Primate DRB6 pseudogenes: clue to the evolutionary origin of the HLA-DR2 haplotype

Felipe Figueroa¹, Colm O’hUigin¹, Hidetoshi Inoki², and Jan Klein¹,³

¹ Max-Planck-Institut für Biologie, Abteilung Immunogenetik, Tübingen, Federal Republic of Germany
² Department of Transplantation, Tokai University School of Medicine, Isehara, Kanagawa, Japan
³ Department of Microbiology and Immunology, University of Miami School of Medicine, Miami, Florida 33101, USA

Received June 12, 1991

Abstract. The HLA-DR2 haplotype contains three β-chain encoding DRB genes and one α-chain encoding DRA gene. Of the three DRB genes, two are presumably functional (HLA-DRB1 and HLA-DRB5), whereas the third (HLA-DRBVI) is a pseudogene. A pseudogene closely related to HLA-DRBVI is present in the chimpanzee (Patr-DRB6) and in the gorilla (Gogo-DRB6). We sequenced the HLA-DRBVI and Patr-DRB6 pseudogenes (all exons and most of the introns), and compared the sequence to that of the Gogo-DRB6 gene (of which only the exon sequence is available). All three pseudogenes seem to lack exon 1 and contain other deletions responsible for shifts in the translational reading frame. At least the HLA-DRBVI pseudogene, however, seems to be transcribed nevertheless. The chimpanzee pseudogene contains two inserts in intron 2, one of which is an Alu repeat belonging to the Sb subfamily, while the other remains unidentified. These inserts are lacking in the human gene. A comparison with sequences published by other investigators revealed the presence of the HLA-DRBVI pseudogene also in the DR1 and DRw10 haplotypes. Measurements of genetic distances indicate DRB6 to be closely related to the DRB2 pseudogene and to the HLA-DRB4 functional gene. In humans, gorillas, and chimpanzees, the DRB6 pseudogene is associated with the same functional gene (DRB5) indicating that this linkage disequilibrium is at least six million years old and that DR2 is one of the oldest DR haplotypes in higher primates.

Introduction

The class II (D) region of the HLA complex, the human major histocompatibility complex (Mhc), is divided into five subregions: DP, DN, DO, DQ, and DR (Klein 1986; Carson and Trowsdale 1986). The DP and DQ subregions each contain two A loci coding for the α chain of the class II αβ heterodimer, and two B loci coding for the β chain. The DR subregion contains one A locus and one or more B loci. In the DN subregion only an A locus has been identified thus far; in the DO subregion only one B locus appears to exist. The DR subregion of the HLA complex displays two kinds of polymorphism — allelic and haplotype. Extensive allelic DR polymorphism is restricted largely to one locus in this subregion — the DRB1 locus; the DR4 locus and the extra DRB loci (if present) show only limited intraspecies variability (Dupont 1989). The haplotype polymorphism is manifested in different cis combinations of alleles at different DRB loci and in different organization and number of DRB genes borne by a single chromosome. Five haplotype groups have been identified. Haplotypes of the DR1 group (DR1, DRw10) contain two loci, DRB1 and an as yet unidentified second locus (Böhme et al. 1985). The DR2 haplotype contains three loci — two presumably functional and polymorphic loci (DRB1 and DRB5), and one pseudogene (DRBVI) (Kawai et al. 1989). Of the human DRB pseudogenes, only one (DRB2) has thus far been given an official designation. For the remaining pseudogenes, we use here tentative designations based on Roman rather than Arabic numerals]. Up until now, four alleles have been recognized officially at each of the two functional loci: HLA-DRB1*1501, *1502, *1601, *1602, HLA-DRB5*0101, *0102, *0201, *0202 (The WHO Nomenclature Committee for factors of the HLA System 1991). Haplotypes of the DR3 group (DR3, DR5, DRw6) also contain three loci (Rollini et al. 1985) — the polymorphic DRB1 locus, the DRB2 pseudogene, and the oligomorphic DRB3 locus. The DRB3 gene codes for the supertypic, serologically detectable DRw52 determinant characteristic of this haplotype group (Dupont 1989). The DRw8 haplotype, which contains only one DRB gene, is probably derived from a DR3-like haplotype by the fusion of the 5' end of the DRB1 gene with the 3' part of the DRB3
gene and deletion of the intervening segment (Gorski et al. 1984; Jonsson et al. 1989). The D~R3-derived portion of the single D~Rw8 haplotype gene encodes the Drw52 supertypic determinant. Finally, chromosomes in the D~R4 haplotype group (D~R4, D~R7, D~R9) each contain four D~R genes, the highly polymorphic and functional D~RBI, the D~RBVII and D~RBVIII pseudogenes, and the oligomorphic D~RB4 gene (Andersson et al. 1987; Spies et al. 1985).

Both the allelic and the haplotype polymorphism evolve in a trans-specific manner in which the divergence of allelic and haplotype lineages often predated speciation. Evidence for trans-specifics species evolution of D~R alleles was provided by Fan and co-workers (Fan et al. 1989); indications of trans-specifics species evolution of D~R haplotypes are emerging from the studies of Brändle and co-workers (1991) and Kasahara and co-workers (Kasahara et al. 1990, 1991). The high allelic polymorphism and the trans-specific mode of evolution complicate the interpretation of the relationships between individual D~R haplotypes, which may be important for understanding the association of certain haplotypes with disease susceptibility (Dawkins et al. 1983). To avoid these complications, we have focused our efforts on deciphering the relationships between the HLA pseudogenes which are presumably largely unaffected by the selection pressures acting on the functional D~R loci (Hughes and Nei 1989). Moreover, we have approached the question of haplotype interrelationship from a historical perspective by studying the pseudogenes in the nearest human relatives, the gorilla and the chimpanzee (Klein et al. 1991; Vincek et al. 1991). In the present communication we describe how this approach helps to understand the origin of the human D~R2, D~R1, and D~Rw10 haplotypes.

Materials and methods

Isolation of cosmid clones containing the HLA-D~RBVI pseudogene. Genomic library prepared from the DNA of the chimpanzee Pan troglodytes B-cell line Hugo, as reported elsewhere (Brändle et al. 1991), and screened with the chimpanzee cDNA probe C4-2 which covers almost the entire coding sequence of the HLA-D~RBVI gene (Fan et al. 1989). Hybridizing cosmid clones were isolated and their restriction maps constructed (for details, see Brändle et al. 1991). Hybridizations of four overlapping cosmid clones with exon-specific probes revealed the presence of a truncated D~R gene which, like the human HLA-D~RBVI, lacked exon 1 but had all the remaining exons. In preparation for sequencing, three EcoRI fragments were isolated from the cosmid 2.12.1c: a 3.7 kb long fragment encompassing exon 2 and the flanking introns; a 2.5 kb long fragment encompassing exons 3 and the flanking introns; and a 2.8 kb long fragment containing the rest of the gene including the 3'UT region. All three fragments were subcloned into the Bluescript SKII vector.

DNA sequencing. Single- or double-stranded DNAs were prepared and sequenced by the dideoxy chain-termination method (Sanger et al. 1977) using the Sequenase version 2.0 DNA sequencing kit (US Biochemicals, Cleveland, OH). The sequencing strategy for both D~R genes is depicted in Figure 1. To prepare single-stranded DNA, bacterial cultures were grown overnight in the presence of a helper phage (Stratagene). The cultures were then spun down for 20 min at 11,000 g and the DNA was precipitated from the supernatant in 4% of polyethyleneglycol (PEG). The DNA was resuspended in 400 μl of Tris EDTA (TE; 10 mM Tris-HCl, pH 8.0, 1 mM ethylenediaminetetraacetic acid, EDTA), extracted four times with phenol and once with chloroform, and precipitated with absolute ethanol. One microgram of single-stranded DNA was used in each sequencing reaction. Double-stranded DNA was prepared by the standard Trizol-lysozyme method (Davis et al. 1986). Five microliters of DNA were denatured in 0.2 M NaOH, 0.2 mM EDTA for 30 min at 37 °C, and then ethanol-precipitated in the presence of 0.3 M sodium acetate, pH 6.0. The DNA was dissolved in 7 μl of distilled water at 37 °C and sequenced. The reactions were run in 6% Accagel (National Diagnostics, Manville, NJ) at a constant 65 W current. Synthetic oligonucleotides corresponding to specific sequences along the D~R gene were synthesized in the GenAssembler Plus Synthesizer (Pharmacia LKB, Freiburg, FRG) and used as primers together with the M13 and reverse M13 primers (Table 1 and Fig. 1).

Isolation of the cosmid clone containing the Patr-D~RB6 pseudogene. Genomic library was prepared from the DNA of the chimpanzee Pan troglodytes B-cell line Hugo, as reported elsewhere (Brändle et al. 1991), and screened with the chimpanzee cDNA probe C4-2 which covers almost the entire coding sequence of the HLA-D~RBVI gene (Fan et al. 1989). Hybridizing cosmid clones were isolated and their restriction maps constructed (for details, see Brändle et al. 1991). Hybridizations of four overlapping cosmid clones with exon-specific probes revealed the presence of a truncated D~R gene which, like the human HLA-D~RBVI, lacked exon 1 but had all the remaining exons. In preparation for sequencing, three EcoRI fragments were isolated from the cosmid 2.12.1c: a 3.7 kb long fragment encompassing exon 2 and the flanking introns; a 2.5 kb long fragment encompassing exons 3 and the flanking introns; and a 2.8 kb long fragment containing the rest of the gene including the 3'UT region. All three fragments were subcloned into the Bluescript SKII vector.

Isolation of the cosmid clone containing the Patr-D~RB6 pseudogene. Genomic library prepared from the DNA of the chimpanzee Pan troglodytes B-cell line Hugo, as reported elsewhere (Brändle et al. 1991), and screened with the chimpanzee cDNA probe C4-2 which covers almost the entire coding sequence of the HLA-D~RBVI gene (Fan et al. 1989). Hybridizing cosmid clones were isolated and their restriction maps constructed (for details, see Brändle et al. 1991). Hybridizations of four overlapping cosmid clones with exon-specific probes revealed the presence of a truncated D~R gene which, like the human HLA-D~RBVI, lacked exon 1 but had all the remaining exons. In preparation for sequencing, three EcoRI fragments were isolated from the cosmid 2.12.1c: a 3.7 kb long fragment encompassing exon 2 and the flanking introns; a 2.5 kb long fragment encompassing exons 3 and the flanking introns; and a 2.8 kb long fragment containing the rest of the gene including the 3'UT region. All three fragments were subcloned into the Bluescript SKII vector.

Results

While the sequences of the D~R1*15, D~R1*16, and the D~R5 genes of the HLA-D~R2 haplotype are known (Lee et al. 1987; Wu et al. 1987; Hurley et al. 1988; Liu et al. 1988), the sequence of the D~R2 haplotype pseudogene has not been available. Since it is the pseudogene in particular that can shed light on the origin of the D~R2 haplotype and its relationship to other HLA-D~R haplotypes, we resolved to sequence it. Furthermore, during our search for the evolutionary origins of the human D~R2 haplotype, we discovered a gene in the chimpanzee Hugo which, superficially at least, resembled the HLA-D~RBVI pseudogene in that both genes seemed to lack exon 1. Expecting information on the chimpanzee gene to help in the interpretation of the phylogenetic relationships be-