Population Biology of Antigen Presentation by MHC Class I Molecules

Peter Parham and Tomoko Ohta

In principle, the function of major histocompatibility complex (MHC) molecules is simple: to bind a peptide and engage a T cell. In practice, placing this function within the context of the immune response begs questions of population biology: How does the immune response emerge from the interactions among populations of peptides, T cells, and MHC molecules? Within a population of vertebrates, how does MHC polymorphism stamp individuality on the response? Does polymorphism confer differential advantages in responding to parasites? How are the pressures on the MHC reflected in turnover of alleles? The role of mutation, recombination, selection, and drift in the generation and maintenance of MHC class I polymorphism are considered.

Tissues of the vertebrate body are composed of different cell types, each specialized for a particular function and collectively forming an interdependent and sophisticated society. Such complexity provides a rich environment for plunder by parasites (*1*), which themselves exist in various forms and possess diverse strategies for subverting particular cell types. Upon infection, the community of vertebrate cells becomes as vulnerable as its weakest member. Countering adversity are the cells of the immune system, dedicated to suppressing infection and maintaining tissue integrity (reviewed in 2). Specificity, repertoire, and memory in the immune response are determined by families of antigen binding molecules: immunoglobulins, T cell receptors (TCRs), and MHC class I and II glycoproteins. Somatic rearrangement of their immunoglobulin or TCR genes restricts individual B and T cells to a single specificity for antigen, but as a population the cells of the immune system possess millions of specificities. Immunoglobulin receptors on B cells bind to native protein antigens, whereas TCRs recognize short peptide fragments bound by polymorphic MHC glycoproteins. Of all structural polymorphisms, the alleles of MHC class I and II genes are most numerous, divergent, and evenly distributed within populations.

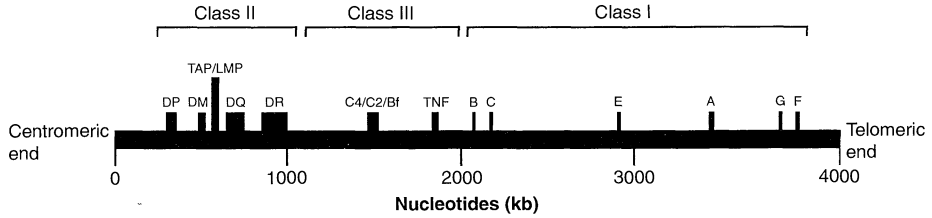
Because they lack wild-type alleles, MHC class I and II genes cannot be described in the simple manner that suffices for monomorphic genes. Statistical methods of population biology are required, some of which have already proved useful in the analysis of MHC structure and evolution (*3*). Such approaches may find application in the description of peptide populations presented by MHC molecules (*4*), the T cell populations that survey them, and the interactions among these populations—a development that will require sufficiently large sets of immunological data and analytical methods tailored to the peculiar properties of the antigen binding molecules of the immune system.

P. Parham is in the Department of Structural Biology, Stanford University, Stanford, CA 94305, USA. T. Ohta is at the National Institute of Genetics, Graduate University for Advanced Studies, Mishima, Shizuoka-ken 411, Japan.

MHC Polymorphism

MHC polymorphisms elicit strong responses when tissues are transplanted between MHC-disparate individuals. Recognition of this phenomenon, initially from experiments to propagate tumors in outbred mice, stimulated study of the immunogenetics of transplantation and the creation of inbred strains of mice essential to modern immunological research (*5*). Although model studies in mice can selectively focus on the small number of MHC types expressed by the inbred strains, the importance of MHC differences for clinical transplantation necessitated a more inclusive approach to the study of human populations. Typing for genes of the human MHC [the human leukocyte antigen (HLA) region] is now routine at transplant centers throughout the world, and includes typing of human populations of diverse geographical origin and ethnicity (*6*).

As a result of HLA typing and matching in clinical transplantation, knowledge of the population genetics of the human MHC far exceeds that of any other species (Fig. 1). For much of its 30-year history, HLA typing was based solely on the cytotoxic reactions of human alloantisera with live lymphocytes (*7*). However, this approach does not distinguish between all alleles, and typing based on nucleotide sequences now replaces serology. For HLA class II genes the transition is nearing completion, whereas for class I genes it is just beginning (*8*). Although population studies based on DNA-typing methods are at an early stage, it is already apparent that serological HLA data sets (*6*) have systematically underestimated population differences. The serologically indistinguishable HLA subtypes commonly mark populations or ethnic groups, and an overall insensitivity to the antigens



HLA class II alleles

Locus	Number of alleles
HLA-DRA	2
HAL-DRB	179
HLA-DQA	18
HLA-DQB	29
HLA-DPA	8
HLA-DPB	69
HLA-DMA	4
HLA-DMB	5
Total	314

HLA class I alleles

Locus	Number of alleles
HLA-A	67
HLA-B	149
HLA-C	39
HLA-E	5
HLA-F	1
HLA-G	6
Total	266

Fig. 1. HLA polymorphism. Linear diagram of the HLA region showing the relative positions of the major loci involved in antigen presentation. TAP/LMP, are the genes encoding the transporter for antigen processing (TAP), which transports peptides from the cytoplasm to the endoplasmic reticulum, and two low molecular weight polypeptides (LMP) of the proteasome, which produce peptides in the cytoplasm. The locations of the C4, C2, and factor B (Bf) complement loci and the tumor necrosis factor (TNF) genes in the class III region are also indicated. Tabulated beneath the class I and II regions is the number of alleles for each locus that are currently defined on the basis of nucleotide sequence (*72*). The definition of class II allelic sequences is more advanced than that of class I, and during the next few years the number of class I alleles can be expected to exceed that of class II.

of non-Caucasian populations exists because serological typing was developed primarily with alloantisera and cells obtained from Caucasians. The limitations of serological typing are apparent from a recent study reporting one or more differences in HLA-A, -B, and -DR antigen assignment for 48.3% of individuals (total, 2308) for whom independent typings were sought (*9*). DNA typing has the potential for greater sensitivity and precision and may soon offer more incisive sets of HLA data to the population geneticist (*8*).

Generation of MHC diversity. Ever since the identification of MHC molecules as "transplantation antigens," the mechanisms that generate and maintain MHC diversity have been debated (*10*). After the discovery in the 1970s that MHC polymorphisms determine the antigen specificity of T cells, it was proposed that the benefits of MHC polymorphism are obtained through heterozygosity, thereby increasing the number of parasite antigens to which an individual's T cells can respond (*11*). The neutral theory of molecular evolution (*12*), which holds that most evolutionary change at the molecular level is caused by random drift of selectively neutral or nearly neutral mutations, was also considered as a potential explanation for MHC polymorphism. This mode of evolution is now thought to apply to those parts of the genome that do not specify genetic information, including most synonymous (silent or noncoding) substitutions, which steadily accumulate with time. Natural selection on nonsynonymous (coding) substitutions may therefore be detected by comparison of the numbers of synonymous and nonsynonymous substitutions. If certain sites within an MHC gene have been selected for polymorphism of amino acid residues, then an increase in nonsynonymous substitutions within these sites, relative to other sites, is expected.

Testing the null hypothesis of neutrality requires knowledge of the nucleotide sequences for MHC alleles, which began to accrue in the 1980s (Fig. 1). Inspection of these sequences revealed regions in which nonsynonymous substitutions were overrepresented (*13*), but understanding their full significance awaited determination of the three-dimensional structures for MHC class I and II molecules (*14*). Elucidation of these structures resulted in definition of the peptide binding sites and the MHC residues (sites) that interact directly with bound peptide and the TCR. The ratio of nonsynonymous to synonymous substitutions for these functional sites was indeed found to be higher than for other sites, showing that natural selection has contributed to MHC polymorphism (*15*). Further application of this method revealed that the evidence for selection and its statistical significance varies with locus and species of origin (*16*).

Evolution of polymorphism. Investigation into the mechanism whereby MHC polymorphism evolves has centered around the events accompanying speciation. Within the paradigm that new species are formed from small populations of existing species (*17*) (in the extreme case from two individuals, as in Noah's ark), most preexisting MHC polymorphisms will be denied to the new species, which then has to develop polymorphism anew. Alternatively, if large populations of existing species evolve into new species, then much of the existing polymorphism would be inherited by the new species in a transspecies manner (*18*). The critical difference between these two models is that the former necessitates a mechanism for the rapid evolution of MHC genes, whereas the latter relies on the steady acquisition of diversity through the accumulation of substitutions as a result of a normal rate of point mutation.

The first comparisons of MHC alleles in different species were of humans and mice. The lack of correspondence between individual alleles in the two species (*19*) shows that 80 million years of separation have been sufficient to reconfigure all class I and II alleles. Subsequent studies have concentrated on comparison of humans and higher primates (*20*). The data show that closely related species share a high proportion of the nucleotide substitutions that form the polymorphism and that the extent of sharing decreases with increased time of separation. Invoking parsimony, the shared substitutions are assumed to have been present in the alleles of a common ancestor and to have been inherited in a transspecies manner by the lineages that gave rise to the modern species. Consistent with the theory of transspecies evolution, the rate at which new point substitutions are acquired is consistent with normal rates of mutation (*21*).

Similar MHC alleles are found in different species of higher primate and represent lineages that existed in the common ancestor and have persisted over long periods of time (up to 45 million years), being inherited through successive speciation events. Certain lineages are shared by humans and Old World monkeys, but are not observed in New World monkeys (*20, 22*). However, in no instance has an identical allele been found in two species, showing that the alleles are constantly undergoing modification during their transspecies evolution. The extent to which allelic lineages can be traced varies with the locus; for example, inspection of the class II locus HLA-DRB and its orthologs reveals extensive relationships among lineages (*23*), whereas for the class I locus HLA-B, lineages are rarely discerned (*24*).

Role of conversion. Comparison of the nucleotide sequences of HLA alleles reveals many examples of two alleles that differ by a short segment of sequence (rarely longer than 100 nucleotides) that is itself identical to the homologous sequence of a third allele (*23, 25, 26*). This pattern (Fig. 2) implies that recombination between alleles in a conversion-like manner (interallelic conversion) contributes to the formation of new alleles. Less common are conversions between alleles of different HLA loci (intergenic conversions), although their role in generating mutant mouse MHC class I genes provided the first evidence for conversion (*27*). The role of conversion in MHC polymorphism depends on the particular locus and is most apparent for HLA-B. The large numbers of HLA alleles formed by recombination (*28*) have persuaded some investigators that recombination is a significant player in the generation of MHC diversity; others have yet to be convinced.

Most mathematical methods used to study molecular evolution have assumed that nucleotide substitutions within a gene are accumulated in a stepwise manner through point mutation (*29*). Such methods do not accommodate intergenic or interallelic recombination of the type observed for MHC genes. Because of differences in the evolution of individual MHC loci, and of opinions regarding the importance of recombination, some investigators have applied conventional methods to calculate mutation rates and divergence times for MHC alleles, and the size of the founding human population (*30*). In contrast, on the basis of MHC class I data, we have been impressed with the importance of recombination and the need to incorporate its effects into the analysis (*31, 32*).

The result of the debate can be viewed as either a compromise or a standoff. The rates of point mutation for MHC genes are not unusual, and extensive polymorphism results directly from their accumulation over many millions of years of transspecies evolution (*33*). The recombinational mechanism of conversion is superimposed upon this process. Conversion cannot by itself generate diversity, but once alleles have been built through accumulation of point substitutions, conversion can substantially increase their numbers and extent of divergence (*34*). Whether this mechanism for rapid generation of new alleles is a specific property of MHC genes is unknown, because its detection is only likely for polymorphic genes that have a high density of substitutions. For most genes, conversion would have no phenotypic effect, and in the absence of selection could act to reduce polymorphism, as may occur for the nonpolymorphic regions of some MHC genes (*28*). Although most attention has been focused on species with high MHC polymorphism, species with low polymor-

phism exist. Whether their homogeneity stems from a Noah's ark effect during speciation or from subsequent contractions in population size is unknown (*35*). When MHC diversity of a population becomes drastically reduced, the generation of new alleles becomes increasingly dependent on the slow point mutation rate, a feature of potential importance for biological conservation (*36*).

Selection and Recombination

In both structure and function, MHC class I and II represent variations on a theme and fulfill complementary roles in immunity. They differ in the intracellular site of peptide binding (endoplasmic reticulum for class I and endosomes for class II), the length of the peptides to which they bind (8 to 10 amino acids for class I, 12 to 24 for class II), and the type of T cells to which they present antigen (cytolytic $CD8^+$ T cells for class I, helper $CD4^+$ T cells for class II) (*2*). MHC class I and II also differ in the partition of the four extracellular domains that exist between two constituent polypeptides. Circumstantial evidence suggests that the common ancestor of class I and II molecules had an organization similar to that of a modern class II molecule (*32, 37*), and a similar conservatism is apparent in the organization of mammalian MHC class II genes. In contrast, the organization of MHC class I genes and their numbers vary greatly: No orthologous relationships exist between individual class I genes of humans and mice (*19, 38*). Indeed, it appears that many rounds of duplication, differentiation, and deletion have been visited on the class I genes of the rodent and primate lineages since their separation ~80 million years ago (*39*).

The number of MHC class I genes in mammalian species varies from 6 to over 1000 (*38*). High polymorphism is confined to one to three "classical" (class Ia) genes that function in antigen presentation and are expressed by a wide range of cell types. The less polymorphic genes consist of pseudogenes and expressed "nonclassical" (class Ib) genes with varying tissue distributions (*40*). Certain nonclassical genes have specialized antigen-presenting functions—for example, presentation of *N*-formyl-methionine–containing peptides by the *Hmt2* gene of mice (*41*)—and may originally have derived from selected alleles of classical class I loci. The three classical human genes, HLA-A, -B, and -C, have differences in regulation, cell surface expression, and patterns of polymorphism, suggesting that they too have evolved specializations.

Peptides bound by HLA class I molecules are the products of cytoplasmic degradation of cellular proteins that are transported into the endoplasmic reticulum. There they associate with a polymorphic HLA-A, -B, or -C heavy chain and the invariant β_2-microglobulin (β_2M) to form a stable trimer that moves to the cell surface (*42*) (Fig. 3A). In healthy cells the peptides are derived from normal cellular proteins, and the immune system is rendered tolerant to these peptides during development. Upon infection, class I molecules loaded with pathogen-derived peptides are recognized by cytolytic $CD8^+$ T cells that then kill the infected cells (*43*). The peptides bound by a particular class I allotype are defined by positions within the peptide sequence that are restricted to one or a few amino acids. The preferred residues are termed anchors because their side chains extend into pockets of the binding site. In

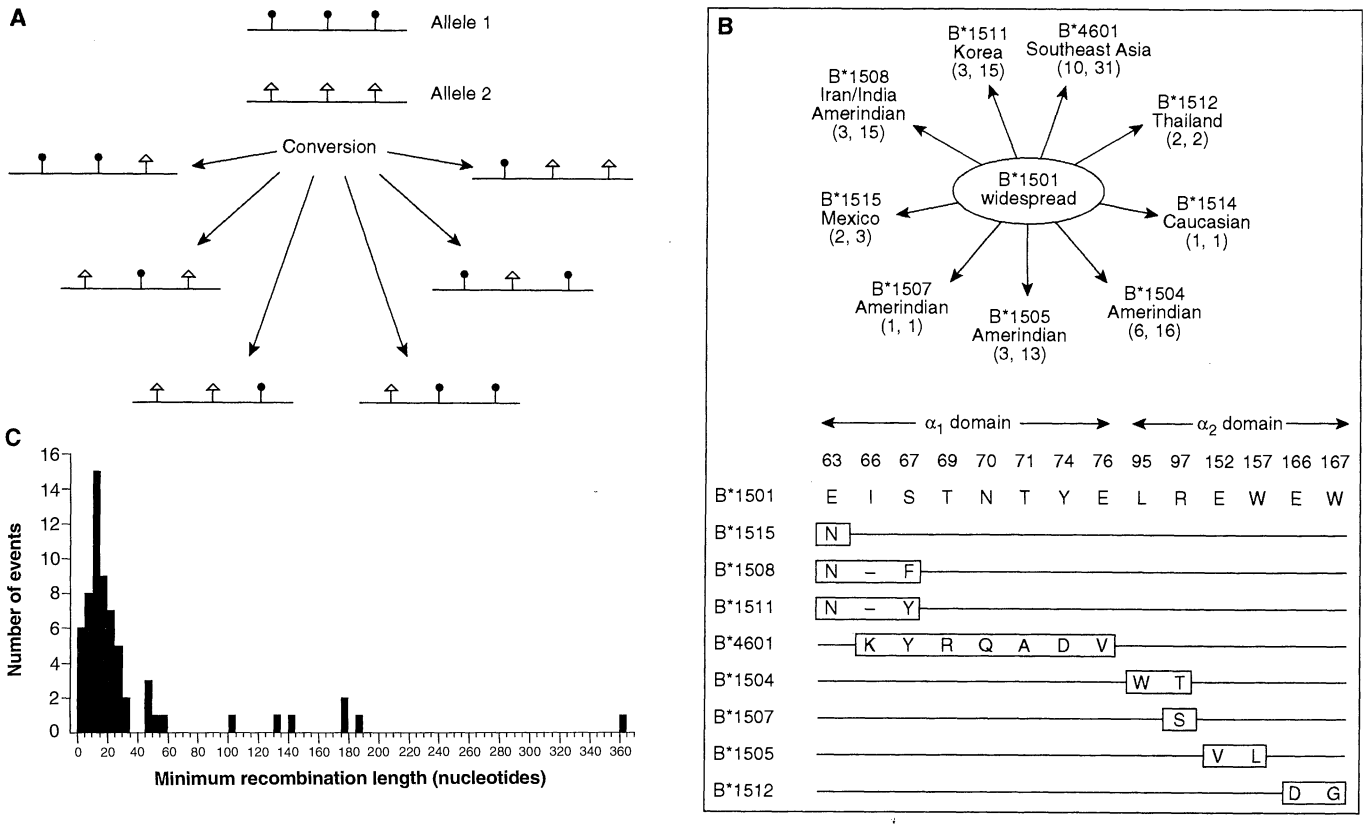


Fig. 2. Gene conversion. (**A**) Diagram showing how three substitutions in two alleles of the same or different MHC class I genes can generate six new alleles under the action of gene conversion. (**B**) Examples of alleles in the B*15 allele family that have been generated from the common and widespread B*1501 allele by conversion with other HLA class I alleles. (Top) Geographic or ethnic origins of the individuals from whom the alleles were characterized, followed by the number of nucleotide substitution and the minimum length of the conversion segment (in parentheses). (Bottom) The amino acids changed by the conversions and their juxtaposition within the α_1 and α_2 domains, producing the characteristic patchwork pattern. All the amino acid changes involve contact residues of the antigen recognition site. (**C**) Histogram showing the distribution of the minimum lengths involved in the conversions identified from simple comparison of HLA-A, -B, and -C sequences [reprinted from (*25*)].

aggregate, HLA class I allotypes have a range of class I peptide binding motifs that covers the spectrum of peptide characteristics: acidic, basic, neutral, and hydrophobic (*44*).

Balancing selection is thought to be responsible for MHC polymorphism and stems from interactions of the immune system with parasites (*1*, *15*, *25*, *26*, *45*). The importance of different types of selection, all of which act to expand the presentation of antigens, has been the subject of debate, and models, simulations, and statistical analyses have yet to distinguish between them (*46*). Two broad categories of selection, typified by overdominant selection (heterozygote advantage) and frequency-dependent selection, focus on immunological problems that are, respectively, of a general or specific nature and of long or short duration, and represent issues of thermodynamics (states at equilibrium) versus kinetics (reaction pathways between states).

Overdominant selection. The general and unchanging problem addressed by overdominant selection is that survival and reproduction of an individual, regardless of time or place, requires effective immunity against a wide range of parasites. That MHC class I and II molecules have sites with highly degenerate binding specificities constitutes a basic solution to this problem as it pertains to antigen presentation. By driving the evolution of allotypes that bind increasingly nonoverlapping populations of peptides, overdominant selection allows heterozygotes to present up to twice the number of peptides from a parasite as can homozygotes—a difference that should result in a more potent T cell response. For individuals, the benefits of heterozygosity will depend on the allotypes they express and will diminish according to the overlap in the repertoires of peptides they present. Overlap between pairs of allotypes varies markedly, correlating roughly with sequence divergence within the peptide binding site. Overdominant selection is therefore expected to favor the formation of sets of divergent peptide binding specificities through accumulation of numbers of amino acid substitutions—a trend that is evident in the lineages of divergent MHC alleles found within species (*33*). The effect is pronounced in the six "families" of HLA-A alleles, but no such divisions can be made in the larger numbers of HLA-B alleles (*47*).

The vast majority of human beings are indeed HLA heterozygous. Six populations of class I molecules are expressed, each defined by a specific HLA-A, -B, or -C heavy chain allotype. Up to 250,000 copies of each molecule are present at the cell surface, the number varying with cell type. The allotype-defined populations are further divided into thousands of functional subpopulations, each defined by a particular peptide species (*48*) (Fig. 3B). This diversity, which derives from incorporation of a variable peptide subunit into the class I molecule, enables a small number of class I heavy chain allotypes to select a TCR repertoire estimated to be in the millions (*49*). In principal, every structural gene within the human genome can contribute to the microheterogeneity of HLA class I molecules and the T cell response. Similarly, subpopulations of HLA class I molecules defined by individual parasite-derived peptides of infected cells can stimulate individual clones of cytotoxic T cells.

Frequency-dependent selection. Overdominant selection can be considered as a constant force that is relatively indifferent to individual parasites and the special demands they pose. Populations faced with epidemic disease and severe loss in numbers are under strong selection pressure from a single parasite; under these conditions any allele that, by chance, provides better immunity against that parasite will increase in frequency. The increased frequency of B*3501 in the Gambian population seems to be a direct result of selection by the malarial parasite and the disease it causes (*50*)—selection that is dependent on time, place, and the parasite. An individual episode of selection is transient compared with the time frame of overdominant selection. Under certain conditions, more persistent selection might increase the frequency of an allele to a point at which it dominates the locus. However, parasites can adapt to the MHC allotypes of their host by mutation of the peptide sequences to which the immune system responds. Such adaptation is likely to be directed at the most frequent alleles in the host population, providing selective advantage for hosts that express rare alleles (*46*, *51*). This mode of selection, which may underlie the association of HLA-B*5301 and HLA-DRB1*1302 with malaria, is frequency dependent and by nature transient: As the selected allele increases in frequency, it becomes a more likely target for parasite adaptation.

Recombination. Selection by parasites, whether individually or as a group, is not the only force determining the evolution of MHC class I genes. Recombination is a second important factor. The evolution of single-copy genes occurs by stepwise accumulation of point substitutions in a wild-type allele. Point substitutions have been a focal point of study in molecular evolution and are perhaps epitomized by mitochondrial DNA, which mutates rapidly and does not recombine (*33*). MHC class I genes, however, do not obey these rules. A variety of recombinational mechanisms serve to alter the diversity beyond that arising from point substitution (*28*, *32*, *52*). Unequal crossing-over can serve to duplicate and delete genes; recombination and gene conversion reassort the substitutions that arise separately in individual alleles and genes (*31*, *53*).

Population genetic models that incorporate unequal crossing-over and gene conversion, in addition to natural selection, mutation, and random genetic drift, have been used to simulate the evolution of a class I gene family (*54*). Whether selection is applied to all or a subset of the loci, diversity appears to be acquired rapidly compared with acquisition at a single locus. Furthermore, adjusting the rates of gene conversion can result in markedly different types of class I gene family, as are observed in nature (*38*). By use of a model based on the organization of class I genes in the mouse, many features of the polymorphism can be simulated with relatively weak selection per individual site at a level that is "near neutrality" (*55*). The estimated selection coefficient for B*5301 in the Gambian population (0.028) and calculations based on estimated substitution rates in MHC class I and II genes (*56*) are also consistent with weak selection on MHC alleles.

Furthermore, linkage disequilibria are formed between sets of substitutions that tend to be converted en bloc, producing a patchwork pattern that is characteristic of MHC alleles (*31*). These simulations indicate that interaction among diversifying se-

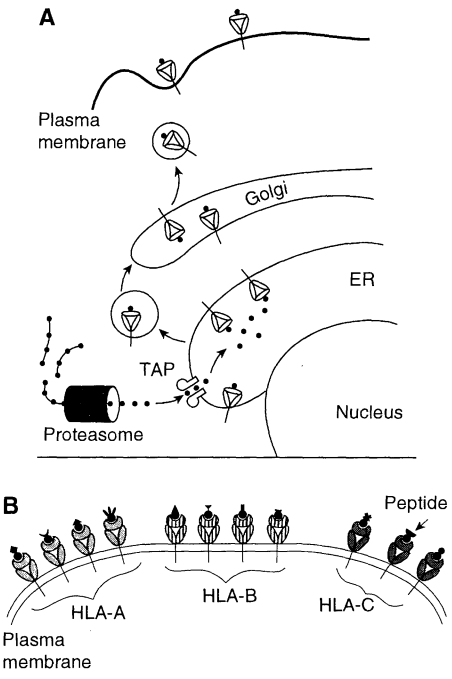


Fig. 3. (**A**) Cartoon showing the class I pathway for processing and presentation of peptide antigens. (**B**) Functional subpopulations of HLA-A, -B, and -C molecules are defined by the individual peptides that are bound. For simplicity, the cell shown is homozygous for the HLA region.

lection, gene conversion, and random drift is important for the acquisition and maintenance of polymorphism in MHC class I gene families. In particular, relatively mild selection at individual sites within the peptide binding site is effective for enhancing polymorphism under gene conversion (*54, 55*). This process resembles the formation of complementary blocks of linked genes in Franklin and Lewontin's (*57*) simulation experiments. As a way of conceptualizing a problem of population genetics, they produced a model of linked overdominant loci and found that co-adapted gene blocks are formed under the joint effects of overdominance, limited recombination, and random genetic drift—a process that might be termed crystallization, but that requires further investigation because of the abundance of interactive systems in biology.

Studies of Native American Populations

The Americas, in contrast to the other continents, were colonized relatively recently (~10,000 to 40,000 years before present) by small, human populations migrating from Asia across the Behringian land bridge to Alaska (*58*). By southward migration their progeny populated much of North and South America, where they evolved in relative isolation until the arrival of Europeans in the 16th century. Evidence from archeology, linguistics, and genetics is consistent with this scenario (*59*) and is further supported by the distribution of HLA class I alleles in modern Amerindian populations (*60*). Owing to the manner of the colonization and its time frame, study of modern Amerindian populations has revealed unique insights into the evolution of HLA class I polymorphism in human populations (Fig. 4).

Amerindian populations have much smaller numbers of HLA-A, -B, and -C alleles (representative numbers being three, eight, and five respectively) than the urban populations studied by most HLA typing laboratories (*60*). Furthermore, the alleles found in Amerindian populations are identical or else evolved recently from four HLA-A, nine HLA-B, and seven HLA-C alleles that are shared with European and Asian populations and that probably represent alleles present in the founding population. However, there are notable differences between Amerindian populations in North and South America. For modern North Amerindian populations, almost all HLA-A, -B, and -C alleles are identical to one of the founding alleles. In contrast, South Amerindian populations express many "new" alleles that derive from founding alleles either by point mutation or interallelic conversion. A marked feature is the large number of recombinant B alleles, which for South American populations can constitute a majority of the HLA-B alleles (*61*). In contrast, HLA-A is represented mainly by founding alleles together with a small number of recombinants. HLA-C also exhibits relatively few "new" alleles (*62*).

On the basis of these comparisons, several observations emerge: (i) Over a comparable period of time, the cohort of HLA class I alleles of a population may remain constant, as in North America, or may change, as in South America. (ii) The HLA-A, -B, and -C loci have evolved in America at a different rate and with differential use of interallelic conversion. (iii) Recombination generates new alleles at a substantially greater rate than point substitution. These conclusions are not unique to Amerindian populations, because comparison of all known HLA-A, -B, and -C alleles reveals evidence of similar evolutionary patterns (*25*).

Despite the presence of new HLA class I alleles, the total number of alleles in South Amerindian populations has not increased relative to those found in North Amerindian populations or that of the hypothetical founding population. Thus, there is no evidence that selection increases the number of alleles relative to that present in the founding population, and it may therefore be erroneous to perceive the "small" number of alleles present in the founding population as disadvantageous. For scientists and clinicians used to wrestling with the

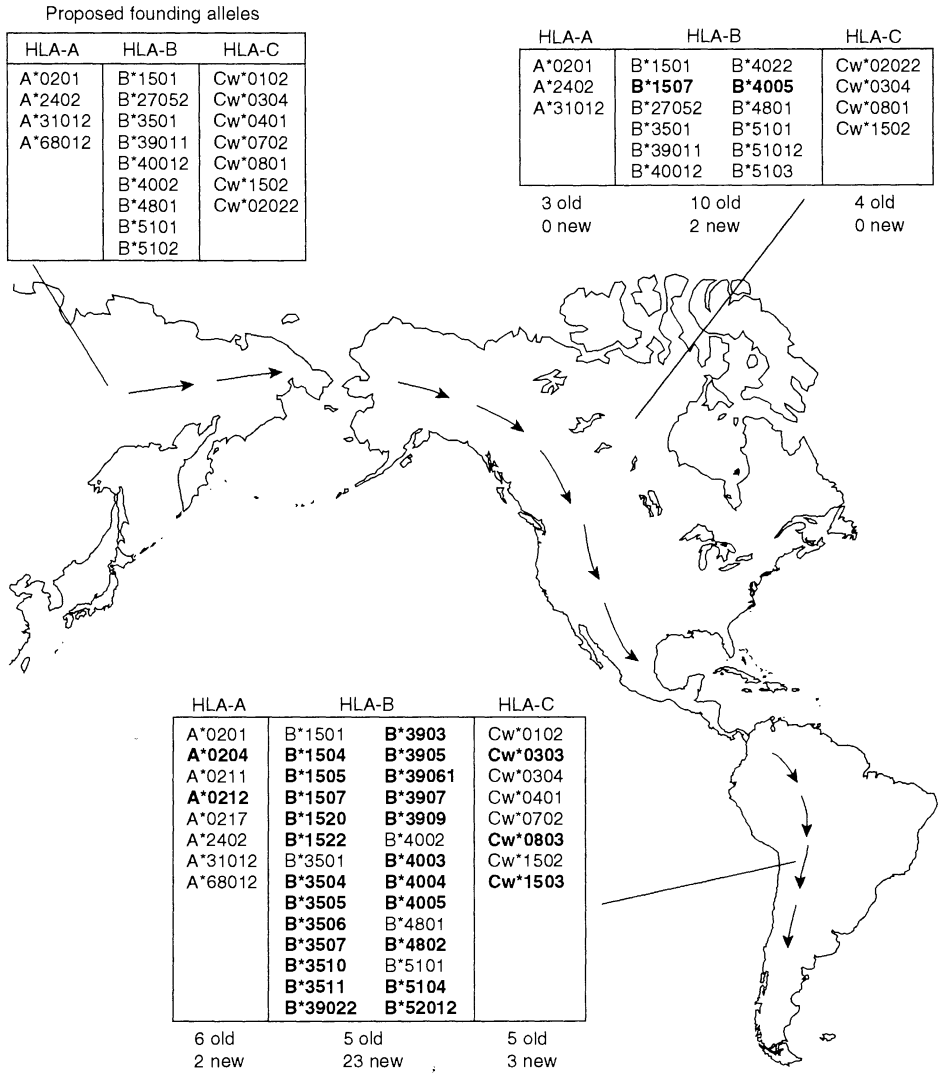


Fig. 4. Sketch map showing the route by which humans first colonized the Americas by migration from eastern Asia. The boxes next to North and South America show, respectively, the HLA-A, -B, and -C alleles that have been detected in four North Amerindian and six South Amerindian populations (*61, 62*). Alleles in bold type are those not detected in non-American populations. All the novel HLA-A and -B alleles are recombinants, whereas the novel HLA-C alleles appear to have arisen by point mutation. The box next to eastern Asia contains a hypothetical set of founding alleles that could account for all the alleles present in modern Amerindian populations. The outlines of the continents are adapted from (*71*).

complexity of HLA type characteristic of urban populations, the subset of HLA alleles present in Amerindian populations may seem unusually small in number and has been proposed as an explanation for the vulnerability of this population to diseases brought from Europe and Africa during the post-Columbian colonization of America (*51*). However, the number of alleles in Amerindians is comparable to those present in indigenous populations from other continents (*63*) and may more accurately represent those arising historically in human populations under natural selection by parasites.

Cities are a relatively recent product of human civilization, and only in the last 150 years have they managed to sustain their populations (*64*). In the process of urbanization, migrants drawn from rural areas become assimilated through commerce, competition, and conflict. South Amerindian populations have revealed the intrinsic capacity for generating new HLA class I alleles, although in these traditional communities the total number of alleles remains constant. If the various South Amerindian populations became assimilated into a single population, in a process approximating that historically involved in the formation of urban populations, then the total number of HLA class I alleles in the resulting admixture would be much greater than that of any of the contributing populations. Such a thought experiment [performed for other markers by Neel (65)] suggests that the large numbers of HLA class I and II alleles found in modern urban populations are predominantly the result of admixture bringing together alleles that evolved under natural selection in previously separated populations (66). With the invention of faster and more economical methods of transportation, the catchment areas for cities have continually expanded. In the last 500 years, cities in the United States have provided destinations for tens of millions of people arriving from an area encompassing much of the world. As a result, these urban populations have an unprecedented diversity of HLA type that presents unique challenges for transplant clinicians seeking to type and match for HLA. This diversity, however, has as much to do with pressures of economy as with those from parasites.

In Amerindian populations, the ascendancy of new recombinant HLA-B alleles is accompanied by the loss of founding alleles. A striking example of such turnover is provided by the Guarani population of southern Brazil, for whom all eight HLA-B alleles are new recombinants (*62*). Comparison of HLA-A, -B, and -C alleles with their orthologs (Patr-A, -B, and -C) from the common chimpanzee (*Pan troglodytes*) demonstrates the tendency to modify and turn over class I alleles. The ~5 million years since divergence from a common ancestor (*67*) has not led to appreciable differences in the number, organization, and identity of the MHC class I genes. However, no single HLA-A, -B, or -C allele is shared by human and chimpanzee populations, although numerous similarities in lineage, polymorphic motifs, and individual substitutions can be discerned (*20*). The differences always include changes in the peptide binding site and, consistent with the rate of evolution in human populations, the B alleles of the two species are most divergent (*22*). Thus, few if any of the MHC class I alleles present in the common ancestral population have survived unmodified to the present day. A time frame of ~5 million years is therefore sufficient to produce complete turnover of a population of MHC class I alleles, but not to produce changes of gene organization of the type that distinguishes humans and mice (*38*).

Selection Versus Drift: A Synthesis

The participation of recombination in the generation of HLA diversity is apparent at different levels. Sexual reproduction produces new combinations of HLA haplotypes for each generation, and meiotic recombination creates new haplotypes. During the lifetimes of species, conversion creates new alleles; unequal crossing-over does the same for the loci, but on a longer time frame. Recombination has the potential to produce rapid diversification, but requires a preexisting accumulation of nucleotide substitutions that can only be achieved over tens of millions of years. The evolutionary interplay between mutation and recombination varies with MHC locus, and even for subregions within a gene (*23*). Statistical analysis of allelic sequences reveals the actions of selection as well as differences between loci and species (*15, 16*). The statistical approach, however, does not assess the selective advantages conferred by the individual events of gene conversion or point substitution that create new alleles.

Analysis of the immune response emphasizes the functional differences between MHC alleles: differences in the populations of the peptides bound, the T cell repertoire selected, and the T cells activated in response to specific antigens. The combination of HLA alleles possessed by an individual stamps a uniqueness upon the immune system that determines the fine specificity of the immune response to a particular pathogen, tumor or transplant. However, from work in the field (*46, 50*), correlations of HLA type with susceptibility to a particular infectious disease are conspicuous by their rarity, suggesting that the persistent and enduring selection on MHC genes has favored presentation of a peptide repertoire that provides coverage of all pathogens. This argument reflects basic aspects of protein structure, rather than disease-specific characteristics, and is consistent with an overdominant selection: the evolution of multiple loci having divergent alleles with complementary peptide binding specificities.

Diseases tend to be episodic and in the short term can have a profound impact on both individuals and populations (*64*). In this instance, rare alleles (of which one newly formed is the rarest) can be advantageous if they provide patterns of antigen presentation to which the disease-causing parasite has not accommodated. Only four HLA class I alleles differing by novel point substitutions have been found (*25*), indicating that point mutation has played a negligible role in diversifying HLA function during the history of the human species. This hypothesis is consistent with estimates of 2 million years to fix a point mutation in an MHC gene and is the nub of transspecies evolution of MHC polymorphism (*18*). In contrast, gene conversion has produced over 80 new HLA-A, -B, and -C alleles during the lifetime of human populations, and many more should be discovered once DNA typing has been applied to the understudied populations of Africa and Asia. At some time and place, each of these recombinants could have had a selective advantage in raising immunity against a specific parasite. In this context, the studies of malaria and HLA-B*5301 set a precedent (*46, 50*). HLA-B*5301 is the product of a single gene-conversion event that occurred in West Africa, and currently it has the edge over the common and widespread B*3501 allele from which it derived, in its ability to protect against malaria. The edge is slight, as evidenced by the estimated selection coefficient of 0.028 (*46*), and could easily be lost once the malarial parasite adapts to evade antigen presentation by B*5301. Such an adaptation seems to have taken place in an East African population in which HLA-B*5301 offers no measurable protection to malaria (*46*). Furthermore, the protection afforded by DRB*1301 in West Africa has been lost in East Africa, where protection is conferred by another DR allele (*46*). The duration of such disease-specific selection can be short-lived, but its impact on HLA allele frequencies will be felt over much longer periods of time. The ephemeral nature of the protection offered by HLA alleles against malaria in Africa today illustrates how independent episodes of disease in the tribal populations of South America could have produced frequency-dependent selection for rare recombinant alleles, a consequent loss of older alleles, and a continuing process of allele turnover.

HLA homozygotes are generally healthy.

Such observations have indicated that the benefits of MHC polymorphism are subtle and not omnipresent. The low selection coefficient estimated for the protective effect of the B*5301 recombinant is consistent with this hypothesis, as are the simulations showing that gene conversion in combination with weak selection is sufficient to produce and maintain MHC polymorphism (54). For the time frame spanning the history and future of human populations, gene conversion is probably the dominant mechanism producing new and functionally distinctive HLA alleles. Conversion involves the reassortment of clusters of substitutions that have been tested by prior selection and is thus intrinsically more efficient than point mutation in producing functionally viable but distinctive products. Recent data indicate that the rate of conversion is also much higher than that of point mutation (68).

Pairs of HLA class I subtypes that are related by gene conversion can have distinctive or similar peptide binding specificities (69). The latter type may correspond to conversions that are selectively neutral or nearly neutral. Populations can lose old alleles and fix new ones through the action of genetic drift, an effect that increases with decreasing population size. Thus, the combined action of gene conversion and genetic drift could have contributed to the turnover of alleles at the HLA-B locus in South Amerindian populations. When alleles are lost through genetic drift, homozygosity increases. The frequency of rare alleles can then increase (under overdominant selection, rather than parasite-specific selection, which in these circumstances becomes a form of frequency-dependent selection) and contribute to allele turnover. Thus, the effects of specific parasites, genetic drift, and overdominant selection can reinforce each other to turn over the recombinant alleles formed by gene conversion. Comparison of humans and chimpanzees shows that in the long term such turnover is relentless. The temporary stability of North Amerindian populations compared with those of South America is likely to stem both from differences in their exposure to parasites and in the histories of their population structure.

Future progress in the population biology of antigen presentation will require application of mathematical approaches that incorporate gene conversion, drift, and the selection by parasites. Simulation of the history of the human MHC in the indigenous people of America would seem to be a rich but manageable system for testing such models. Greater knowledge of the distribution of HLA alleles in human populations should facilitate matching of donors and recipients for transplantation, particularly through international cooperation, and also facilitate tracing of the history and origins of human populations (70). The more ambitious and long-term goal will be to characterize the interactions among populations of peptides, TCRs, and MHC molecules in order to predict and manipulate the immune response to current and future parasites.

REFERENCES AND NOTES

1. Here, parasite is used as a generic term for any agent that colonizes tissues or infects cells. J. Klein and C. O'hUigin, *Philos. Trans. R. Soc. London Ser. B* **346**, 351 (1994); A. V. S. Hill, C. E. M. Allsopp, S. R. N. Yates, A. J. McMichael, in *HLA 1991: Proceedings of the Eleventh International Histocompatibility Workshop and Conference*, K. Tsuji, M. Aizawa, T. Sasazuki, Eds. (Oxford Univ. Press, Oxford, 1992), vol. 2, pp. 63–72.
2. W. E. Paul, Ed., *Fundamental Immunology* (Raven, New York, 1993).
3. A. L. Hughes and M. K. Hughes, *Immunogenetics* **42**, 233 (1995).
4. ———, *ibid.* **41**, p. 257.
5. J. Klein, *Biology of the Mouse Histocompatibility-2 Complex* (Springer-Verlag, New York, 1975).
6. T. Imanishi, T. Akaza, A. Kimura, K. Tokunaga, T. Gojobori, in *HLA 1991, Proceedings of the Eleventh International Histocompatibility Workshop and Conference*, K. Tsuji, M. Aizawa, T. Sasazuki, Eds. (Oxford Univ. Press, Oxford, 1992), vol. 1, pp. 1065–1220.
7. P. I. Teresaki and J. D. McClelland, *Nature* **204**, 998 (1964); W. F. Bodmer, M. Tripp, J. Bodmer, in *Histocompatibility Testing*, D. B. Amos and J. J. van Rood, Eds. (Munksgaard, Copenhagen, 1967), pp. 141–149; *History of HLA: Ten Recollections*, P. I. Terasaki, Ed. (UCLA Tissue Typing Laboratory, Los Angeles, CA, 1990).
8. J. Bidwell *Immunol. Today* **15**, 303 (1994).
9. J. Hegland *et al.*, *Hum. Immunol.* **44** (suppl.), 126 (1995).
10. B. Clarke and D. R. S. Kirby, *Nature* **211**, 999 (1966); G. D. Snell, *Folia Biol.* **14**, 335 (1968); F. M. Burnet, *Nature* **226**, 123 (1970); *ibid.* **232**, 230 (1971); W. F. Bodmer, *ibid.* **237**, 139 (1972).
11. R. M. Zinkernagel and P. C. Doherty, *Nature* **248**, 701 (1974); *ibid.* **256**, 50 (1975); *Lancet* **i**, 1406 (1975).
12. M. Kimura, *ibid.* **217**, 624 (1968); *ibid.* **267**, 275 (1977); *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983).
13. C. Jaulin *et al.*, *Immunogenetics* **22**, 453 (1985); C. N'Guyen, R. Sodoyer, J. Trucy, T. Strachan, B. R. Jordan, *ibid.* **21**, 479 (1985).
14. P. J. Bjorkman *et al.*, *Nature*, **329**, 506 (1987); P. J. Bjorkman *et al.*, *ibid.*, p. 512; P. J. Bjorkman and P. Parham, *Annu. Rev. Biochem.* **59**, 253 (1990); M. A. Saper, P. J. Bjorkman, D. C.Wiley, *J. Mol. Biol.* **219**, 277 (1991); J. H. Brown *et al.*, *Nature* **364**, 33 (1993).
15. A. L. Hughes and M. Nei, *Nature* **335**, 167 (1988); A. L. Hughes, M. K. Hughes, C. Y. Howell, M. Nei, *Philos. Trans. R. Soc. London Ser. B* **346**, 359 (1994); A. L. Hughes and M. Nei, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 958 (1989).
16. J. Zemmour and P. Parham, *J. Exp. Med.* **176**, 937 (1992); T. Ohta, *J. Mol. Evol.* **41**, 115 (1995); N. Yuhki and S. J. O'Brien, *J. Exp. Med.* **172**, 621 (1990).
17. E. Mayr, *Systematics and the Origin of Species* (Columbia, New York, 1942); *Change of Genetic Environment and Evolution*, J. Huxley, A. C. Hardy, E. B. Ford, Eds. (Allen Unwin, London, 1954), pp. 157–180; E. Mayr, *Populations, Species and Evolution* (Harvard Univ. Press, Cambridge, MA).
18. B. Arden and J. Klein, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2342 (1979); J. Klein, in *Immunology 80*, M. Fougereau and J. Dausset, Eds. (Academic Press, London, 1980), pp. 239–253; ——— and F. Figueroa, *CRC Crit. Rev. Immunol.* **6**, 295 (1986); J. Klein, *Hum. Immunol.* **19**, 155 (1987).
19. J. H. Rogers, *EMBO J.* **4**, 749 (1985); A. L. Hughes and M. Nei, *Mol. Biol. Evol.* **12**, 247 (1989).
20. W. E. Mayer *et al.*, *EMBO J.*, **7**, 2765 (1988); D. A. Lawlor, F. E. Ward, P. D. Ennis, A. P. Jackson, P. Parham, *Nature* **335**, 268 (1988); H. A. Erlich and U. Gyllensten, *Hum. Immunol.* **30**, 110 (1989); D. A. Lawlor, E. Warren, F. E. Ward, P. Parham, *Immunol. Rev.* **113**, 147 (1990); J. Klein, Y. Satta, N. Takahata, C. O'hUigin, *J. Med. Primatol.* **22**, 57 (1993).
21. Y. Satta, N. Takahata, C. Schonbach, J. Gutknecht, J. Klein, in *Molecular Evolution of the Major Histocompatibility Complex*, J. Klein and D. Klein, Eds. (Springer, Heidelberg, 1991), pp. 51–62.
22. D. I. Watkins, *CRC Crit. Rev. Immunol.* **15**, 1 (1995).
23. H. A. Erlich and U. B. Gyllensten, *Immunol. Today* **12**, 411 (1991); U. B. Gyllensten, E. Sundvall, M. Ezcurra, H. A. Erlich, *J. Immunol.* **156**, 4368 (1991).
24. K. Kato, J. A Trapani, J. Allopenna, B. Dupont S. Y. Yang, *J. Immunol.* **143**, 3371 (1989); D. A. Lawlor, E. Warren, P. Taylor, P. Parham, *J. Exp. Med.* **174**, 1491 (1991); J. D. Domena, W. H. Hildebrand, W. M. Bias, P. Parham, *Tissue Antigens* **42**, 156 (1993); A. L. Hughes, M. K. Hughes, D. I. Watkins, *Genetics* **133**, 669 (1995).
25. P. Parham, E. J. Adams, K. L. Arnett, *Immunol. Rev.* **143**, 141 (1995); W. Hildebrand *et al.*, *Tissue Antigens* **43**, 209 (1994).
26. S. N. McAdam *et al.*, *Proc. Acad. Natl. Acad. Sci. U.S.A.* **91**, 5893 (1994); M. K. Kuhner *et al.*, *Tissue Antigens* **38**, 152 (1991); J. C. Stephens, *Mol. Biol. Evol.* **2**, 539 (1985).
27. R. A. Flavell *et al.*, *Science* **233**, 437 (1986); S. G. Nathenson, J. Geleibter, G. M. Pfaffenbach, R. A. Zeff, *Annu. Rev. Immunol.* **4**, 471 (1986); J. Geleibter and S. G. Nathenson, *Trends Genet.* **3**, 107 (1987).
28. P. Parham *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4005 (1988); P. Parham, D. A. Lawlor, C. L. Lomen, P. D. Ennis, *J. Immunol.* **142**, 3937 (1989).
29. W.-H. Li and D. Graur, *Fundamentals of Molecular Evolution* (Sinauer, Sunderland, MA, 1991).
30. F. Ayala, A. Escalante, C. O'hUigin, J. Klein, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6787 (1994); N. Takahata, *Mol. Biol. Evol.* **10**, 2 (1993); J. Klein, N. Takahata, F. J. Ayala, *Sci. Am.* **269** (December 1993), p. 78; A. Gibbons, *Science* **267**, 35 (1995).
31. T. Ohta, *Theor. Popul. Biol.* **23**, 216 (1983).
32. D. A. Lawlor, J. Zemmour, P. D. Ennis, P. Parham, *Annu. Rev. Immunol.* **8**, 23 (1990).
33. J. Klein, Y. Satta, C. O'hUigin, *Annu. Rev Immunol.* **11**, 269 (1993); F. A. Ayala, *Science* **270**, 1930 (1995).
34. J. B. Walsh, *Genetics* **117**, 543 (1987); T. Ohta, *Theor. Popul. Biol.* **37**, 213 (1990); *J. Mol. Evol.* **33**, 34 (1991).
35. S. J. O'Brien, D. E. Wildt, D. Goldman, C. R. Merril, M. Bush, *Science* **221**, 459 (1983); D. I. Watkins *et al.*, *Immunogenetics* **33**, 79 (1991); J. Trowsdale, V. Groves, A. Arnason, *ibid.* **29**, 19 (1989).
36. S. J. O'Brien, *Proc. Nat. Acad. Sci. U.S.A.* **91**, 5748 (1994); R. M. May, *Nature* **374**, 309 (1995); S. J. O'Brien, *Nat. Med.* **1**, 742 (1995).
37. J. Kaufman, *Dtsch. Zool. Ges.* **81**, 131 (1988); J. Klein and C. O'hUigin, *Curr. Opin. Genet. Dev.* **3**, 923 (1993); A. L. Hughes and M. Nei, *Immunogen* **37**, 337 (1993).
38. J. Trowsdale, *Immunogenetics* **41**, 1 (1995); R. D. Campbell and J. Trowsdale, *Immunol. Today*, **14**, 349 (1993).
39. A. S. Romer, *The Procession of Life* (Weidenfeld Nicolson, London, 1968).
40. I. Stroynowski, *Annu. Rev. Immunol.* **8**, 501 (1990); Non-classical MHC antigens, G. Möller, Ed., *Immunol. Rev.* **147** (September 1995).
41. C.-R. Wang, B. E. Loveland, K. Fischer-Lindahl, *Cell* **66**, 335 (1991).
42. S. Kvist and F. Lévy, *Semin. Immunol.* **5**, 105 (1993).
43. A. Townsend and H. Bodmer, *Annu. Rev. Immunol.* **7**, 601 (1989); R. M. Zinkernagel, *Science* **271**, 173 (1996).
44. O. Rötzschke, K. Falk, O. Stevanovic, G. Jung, H. G. Rammensee, *Nature* **351**, 290 (1991); K. Falk and O. Rötzschke, *Semin. Immunol.* **5**, 81 (1993).
45. W. R. Duncan, E. K. Wakeland, J. Klein, *Nature* **281**, 603 (1979); P. W. Hedrick and G. Thomson, *Genetics* **104**, 449 (1983); F. L. Black and F. M. Salzano, *Am. J. Hum. Genet.* **33**, 894 (1981); F. L. Black, *ibid.* **36**, 1318 (1984); P. W. Hedrick and G. Thomson, *Genetics* **65**, 707 (1987); N. Takahata and M. Nei, *ibid.* **124**, 967 (1990); P. W. Hedrick, *Hum. Hered.* **40**, 213 (1990); ———, T. W. Whittam, P. Parham, *Proc. Natl. Acad. Sci. U.S.A.* **88**,

5897 (1991); T. Markow *et al.*, *Am. J. Hum. Genet.* **53**, 943 (1993).
46. A. L. Hughes and M. Nei, *Nature* **355**, 402 (1992); A. V. S. Hill, D. Kwiatkowski, A. J. McMichael, B. M. Greenwood, S. Bennett, *ibid.*, p. 403; A. V. S. Hill *et al.*, *Philos. Trans. R. Soc. London* **346**, 379 (1994).
47. K. Kato, J. A. Trapani, J. Allopenna, B. Dupont, S. Y. Yang, *J. Immunol.* **143**, 3371 (1989); D. A. Lawlor, E. Warren, P. Taylor, P. Parham, *J. Exp. Med.* **174**, 1491 (1991); J. D. Domena, W. H. Hildebrand, W. B. Bias, P. Parham, *Tissue Antigens* **42**, 156 (1993); S. N. McAdam *et al.*, *J. Immunol.* **154**, 6421 (1995).
48. D. F. Hunt *et al.*, *Science* **255**, 1261 (1992).
49. M. M. Davis and P. J. Bjorkman, *Nature* **334**, 395 (1988).
50. A. V. S. Hill *et al.*, *ibid.* **352**, 595 (1991).
51. W. Bodmer, *ibid.* **237**, 139 (1972); F. L. Black, *Science* **258**, 1739 (1992).
52. P. Parham, *Semin. Immunol.* **6**, 1 (1994).
53. T. Ohta, *J. Mol. Evol.* **33**, 34 (1991).
54. ———, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3509 (1988); *ibid.* **88**, 6716 (1991); in *HLA 1991*, K. Tsuji *et al.*, Eds. (Oxford Univ. Press, Oxford, 1992), vol. 2, pp. 20–27; in *Evolution of Life*, S. Osawa and T. Honjo, Eds. (Springer-Verlag, Tokyo, 1991), pp. 145–159.
55. T. Ohta, *Annu. Rev. Ecol. Syst.* **23**, 263 (1992).
56. Y. Satta , C. N. O'hUigin, N. Takahata, J. Klein, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7184 (1994).
57. I. Franklin and R. C. Lewontin, *Genetics* **65**, 707 (1970).
58. B. M. Fagan, *The Great Journey: The Peopling of Ancient America* (Thames Hudson, London, 1987); J. F. Hoffecker, W. R Powers, T. Goebel, *Science* **259**, 44 (1993); A. Gibbons, *ibid.*, p. 312.
59. R. C. Williams *et al.*, *Am. J. Phys. Anthropol.* **66**, 1 (1985); J. H. Greenberg, H.Turner, S. L. Segura, *Curr. Anthropol.* **27**, 477 (1986); A. Toroni *et al.*, *Genetics* **53**, 563 (1992); E. J. E. Szathmary, *Am. J. Hum. Genet.* **53**, 793 (1993).
60. D. D. Kostyu and D. B. Amos, *Tissue Antigens* **16**, 111 (1981); R. C. Williams and J. E. McAuley, *Hum. Immunol.* **33**, 39 (1992); M. L. Petzl-Erler, R. Luz, V. S. Sotomaior, *ibid.* **41**, 227 (1993).
61. M. P. Belich *et al.*, *Nature* **357**, 326 (1992); D. I. Watkins *et al.*, *ibid.*, p. 329; T. L. Garber *et al.*, *Immunogen* **42**, 19 (1995).
62. K. L. Arnett, E. J. Adams, M. L. Petzl-Erler, T. Markow, P. Parham, unpublished data.
63. K. Lienert, J. McCluskey, G. Bennett, C. Fowler, G. Russ, *Tissue Antigens* **45**, 12 (1995).
64. W. H. McNeill, *Plagues and Peoples* (Norton, New York, 1976).
65. J. V. Neel, *Am. J. Hum. Genet.* **30**, 465 (1978); F. M. Salzano and S. M. Callegari-Jacques, *South American Indians: A Case Study in Evolution* (Oxford Univ. Press, New York, 1988).
66. L. Degos and J. Dausset, *Immunogen.* **3**, 195 (1974); A. Piazza, P. Menozzi, L. L. Cavalli-Sforza, *Hum. Immunol.* **4**, 297 (1994).
67. C. G. Sibley and J. E Ahlquist, *J. Mol. Evol.* **20**, 2 (1984); M. Hasegawa, H. Kishino, T. Yano, *ibid.* **26**, 132 (1987); M. Ruvolo, T. R. Disotell, M. W. Allard, W. M. Brown, R. L. Honeycutt, *Proc. Nat. Acad. Sci. U.S.A.* **88**, 1570 (1991).
68. G. Zangenberg, M.-M. Huang, N. Arnheim, H. Erlich, *Nat. Genet.* **10**, 407 (1995).
69. L. D. Barber *et al.*, *Curr. Biol.* **5**, 179 (1995); L. D. Barber and P. Parham, unpublished data.
70. W. F. Bodmer and L. L. Cavalli-Sforza, *Genetics, Evolution and Man* (Freeman, San Francisco, 1976); L. L Cavallii-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1994).
71. F. E. Leon-S., A. Ariza-Deleon, A. Ariza-Caceido, *Hum. Immunol.* **42**, 348 (1995).
72. We thank S. G. E. Marsh of the Tissue Antigen Laboratory, Imperial Cancer Research Fund, London, for providing the tabulation of HLA sequences shown in Fig. 1. We thank W. Bodmer and J. Bodmer for suggesting this collaboration.

Population Dynamics of Immune Responses to Persistent Viruses

Martin A. Nowak and Charles R. M. Bangham

Mathematical models, which are based on a firm understanding of biological interactions, can provide nonintuitive insights into the dynamics of host responses to infectious agents and can suggest new avenues for experimentation. Here, a simple mathematical approach is developed to explore the relation between antiviral immune responses, virus load, and virus diversity. The model results are compared to data on cytotoxic T cell responses and viral diversity in infections with the human T cell leukemia virus (HTLV-1) and the human immunodeficiency virus (HIV-1).

Molecular techniques have provided fundamental insight into the fine detail of the immune system. But many biologically important questions are not primarily concerned with the molecular mechanisms of immune recognition but with the population dynamics of the immune response. Such questions usually cannot be answered by experimental methods alone but require the help of mathematical models.

These questions arise particularly in the dynamics of host-parasite interactions (*1*). In HIV infection, for example, mathematical models have been devised to describe the slow decline in the numbers of CD4 cells over many years, the interaction between HIV and other opportunistic infections, the emergence of drug-resistant viruses, and the consequences of antigenic diversity and viral evolution during single infections (*2*, *3*). In HIV and hepatitis B virus (HBV) infection, mathematical models of drug treatment dynamics have provided estimates for the turnover rates of infected cells and free virus (*4*, *5*).

The strategy of successful mathematical modeling is akin to Ockham's razor: start with the smallest number of essential assumptions and follow the implications rigorously to their logical conclusions. An elegant model can often have greater intrinsic value than an accurate one overloaded with detail. Mathematical models differ from verbal theories in giving a precise and explicit connection between assumption and conclusion. The act of formulating a model forces one to ask questions that are often overlooked (*6*). Here a simple, but general mathematical framework is presented for viral replication and immune responses. We explore the basic dynamics of virus–host cell interaction and the consequences of immune responses on virus load and antigenic diversity.

M. A. Nowak is in the Department of Zoology, University of Oxford, South Parks Road, OX1 3PS Oxford, UK. E-mail: nowak@vax.ox.ac.uk. C. R. M. Bangham is in the Imperial College School of Medicine at St. Mary's, Norfolk Place, London W2 1PG, UK. E-mail: c.bangham@ic.ac.uk

Parameters That Influence Infection Dynamics

Viruses are intracellular parasites that depend on the host cell to survive and replicate. The host cell can be damaged either directly by the virus or by immune responses to the virus; the balance of good and harm done by the antiviral immune response depends on the amount of virus present, the tissues infected, and the chronicity of the infection (*7*).

The abundance of virus—that is, the virus load—is an important determinant of the outcome of infection with many viruses: for instance, in HIV-1 and other lentivirus infections, virus load is correlated with pathogenicity, disease stage, and progression of disease (*8*, *9*); in HTLV-1, a large provirus load is associated with chronic inflammatory conditions (*10*); in HBV, the level of viremia is correlated with the risk of chronic infection (*11*); in cytomegalovirus infection, the amount of tissue damage is related to virus load (*12*); and in Lassa fever, mortality is correlated with the level of viremia (*13*).

Antibodies, cytokines, natural killer cells, and T cells are essential components of a normal immune response to a virus. But in most virus infections, cytotoxic T lymphocytes (CTLs) play a critical part in antiviral defense by attacking virus-infected cells. It is believed that they are the main host immune factor that limits the extent of virus replication in vivo and thus determines virus load. The clearest evidence for the role of these cells comes from passive transfer of immune CTLs to mice and humans (*14*). Using hu-