Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 6251–6252, July 1994

Commentary

The conundrum of nonclassical major histocompatibility complex genes

Jan Klein† and Colm O'hUigin**

**Max-Planck-Institut für Biologie, Abteilung Immungenetik, D-72076 Tübingen, Germany; and †Department of Microbiology and Immunology, University of Miami School of Medicine, Miami, FL 33101*

In addition to the typical major histocompatibility complex (MHC) genes, which are involved in the initiation of the anticipatory (specific) immune response, there is a group of atypical MHC genes whose functional status is uncertain (1, 2). The two groups are often referred to as classical and nonclassical MHC genes, respectively. The nonclassical MHC genes differ from the classical ones in several features. (*i*) There is considerable sequence dissimilarity between the two groups. (*ii*) While the pattern of expression of classical genes in different tissues is relatively uniform, that of the nonclassical genes is often erratic. (*iii*) Many of the nonclassical sequences are pseudogenes. (*iv*) The polymorphism of the nonclassical genes is usually lower than that of the classical genes. (*v*) If some of the nonclassical genes are functional at all, their function is widely assumed to be of a different nature than that of the classical genes.

The nonclassical MHC genes were originally described in the mouse and related species, but there is now growing evidence for their existence throughout the chordate phylum (3, 4). The term "nonclassical genes" is usually applied to the class I loci, but in fact some of the families of class II loci (e.g., the mammalian *DM* system) also fit the description. The gene discovered by Bahram *et al.* (5), reported in this issue of the *Proceedings*, is the first human nonclassical class I locus by the five criteria. (*i*) It belongs to a new family of class I loci distantly related to the *HLA-A, -B, -C* family, which represents the classical human class I loci. It *is*, however, very clearly a class I locus and we believe there is no justification to introduce a nonstandard designation ("*M*HC class *I* chain-related", or *MIC*, gene), albeit this designation has received provisional approval by the Commission on Nomenclature of the Human Gene Mapping Workshop. A standardized symbol will undoubtedly be assigned by the World Health Organization Nomenclature Committee. We will refer to the new gene as *HLA-X* (*X* for unassigned). Fig. 1 documents that *HLA-X* is indeed a bona fide class I gene. In the phylogenetic tree of representative MHC sequences, *HLA-X* clusters well within the class I family. Its origin pre-dates the separation of placental and marsupial mammals some 140 million years ago, but the gene is probably not much older than that. Most likely it is a mammalian invention. The long intron 1 and the fusion of the exons encoding the cytoplasmic tail and the 3′ untranslated region are by no means indicative of a special status of the *HLA-X* gene: mammalian genes, and the MHC genes in particular, are notorious for their variation in intron length, and variation in the number of exons at the 3′ end has been documented for several MHC genes (13).

(*ii*) The expression of *HLA-X* in fibroblasts and epithelial cell lines, and its apparent nonexpression in B- and T-cell lines, is in keeping with the erratic expression of other nonclassical MHC genes. The pattern may be significant, but it is equally possible, or even more likely, that the vagaries in expression reflect the declining functional importance of the nonclassical genes.

(*iii*) Although *HLA-X* does not appear to be a pseudogene, it would be surprising if some of the other genes of this family, in humans or in other mammals, do not turn out to be inactive.

(*iv*) Without gene frequency data it is premature to regard the allelic variation Bahrem *et al.* (5) found at the *HLA-X* locus as polymorphism. Moreover, the variation is really very modest, not exceeding that found at many non-MHC loci, and it is *not* at sites corresponding to the peptide binding region. Most important, however, polymorphism in the MHC is not necessarily indicative of functionality. The *HLA-DRB6* locus, for example, has probably been occupied by pseudogenes for most of its >60-million-year history, yet the locus is highly polymorphic (14). The *HLA-X* gene is closely linked to the highly polymorphic *HLA-B* gene, and *HLA-X* variants could therefore be maintained in the population by selection for *HLA-B* polymorphism.

(*v*) Barham *et al.* (5) argue that conservation of certain amino acid residues in the HLA-X protein is indicative of functionality. Their sequence comparisons are, however, biased by their restriction to human and mouse proteins only. An entirely different picture emerges when proteins from other vertebrate classes are included. Molecules that diverged >140 million years ago can be expected to have evolved under different functional constraints even if they remained functional during the entire period. Even if HLA-X does bind peptides, why should we expect its binding pockets to be similar to those in HLA-A or -B? A further important point is that even if HLA-X could be shown to bind peptides, this would not constitute proof of its functionality. To demonstrate that nonclassical MHC proteins are functional, it would have to be shown that they present peptides to T lymphocytes and that the presentation is physiologically important. In other words, one would have to prove that the absence of a particular nonclassical MHC protein leaves the host at a serious disadvantage in terms of resistance to specific pathogens. As far as we know, such proof has not been put forward for any of the nonclassical MHC proteins.

What then is to be made of the presence of nonclassical genes in the MHC? To us, it seems that their persistence can be comprehended only in the context of the evolutionary history of the entire system. Evidence is growing that the MHC, more than any other complex of duplicated genes, evolves in fits and spurts, in repeated cycles of expansion (duplication) and contraction (deletion), and that these cycles often coincide with periods of adaptive radiation of the major taxa (15). The cycles may be impelled by opposing demands on the MHC: on the one hand, the MHC has to be variable enough to meet the challenge of the pathogens; on the other hand, it cannot be too variable at the level of the individual because it would otherwise reduce the size of the T-cell repertoire to dangerously low levels. Theoretical considerations indicate that an optimal solution to this quandary is to keep only a small number (two to three) of loci functional and diversify them by polymorphic variation (16).

There appears, however, to be a need to replace, from time to time, the functional genes by new ones—to change guards, so to say. This need seems par-

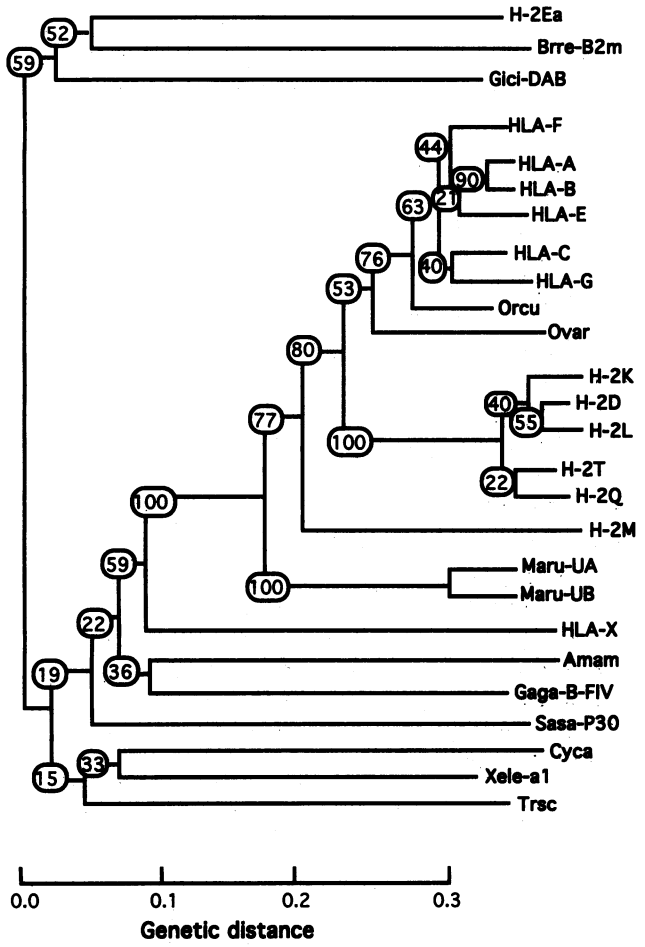

FIG. 1. Dendrogram showing the relationship of the human nonclassical gene *HLA-X* to selected vertebrate class I genes in the $\alpha 3$ domain. The dendrogram is rooted by using β_2-microglobulin (Brre-B2m; ref. 6) as well as class II α (H-2Ea; ref. 7) and β (Gici-DAB; ref. 8) chains. Distances are based on percent protein sequence identity and the dendrogram was constructed by the neighbor-joining method of Saitou and Nei (9). Figures on nodes indicate bootstrap percent recoveries from each node over 500 replications. References to sequences can be found in refs. 10, 11 (H-2M), and 12 (Trsc).

ticularly acute in periods when groups of vertebrates colonize new ecological niches vacated by other groups undergoing extinction, the periods of adaptive radiation. This timing makes sense because adaptation to a new niche almost certainly includes adjustment to new sets of parasites, and this could require a dramatic reorganization of the MHC. The circumstances of adaptive radiation might promote reorganization in that the fragmented populations, first reduced and then expanded in size, may provide favorable conditions for the fixation of the reorganized chromosomes. The period of reorganization, however, may generate more genes than the system could tolerate if all of them were to remain functional. To solve this problem, only a few genes of the set are chosen to be functional (those that optimally satisfy the rising demands) and all the others are kept in limbo—they become the group of nonclassical MHC genes.

The genes in the nonclassical category may come from two sources: first, from genes that were previously functional but were subsequently replaced by the new guard, and second, from genes that were generated *de novo* by duplication in the expansion cycle and that were rejected as candidates for the post of the primary functional genes. Nonclassical genes of different MHCs can therefore be generated at different times and so be unmatched in age. The *H-2Q*, *H-2T*, and *H-2M* genes may have arisen at the time of rodent (or murid) radiation; the *HLA-X* family of genes may have emerged during the period of the major mammalian radiation; and the *HLA-DM* genes may be leftovers from an even older radiation.

The nonclassical genes in limbo can be expected to have a variety of fates. Most of them eventually deteriorate into pseudogenes, although this process may take a very long time. The widespread belief that a nonutilized gene becomes rapidly garbled is not substantiated by actual data. The long persistence of genes for a large number of atavisms (17) attests to the fact that genes remain potentially functional millions of years after they became superfluous. Some other nonclassical genes in limbo may assume auxiliary functions. They may, for example (as is widely believed but not proven), specialize in the presentation of specific peptides derived from a particularly importune parasite. Finally, a third category of nonclassical genes may even be recruited at the next round of MHC reorganization to become the new guard of functional loci.

We thank Ms. Lynne Yakes for editorial assistance.

1. Klein, J. (1986) *Natural History of the Major Histocompatibility Complex* (Wiley, New York).
2. Wang, C.-R., Livingstone, A., Butcher, G. W., Hermel, E., Howard, J. C. & Fischer Lindahl, K. (1991) in *Molecular Evolution of the Major Histocompatibility Complex*, eds. Klein, J. & Klein, D. (Springer, Heidelberg), pp. 441–462.
3. Hashimoto, K., Nakanishi, T. & Kurosawa, Y. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6863–6867.
4. Flajnik, M. F., Kasahara, M., Shum, B. P., Salter-Cid, L., Taylor, E. & Du Pasquier, L. (1993) *EMBO J.* **11**, 4385–4396.
5. Bahram, S., Bresnahan, M., Geraghty, D. E. & Spies, T. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6259–6263.
6. Ono, H., Figueroa, F., O'hUigin, C. & Klein, J. (1993) *Immunogenetics* **38**, 1–10.
7. Hyldig-Nielsen, J. J., Schenning, L., Hämmerling, U., Widmark, E., Heldin, E., Lind, P., Servenius, B., Lund, T., Flavell, R., Lee, J. S., Trowsdale, J., Schreier, P. H., Zablitzky, F., Larhammar, D., Peterson, P. A. & Rask, L. (1983) *Nucleic Acids Res.* **11**, 5055–5071.
8. Bartl, S. & Weissman, I. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 262–266.
9. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
10. Mayer, W. E., Sevilir Williams, N., O'hUigin, C., Vincek, V., Zaleska-Rutczynska, Z. & Klein, J. (1993) *Mol. Phylogenet. Evol.* **2**, 23–30.
11. Singer, D. S., Hare, J., Golding, H., Flaherty, L. & Rudikoff, S. (1988) *Immunogenetics* **28**, 13–21.
12. Hashimoto, K., Nakanishi, T. & Kurosawa, Y. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2209–2212.
13. Steinmetz, M. & Hood, L. (1983) *Science* **222**, 727–733.
14. Figueroa, F., O'hUigin, C., Tichy, H. & Klein, J. (1994) *J. Immunol.*, in press.
15. Klein, J., Ono, H., Klein, D. & O'hUigin, C. (1993) in *Progress in Immunology*, eds. Gergely, J. & Petranyi, G. (Springer, Heidelberg), pp. 137–143.
16. Takahata, N. (1994) in *Non-Neutral Evolution: Theories and Molecular Data*, ed. B. Golding (Chapman & Hall, London), in press.
17. Hall, B. K. (1984) *Biol. Res.* **59**, 89–124.