported by the NIH.

Development and function of T cells in mice rendered interleukin-2 deficient by gene targeting

Hubert Schorle, Thomas Holtschke, Thomas Hünig, Anneliese Schimpl & Ivan Horak*

Institute of Virology and Immunobiology, University of Würzburg, Versbacherstrasse 7, 8700 Würzburg, Germany

INTERLEUKIN-2 (IL-2) is a lymphocytotropic hormone which is thought to have a key role in the immune response of mammalian cells. It is produced by a subpopulation of activated T-lymphocytes and acts *in vitro* as the principal auto- and paracrine T-cell growth factor (for reviews see refs 1–3). IL-2 is, however, not the sole T-cell growth factor[4,5], nor does it act exclusively on T cells, also promoting growth of NK cells[6] and differentiation of B cells[7]. A role for IL-2 in T-cell development has been postulated but remains controversial[8–12]. Here we test the requirement for IL-2 *in vivo* using IL-2-deficient mice generated by targeted recombination. We find that mice homozygous for the IL-2 gene mutation are normal with regard to thymocyte and peripheral T-cell subset composition, but that a dysregulation of the immune system is manifested by reduced polyclonal *in vitro* T-cell responses and by dramatic changes in the isotype levels of serum immunoglobulins.

To inactivate the endogenous IL-2 gene by targeted recombination[13–16], a neomycin-resistance (*neo*^{r}) gene was inserted in reverse orientation into the third exon of the IL-2 genomic clone[17] (Fig. 1). This insertion introduces several stop codons in all reading frames which abolish the biological activity of the IL-2 protein completely[18]. The mutated fragment was introduced into embryonic stems cells (ES cells)[19] and *neo*^{r} colonies were selected. Clones with the correct homologous recombination event at the IL-2 locus were identified by polymerase chain reaction (PCR) and Southern blot analysis. The frequency of the correctly targeted allele was one in 200.

Two ES cell clones carrying the disrupted IL-2 gene on one allele were injected into blastocysts of C57BL/6 mice and transferred to CD-1 foster mothers. Eight out of 40 animals born were chimaeras. To test for germ-line transmission of the IL-2 mutation, the chimaeras were mated with C57BL/6 mice. Three out of five chimaeric mice derived from the same clone transmitted the ES cell genome to their progeny. Mice heterozygous for the disrupted IL-2 gene were mated and their progeny analysed to detect animals homozygous for the mutation (IL-2$^{-/-}$) (Fig. 1*d*). As expected, the IL-2 mutation was transmitted in a mendelian fashion.

Intrathymic differentiation of T cells expressing the T-cell antigen receptor (TCR) proceeds from a TCR-negative, CD4$^-$8$^-$ through a TCR-low, CD4$^+$8$^+$ stage towards the mature TCR-high, CD4$^+$8$^-$ and CD4$^-$8$^+$ phenotypes that seed the peripheral lymphoid organs (reviewed in ref. 20). Both positive selection (for self MHC-restriction) and negative selection (for self tolerance) occur during passage through the CD4$^+$8$^+$ compartment[21]. To compare the representation of these thymocyte subsets in normal and in IL-2-deficient animals, a litter of seven mice derived from IL-2$^{(+/-)}$ parents was analysed by two-colour flow cytometry at 4 weeks of age (Fig. 2). It is apparent that disruption of the IL-2 gene had no major effect on thymocyte subset composition. Both immature and mature subsets were present at normal levels in animals homozygous for the inactivated gene. Also, the small contribution of γ/δ T cells was unaffected (not shown). Finally, the average total cell number per thymus was similar in normal and mutant mice (2.20×10^8 ($n = 7$) versus 2.18×10^8 ($n = 5$); data from two litters). These findings rule out an essential role for the IL-2/IL-2R-pathway in thymopoiesis and are in full agreement with a reported case of human IL-2 deficiency[22].

Cell-surface marker analyses of spleen (not shown) and lymph node cells (Fig. 2) also revealed no striking differences between normal and IL-2-deficient mice. These observations clearly indicate that IL-2 is neither essential for thymocyte maturation nor for seeding of the periphery with mature T cells. IL-2 deficiency was, however, manifested at the functional level. Cells from thymus, lymph node and spleen of IL-2-deficient mice responded less well than those from normal littermates to polyclonal T-cell activators (concanavalin A, Table 1; and anti-CD3, data not shown) unless IL-2 was added. Because, as expected, no IL-2 activity was detected in supernatants of concanavalin A-activated spleen and lymph node cells from IL-2-deficient mice the residual responses after polyclonal activation suggest the use of alternative growth factor pathways.

* To whom correspondence should be addressed.

less its present high rate. Not surprisingly, their model of core–mantle interaction shows some significant differences from Larson's, which directly relates heat loss from the lower mantle to simultaneous stabilization of the core. From their evidence, slow TPW — that is, reduced convective activity — corresponds to a decreasing reversal rate. Conversely, fast TPW is accompanied by an increasing reversal rate, so, as in Larson's model, increased mantle activity is somehow related to core stabilization. But as the figure shows, some substantial time-lags have to be built into Courtillot and Besse's model, with events in the mantle significantly preceding changes in the core.

There is no doubt that the issues under discussion are of supreme importance to the earth sciences. If the coupling of core, mantle and surface events (including climatic) can be firmly established, a substantial unifying framework will be available for the study of most major geological processes, though events of extraterrestrial origin may yet muddy the waters. The comparison of Larson's and Courtillot and Besse's papers, however, illustrates some of the obstacles to further progress.

First, at present it is permissible to postulate many different things about the physical nature of happenings at the core–mantle boundary, within the mantle itself, and within the core. The generally acceptable theoretical models, against which the observational facts ultimately have to be tested (because many of the phenomena are outside the range of direct observation), do not yet exist in sufficient detail. Superplumes demand supercomputers.

Second (within the observational field) is the problem of correlation of events in time, which is so often (and not always wisely) the basis of geological hypothesis. Larson, for example, demonstrates a superb one-to-one correlation, but unfortunately there is only one period of increased production of oceanic crust to correlate with one period of geomagnetic stability. What is the level of significance? Ultimately, if time correlation is the key to the argument, we need to define as many separate volcanic events as possible, or whatever other events are used to measure mantle convective activity. □

K. G. Cox is in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

1. Larson, R. L. *Geology* **19**, 547–550 (1991).
2. Wilson, J. T. *Can. J. Phys.* **41**, 863–870 (1963).
3. White, R. S. & McKenzie, D. P. *J. geophys. Res.* **94**, 7685–7730 (1989).
4. Caldeira, K. & Rampino, M. R. *Geophys. Res. Lett.* **18**, 987–990 (1991).
5. Courtillot V. & Besse J. *Science* **237**, 1140–1147 (1987).

IMMUNOLOGY

Disease and evolution

J. C. Howard

"Now every species of mammal and bird so far investigated has shown a quite surprising biochemical diversity revealed by serological tests. The antigens concerned seem to be proteins to which polysaccharide groups are attached. We do not know their functions in the organism, though some of them seem to be part of the structure of the cell membrane. I wish to suggest that they may play a part in disease resistance, a particular race of bacteria or virus being adapted to individuals of a certain range of biochemical constitutions, while those of other constitutions are relatively resistant" (J. B. S. Haldane, 1949; ref. 1).

THIS passage is typical Haldane — clear, original, synthetic, precocious. But was he right? Haldane was aiming to explain not the existence of membrane-bound glycoproteins as such, but rather their polymorphic diversity. By virtue of their overwhelming advantage in rate of evolution, microbes seemed always to have the upper hand against host attempts to evolve a resistance mechanism common to all members of the species. Haldane conceived of a form of equilibrium in which host resistance factors vary from individual to individual. Infectious pathogens would then tend to learn to deal with the commonest type, leaving the way clear for rare types to persist, and eventually flourish for a time.

The membrane glycoproteins coded by the major histocompatibility complex (MHC) were some of the main contributors to Haldane's "quite surprising biochemical diversity". We now know that the MHC indeed seems to function as a resistance system, apparently exactly as Haldane conceived it. Polymorphic glycoproteins encoded by the MHC play a central role in the immune system; peptide fragments of infectious origin can be captured in the peptide-binding grooves of both class I and class II MHC molecules and presented by them to the T-cell immune system. Furthermore, the astonishing polymorphism of MHC molecules largely centres upon those residues which interact with peptides[2], and the large excess of coding substitutions at exactly these positions[3] suggests, exactly as Haldane's theory predicts, that there is strong selective pressure in favour of variation.

All this is good physiology and good genetics, but does it constitute proof of Haldane's theory? Surely, what is needed is explicit evidence that MHC allele frequencies are driven by pathogens? On pages 595 and 619, *Nature* now publishes two outstanding studies on the selective pressures operating on MHC polymorphism in the wild[4,5], and they reach what at first sight seem to be opposed conclusions, one wholly for Haldane, the other wholly against. I shall try to reconcile these two studies, and will argue that they should in fact both be seen as strengthening rather than weakening the generalization that infectious disease is the principal motor for polymorphism.

First, for Haldane, Hill and colleagues[4] face the problem directly. Find a disease associated with a high mortality in the wild, a disease having a high prevalence and in a species in which it is possible to identify most histocompatibility alleles with precision. In addition, it must be possible to assess disease-related mortality at all stages in the life cycle. The species is the challenge of course. Only the human qualifies. The disease is falciparum malaria, hyperendemic in West Africa, a potent killer of the young and already established as the exemplar for natural selection through its effect in elevating the frequency of HbS, the sickle cell haemoglobin allele[6] (still, after more than 30 years, the only balanced polymorphism in man for which we have an explanation).

Hill *et al.* compared human leukocyte antigen types in children desperately ill with severe malarial anaemia or cerebral malaria against a number of control groups. This was a brave and correct analysis, for the story is about natural selection, and natural selection in this context means life-threatening illness. Out of 45 class I alleles (assayed first by serology, then confirmed in a different sample by the polymerase chain reaction), one alone, HLA-Bw53, was significantly reduced in frequency in the severely ill children. The hypothesis that HLA-Bw53 is indeed a protective allele against the severest forms of malaria is confirmed by its geographical distribution; vanishingly rare elsewhere in the world, HLA-Bw53 reaches a frequency as high as 25% in malarial regions of Africa. Out of 13 class II haplotypes assayed, again one, DRB1*1302-DQB1*0501, was significantly reduced in frequency in one category of the severely ill children, namely those with severe malarial anaemia. The class I and class II protective mechanisms are presumably different, and suggest new approaches to vaccine development.

Although the degree of protection conferred on an individual by the protective HLA alleles was less than that provided by HbS, the higher frequency

of the HLA alleles means that, overall, their protective effect must be greater, representing a total equivalent to about 15% of all severe malaria cases in the Gambia, against 12% for HbS. The population of West Africa has been dense enough for efficient malarial transmission for no more than 10,000 years; presumably HLA-Bw53 has reached its 25% frequency from close to zero in that time by simple directional selection. Its now-known selective advantage in a malarial environment is easily large enough to achieve this incidence.

The premise of Haldane's theory is that parasites always out-evolve their hosts. Will *Plasmodium falciparum* therefore soon take avoiding action? If not, should we expect HLA-Bw53 to carry on increasing in frequency at the expense of other HLA alleles? Presumably (but it now seems a fair presumption), other pathogens will take advantage of Bw53 as its frequency rises, though whether, in a modern medical context, we shall ever see the consequence is open to doubt. In any event, the results of Hill *et al.* clearly demonstrate the essential validity of Haldane's claim, and reinforce other evidence in its favour both in human infectious disease[7], and also, in one well-documented case, in chickens[8,9].

The argument for infectious disease as the driving force for MHC polymorphism now seems overwhelming. So when Potts and colleagues[5] established populations of half-wild mice carrying known histocompatibility types at known frequencies in an enclosed but naturalistic environment, they can hardly have anticipated what would happen. Essentially right away, genotype frequencies departed radically from expectation, with far too many MHC heterozygotes being found in the progeny. Were MHC homozygotes peculiarly susceptible to infectious disease, and dying young? Potts *et al.* showed that this was definitely not so, and that the whole excess of heterozygotes arose from mating choices made by the females. Not only did females initially tend to establish territorial relationships with males differing at the MHC, but they also devoted more than half of their reproductive activity to contracting extraterritorial liaisons with further MHC-incompatible males. So strong was the disassortative mating that Potts *et al.* were able to demonstrate that this force, by itself, could be responsible for maintaining the known polymorphism of MHC alleles in the wild mouse population.

The ability of mice and rats to use MHC alleles as individuality cues has been known for some time[10,11]. Different MHC alleles are associated with distinct urinary odours which are a basis for mating preference, though the pattern of preference has not been formally established. It is striking that the effect of the MHC should so predominate, in Potts *et al.*'s study, that it could not be overridden by segregation at the many hundreds of other loci by which the mice must have differed from each other.

What does this result mean? It can be reconciled with Haldane's theory by Potts *et al.*'s own suggestion that female disassortative mating preference, by ensuring MHC heterozygosity, confers enhanced fitness on the progeny through increased disease resistance. Their alternative proposal is that the MHC is serving as a polymorphic marker in a genetic incompatibility system for avoiding the generalized deleterious consequences of genome-wide inbreeding. One should note, however, that the first hypothesis is in fact a special case of the second, where the relevant locus for inbreeding avoidance is the MHC, and the deleterious consequences are decreased resistance to infectious disease. Furthermore, although a single-locus marker for genome-wide diversity can work, it is most effective at loci linked to, or identical to, the incompatibility locus itself. The simple teleology of the matter therefore suggests that the MHC is itself the genomic locus whose heterozygosity is of principal evolutionary concern.

A Spanish cock in full plumage, with a fine fat comb and wattles (from Darwin's *Variation of Animals and Plants under Domestication*). A lot of what is on display between animals indicates health status — 'feeling off colour', 'bright-eyed and bushy-tailed', 'in the pink of condition', these phrases all tell us how quickly we perceive disease and health. Infectious disease may drive sexual selection too[18].

So we return to infectious disease, but now at one remove. Not only does infectious disease apparently directly drive polymorphism in the MHC by differential mortality, but we must now conclude that it also regulates mating preference in order to maximize heterozygosity at these all-important loci for disease resistance. Heterozygotes may well flourish at the expense of homozygotes, as a direct result of the ravages of disease; however, by disassortative mating, wild mice seem to be achieving the same goal at a lower cost. Not only that, but rodents may even have hijacked the infectious micro-organisms themselves in order to provide the volatile urinary odorants that serve to identify MHC alleles[12]. There is evidence that the urine of germ-free rats of different genotype cannot be distinguished[13], but contrary results have been obtained in mice[14] and the relevance of these observations to MHC-related mating preference is yet to be established. If the pressure of infectious disease can influence olfactory preference in mating, perhaps it can affect other cues as well. It no longer seems implausible to suppose that well-developed secondary sexual characters (see figure) may primarily indicate freedom from infectious disease.

Haldane saw, of course, that the selection pressures of infectious disease would be strong enough to bring other adaptation in tow. He proposed, for example, that polymorphic resistance loci would tend to have high mutation rates. Is it an accident that the highest

1. Haldane, J. B. S. *La Ricerca Scientifica* **19** Suppl. 19, 68–75 (1949).
2. Bjorkman, P. J. *et al. Nature* **329**, 512–518 (1987).
3. Hughes, A. L. & Nei, M. *Nature* **335**, 167–170 (1988).
4. Hill, A. V. S. *et al. Nature* **352**, 595–600 (1991).
5. Potts, W. K., Manning, C. J. & Wakeland, E. K. *Nature* **352**, 619–621 (1991).
6. Allison, A. C. *Br. med. J.* **1**, 290–294 (1954).
7. de Vries, R. R., Khan, P. M., Bernini, L. F., van Loghem, E. & van Rood, J. *J. Immunogenet.* **6**, 271–287 (1979).
8. Pazderka, F., Longenecker, B. M., Law, G. R. J., Stone, H. A. & Ruth, R. F. *Immunogenetics* **2**, 93–100 (1975).
9. Briles, W. E., Stone, H. A. & Cole, R. K. *Science* **195**, 193–195 (1977).
10. Yamazaki, K. *et al. J. exp. Med.* **144**, 1324–1335 (1976).
11. Singh, P. B., Brown, R. E. & Roser, B. J. *Nature* **327**, 161–164 (1987).
12. Howard, J. C. *Nature* **266**, 406–408 (1977).
13. Singh, P. B. *et al. J. chem. Ecol.* **16**, 1667–1682 (1990).
14. Yamazaki, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 8413–8416 (1990).
15. Klein, J. *Adv. Immun.* **26**, 55 (1978).
16. Hamilton, W. D. *J. Genet.* **69**, 17–32 (1990).
17. Hamilton, W. D., Axelrod, R. & Tanese, R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3566–3573 (1990).
18. Pomiankowski, A. N. *Oxf. Surv. évol. Biol.* **5**, 136–184 (1988).

recorded mammalian mutation rate is in an allele of a class I gene from the mouse MHC (ref. 15), probably achieved by a specialized mutational mechanism? He did not, however, refer to the ultimate extension of his position, namely, that although infectious disease may be the principal cause of polymorphism, it would be virtually ineffective in the absence of segregation and, for multiple loci, independent assortment. In other words, the sexual mechanism itself is an essential ingredient of the resistance system. Which came first? Most of us have been brought up with the idea that sex causes disease. A longer term view, and one for which there is increasing support, is that the boot belongs on the other foot[16,17]. □

J. C. Howard is in the Department of Immunology, AFRC Institute of Animal Physiology and Genetics, Babraham, Cambridge CB2 4AT, UK.

POPULATION GENETICS

A way to world knowledge

Jared M. Diamond

THINK of what we could deduce if we knew the DNA base-pair sequence of every human now alive. We could reconstruct much of the history of modern human populations in terms of prehistoric migrations and founder effects. We could study the social structure of human populations, the frequencies and types of human mutations, and natural selection in humans along environmental gradients. We could understand our differing genetic susceptibility to many diseases. Although that goal of completely sequencing everyone is unfortunately not feasible, a flurry of papers shows that a much lesser enterprise could yield a great deal of information[1–3].

The immediate impetus comes, of course, from the Human Genome Project, which seeks to sequence completely what amounts to one Caucasian's genome. That endeavour does not in itself address the questions outlined above, but Cavalli-Sforza *et al.*[1] argue that with a mere one per cent of extra effort and expense it could. Two practical issues arise: how to allocate resources among populations (for example do we learn more by sequencing the genomes of 500 individuals from each of 20 populations, or those of 20 from each of 500?); and how large a fraction, and which specific fractions, of each individual's genome to sequence. If we had already sequenced everybody, we could then with hindsight design an efficient sampling strategy. The dilemma is that we don't know the results, yet we have to guess at them to design the strategy.

An essential consideration that Cavalli-Sforza *et al.* point out stems from the main pattern of human history for the past 10,000 years. By virtue of acquiring plants and animals suitable for domestication, certain groups that formerly constituted only tiny fractions of the world's human population have overrun much of the globe. Prime examples are the expansions of Anatolian farmers and proto-Indo-European speakers over Eurasia, of Austronesians over Indonesia and the Australian realm, of Bantus over sub-Saharan Africa, and of modern Europeans over Australia and the Americas. Conversely, other populations that once accounted for much of human diversity and population numbers are now close to vanishing, overrun by the expansions of others. Cavalli-Sforza *et al.* answer the first of the two practical issues by proposing a crash programme for collecting samples (permanent cell lines, DNA or both) from these vanishing populations. With that material secured, the matter of how to study it could then be addressed at leisure.

Any geneticist or anthropologist could quickly come up with a wish-list of declining human populations to sample. High on anyone's list would come the !Kung of Namibia and Botswana, remnants of the population that occupied most of southern Africa until a few millennia ago, and located at the deepest branches in the world tree of mitochondrial DNA (mtDNA)[4]; scattered tribal peoples of southeast Asia and Indonesia, probably remnants of that area's former population related to native Australians and Melanesians and swamped by the Austronesian expansion; and Japan's Ainus, remnants of who-knows-what. Other relict populations persist in regions protected by mountainous terrain such as the Pyrenees and Caucasus. Still other priorities would be areas of high genetic diversity such as New Guinea, home to one-fifth of all the world's languages and (especially in the highlands) to an unknown fraction of the world's genetic diversity.

This wish-list begs the question of whether to sample a few dozen areas in detail or else hundreds in less detail. Here, too, the best strategy depends on the incompletely known patterns of genetic diversity that we seek to discover. For example, mtDNA sequencing detected no geographical variation within the peoples of the Middle East, but abundant variation in those of the New Guinea highlands and among African pygmies[2,4,5]. Similarly, the number of people to sample from each population depends on intrapopulation variability, which surely differs not only among populations but also among the genetic loci studied.

Hence one's sampling scheme will depend on the particular problem to be tackled. For instance, at one extreme mtDNA is especially variable, the control region of mtDNA is hypervariable, and the first 400 base pairs of that region contains most of its hypervariability. Of 88 mtDNA types from the control region derived from 117 Caucasian people, only 12 were shared by two or more individuals, and 10 of those 12 shared types were shared between Sardinian individuals or between Middle Eastern individuals[2]. Again, identical mtDNA types were shared between individual !Kung or Biaka Pygmies or Mbuti Pygmies, but not between two or more of those populations[4]. At the opposite extreme, a study of nuclear DNA in Pygmies, Europeans, Chinese and Melanesians identified a polymorphism (*HP*/*Bam*HI) whose gene frequency ranged only from 0.35 (in Chinese) to 0.46 (in Zaire pygmies)[3]. A larger population would be required to study variation in *HP*/*Bam*HI than in the mtDNA control region.

By analogy, imagine trying to reconstruct the evolution of the animal kingdom if some species such as pigeons and house mice had undergone a population explosion, but if most terrestrial animals weighing over ten kilograms and most organisms inhabiting tropical rainforest had just vanished (which indeed may come to be the case before too long; the argument for sampling the genomes of such species is also compelling). The human problem is partly similar, but only partly. Although some threatened human populations may vanish because of the deaths of most members, more will vanish as their members become absorbed into other populations[6]. But attempts to sample the genomes of animals and humans have in common a need to reconcile our intellectual desire to know with our ethical desire to preserve. Above all, both share a sense of urgency, because so little time remains before it will be too late. □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024, USA.

1. Cavalli-Sforza, L. L., Wilson, A. C., Cantor, C. R., Cook-Deegan, R. & King, M. C. *Genomics* (in the press).
2. Di Rienzo, A. & Wilson, A. C. *Proc. natn. Acad. Sci. U.S.A.* **88,** 1597–1601 (1991).
3. Bowcock, A. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **88,** 839–843 (1991).
4. Vigilant, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **86,** 9350–9354 (1989).
5. Stoneking, M. *et al. Genetics* **124,** 717–733 (1990).
6. Cavalli-Sforza, L. L. *et al. Cold Spring Harbor Symp. quant. Biol.* **51,** 411–417 (1986).