

NEWS AND COMMENTARY

Evolutionary genomics

Foxy MHC selection story

P Hedrick

Heredity (2004) 93, 237–238. doi:10.1038/sj.hdy.6800539

Published online 28 July 2004

Identifying and characterizing genes that have undergone favorable selection is a major focus of current research in evolutionary genomics. Probably, the most well-studied system indicative of adaptive evolution is the major histocompatibility complex (MHC), initially known in humans from matching recipients and donors in organ transplants. The importance of MHC genes in distinguishing self from nonself is thought to have its roots in recognition and elimination of pathogens. Aguilar *et al* (2004) now have found that MHC variation in the otherwise genetically depauperate population of San Nicolas Island foxes is inexplicably high, and suggest that their data support strong balancing selection.

The human MHC genes, known as *HLA*, are the most variable in the genome and, in some regions, are about two orders of magnitude more variable than the overall genomic rate. Incredibly, human MHC genes *HLA* *-A*, *-B*, and *-DRB* have at least 243, 499, and 321 different alleles worldwide, respectively. In addition, their high variation resides primarily in amino acids known to be important in initiating the immune response, with some individual amino-acid sites having heterozygosities greater than 0.7.

Evidence of balancing selection at MHC genes has been found in research using a number of different approaches (Garrigan and Hedrick, 2003). Selection in the current generation has been

identified by measuring survival differences between heterozygotes and homozygotes, correlations of disease resistance with genotype, and deviations from Hardy–Weinberg or random mating proportions. Selection in the recent past has been determined by excess heterozygosity compared to neutral theory expectations, differences in F_{ST} compared to neutral theory, or excess linkage disequilibrium. Selection in the distant past has been documented as an excess of nonsynonymous to synonymous substitutions, by the McDonald–Kreitman or Tajima’s *D*-tests, and as trans-species polymorphism.

The island fox (*Urocyon littoralis dicheyi*) is an endemic canid inhabiting the six largest Channel Islands off the coast of southern California. The San Nicolas Island fox has been characterized as the least genetically variable, sexually reproducing animal population because examination of allozymes, minisatellites, and 18 dinucleotide microsatellite loci showed no genetic variation. The new survey of two MHC loci and three tetranucleotide microsatellite loci closely linked to these MHC genes surprisingly found substantial variation in the San Nicolas Island fox population.

This detailed examination of an important genomic region is compelling and shows that sophisticated evolutionary genetics can be carried out on endangered species. Having worked on MHC variation and selection in a

number of organisms, my predisposition is to loudly applaud these findings. However, one needs to be careful in selling an evolutionary story so that it does not become greater than the facts merit.

To provide a perspective for these data, Table 1 gives the observed and expected heterozygosities for the two MHC loci (*DRB* and *DQB*), the average for the three microsatellite loci linked to the MHC, and 18 unlinked microsatellite loci. Aguilar *et al* (2004) base much of their article on the difference between the observed heterozygosity for *DRB* (0.36) and that for the 18 unlinked microsatellite loci (0.00) on San Nicolas. Their assumption is that both the unlinked microsatellite and MHC loci would have been influenced equivalently by nonselective factors, such as genetic drift and gene flow, but that any differences between these categories of loci could only be the result of selection acting on the MHC loci.

As they show with simulations, the probability that heterozygosity will be retained at the *DRB* locus and lost at the 18 unlinked loci is miniscule if chance alone is involved. However, assuming a single-generation bottleneck of less than 10 individuals and 95% selection against MHC homozygotes, they are able to generate a scenario with the observed level of *DRB* variation and no observable variation at the 18 unlinked loci. Although selection may be substantial for MHC genes, the near lethal homozygosity level of selection assumed here is far beyond previous observations.

The level of heterozygosity at the MHC-linked microsatellite loci (averaging over 0.5 over all the populations) is surprisingly even higher than that for the MHC genes. Contrary to my intuition about the proposed selection scenario, Aguilar *et al* (2004) did not find any linkage disequilibrium between

Table 1 The observed (Obs.) and expected (Exp.) heterozygosity for two MHC loci, three microsatellite loci linked to the MHC, and 18 unlinked microsatellite loci in the Island Fox (asterisks indicate benchmarks used in their simulations)

Island	MHC				Microsatellite loci		
	<i>DRB</i>		<i>DQB</i>		<i>MHC</i> (3)		<i>Other</i> (18)
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	
San Miguel	0.00	0.44	0.00	0.00	0.42	0.35	0.11
Santa Rosa	0.16	0.46	0.00	0.00	0.51	0.68	0.21
Santa Cruz	0.14	0.28	0.21	0.40	0.58	0.68	0.22
San Nicolas	0.36*	0.30	0.00	0.00	0.51	0.47	0.00*
Santa Catalina	0.36	0.41	0.55	0.44	0.41	0.59	0.36
San Clemente	0.00	0.00	0.00	0.00	0.59	0.64	0.26
Mean	0.17	0.32	0.13	0.14	0.50	0.57	0.19

these linked microsatellite loci and *DRB*. This is perplexing because if there were recent, strong selection on *DRB* as they suggest, or even other closely linked loci, then the two closely linked MHC microsatellite loci would be expected to still show linkage disequilibrium with *DRB*. For example, if we generously assume that the recombination rate c between *DRB* and either of these loci is 0.005 and selection occurred even until 20 generations ago (40 years), then linkage disequilibrium would only have decayed $(1-c)^{20} = (0.995)^{20} = 0.905$, around 10% from its high.

Several other quandaries are presented by the data. If *DRB* were the gene under strong balancing selection, then it is surprising that it shows no variation at all on San Clemente Island, a much larger population. Does this mean that somehow this population has not been selected for *DRB* variation? Or why would have the Santa Catalina population recently declined to only 10 individuals from nearly 1000, primarily from canine distemper, when it had high variation

for the *DRB* locus and the other MHC loci surveyed.

One solution for these concerns is to note that for these new data, San Nicolas appears to be a genetic subset of the variation found on nearby Santa Catalina. For *DRB*, Santa Catalina has three alleles and San Nicolas has only the two most common of these alleles, so one could surmise a founder event or genetic drift could explain the difference. Similarly, *DQB* and the linked microsatellite loci have an average expected heterozygosity on San Nicolas and Santa Catalina of 0.35 and 0.55, a difference also consistent with nonselective effects. In other words, perhaps a combination of nonselective effects and not-so-extreme balancing selection could be responsible for the observations.

These data are an exciting look at the findings we may expect in more species using information from genome projects. However, direct evidence of the selective differences at given genes seems fundamental to confirm their importance in contemporary populations. From this study, I think it is too early to redesign captive breeding pro-

grams of endangered species to focus on, or even include, maximization of heterozygosity or retention of alleles at MHC genes as the authors suggest (Hedrick and Miller, 1994; Hedrick, 2003). For example, the four *DQB* alleles are very similar and differ by only a few nucleotides. Which alleles would be favored in a captive breeding program when we do not know anything about selective differences between them? More than anything, this study shows the great complexity of evaluating adaptive variation, even in one of our best-understood genetic systems.

P Hedrick is at School of Life Sciences, Arizona State University, Tempe, AZ 85287-4501, USA.

e-mail: philip.hedrick@asu.edu

Aguilar A et al (2004). *Proc Natl Acad Sci USA* **101**: 3490–3494.

Garrigan D, Hedrick PW (2003). *Evolution* **57**: 1707–1722.

Hedrick PW (2003) In: Holt WV, Pickard AR, Rodger JC, Wildt DE (eds) *Reproduction Science and Integrated Conservation*, Cambridge University Press: Cambridge. pp 97–113.

Hedrick PW, Miller PS (1994) In: Loeschcke V, Tomiuk J, Jain SK (eds) *Conservation Genetics*, Birkhauser: Basel. pp 187–204.