Immunogenetics (1995) 42: 376–385

ORIGINAL PAPER

**Fumiko Kobari · Keisuke Sato · Benny P. Shum
Shin Tochinai · Makoto Katagiri · Teruo Ishibashi
Louis Du Pasquier · Martin F. Flajnik
Masanori Kasahara**

Exon-intron organization of *Xenopus* MHC class II β chain genes

Received: 19 May 1995 / Revised: 3 July 1995

Abstract The amphibian *Xenopus laevis* is the most primitive vertebrate in which the major histocompatibility complex (MHC) has been defined at the biochemical, functional, and molecular genetic levels. We previously described the isolation and characterization of cDNA clones encoding *X. laevis* MHC class II β chains. In the present study, genomic clones encoding class II β chains were isolated from *X. laevis* homozygous for the MHC *f* haplotype. Three class II β chain genes, designated *Xela-DAB, Xela-DBB*, and *Xela-DCB*, were identified. Sequence analysis of these genes showed that *Xela-DBB* and *Xela-DCB* correspond to the previously characterized cDNA clones F3 and F8, respectively, whereas *Xela-DAB* encodes a third, hitherto unidentified class II β chain of the MHC *f* haplotype. As a representative of *X. laevis* class II β chain genes, the *Xela-DAB* gene underwent detailed structural analysis. In addition, the nucleotide sequence of *Xela-DABf* cDNA clones was determined. The *Xela-DAB* gene is made up of at least six exons, with an exon-intron organization similar to that of a typical mammalian class II β chain gene. The 5′-flanking region of the *Xela-DAB* gene contains transcriptional control elements known as X1, X2, and Y, but lacks typical TATA or CCAAT boxes. A notable feature of the *X. laevis* class II β chain genes is that the sizes of the introns are larger than those of their mammalian counterparts. As assessed by northern blot analysis, the three class II β chain genes had similar expression patterns, with the highest level of transcription detected in the intestine. Identification of the *Xela-DAB, -DBB*, and *-DCB* genes is consistent with our previous observations, which suggested that the MHC of the tetraploid frog *X. laevis* is diploidized at the genomic level and contains three class II β chain genes per haplotype that cross-hybridize to one another under reduced stringency conditions.

The nucleotide sequence data reported in this paper have been submitted to the GSDB/DDBJ/EMBL/NCBI nucleotide sequence databases and have been assigned the accession numbers D50035-D50043

F. Kobari · T. Ishibashi · M. Kasahara (✉)
Department of Biochemistry, Hokkaido University School of Medicine, N-15, W-7, Sapporo 060, Japan

F. Kobari · S. Tochinai
Division of Biological Sciences, Hokkaido University Graduate School of Science, Sapporo 060, Japan

K. Sato · M. Katagiri
Department of Pathology, Asahikawa Medical College, Asahikawa 078-11, Japan

B. P. Shum[1] · M. F. Flajnik
Department of Microbiology and Immunology, University of Miami School of Medicine, Miami, FL 33101, USA

L. Du Pasquier
Basel Institute for Immunology, CH-4005, Basel, Switzerland

Present address:
[1] Department of Cell Biology, Stanford University School of Medicine, Stanford, CA, USA

Introduction

The major histocompatibility complex (MHC) of the amphibian *Xenopus laevis* is one of the best characterized non-mammalian MHCs. Our previous biochemical studies showed that *X. laevis* has both MHC class I and class II molecules and that the class I α and class II α/β chains are highly polymorphic (Flajnik et al. 1984; Kaufman et al. 1985a, 1985b). Furthermore, genetic analysis using MHC homozygous strains provided evidence that the class I α chain and both chains of class II molecules are encoded by a single chromosomal region that controls mixed lymphocyte culture reactions, T-B-cell collaboration, cell-mediated lympholysis, and graft rejection (for reviews, see Du Pasquier et al. 1989; Flajnik and Du Pasquier 1990). These results suggest that the essential properties of the MHC, as revealed from the analysis of the mammalian MHC (Klein 1986), were already established when the amphibian emerged ~350 million years ago.

The MHC of *X. laevis* provides not only important information about the phylogeny of adaptive immunity, but has two unique features that could serve as models of immunological interest. First, despite the fact that *X. laevis*

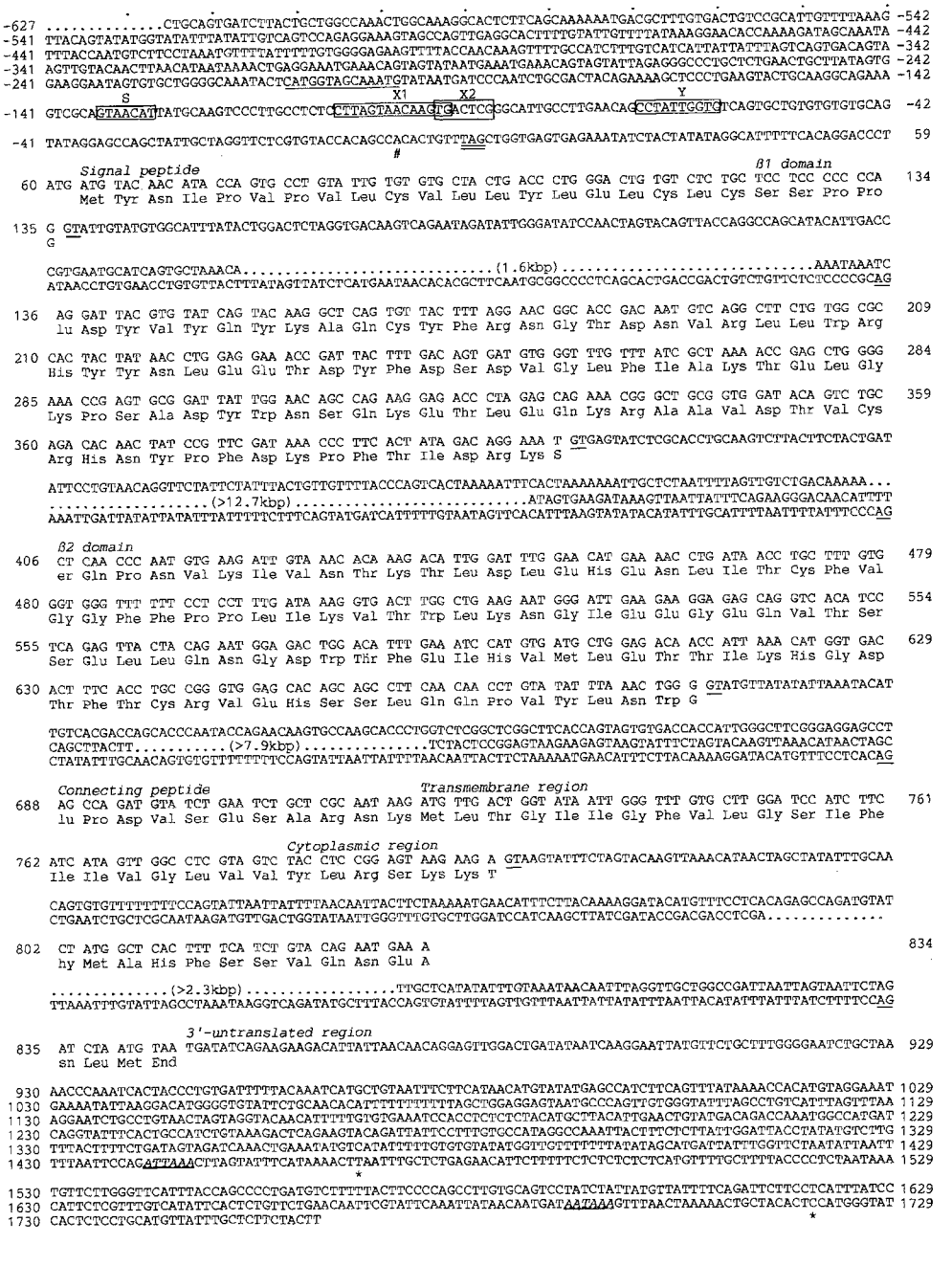

Fig. 1 Nucleotide sequence of the *Xela-DABf* gene. The deduced amino acid sequence is shown *below* the nucleotide sequence in a standard *three-letter code*. Only the coding region and the 5′-flanking region are numbered. An in frame stop codon preceding the translation initiation codon is *doubly underlined*. Putative polyadenylation signals (ATTAAA and AATAAA) are *italicized* and *underlined*. Two transcription termination sites identified by the 3′-RACE technique are indicated by *asterisks*. The major transcription initiation site defined by the 5′-RACE technique is indicated by #. Negative numbers are given for nucleotides located upstream of the major transcription initiation site. Putative regulatory elements in the 5′-flanking region are boxed. A second X1 box-like sequence located upstream of the X1 is *underlined*. The sequence shown as exon 5 (encoding the cytoplasmic region) was obtained from the *Xela-DAB* cDNA clone. The conserved dinucleotides (GT/AG) at the splice donor/acceptor sites (Mount 1982) are *underlined*. The GSDB/DDBJ/EMBL/NCBI accession numbers for the *Xela-DABf* genomic sequences are D50037 (exon 1 and adjacent sequences), D50038 (exon 2 and adjacent intron sequences), D50042 (exon 3 and adjacent intron sequences), D50039 (exon 4 and adjacent intron sequences), and D50041 (the last exon and adjacent sequences). The accession numbers for the *Xela-DABf* cDNA sequences, which can be deduced from the exonic sequences shown in this Figure, are D50036 (for the cDNA clone ending at nucleotide position 1466) and D50035 (for the cDNA clone ending at nucleotide position 1720)

is a tetraploid species that arose ~30 million years ago as a result of allopolyploidization between two diploid species (Bisbee et al. 1977), it expresses only one diploid set of the MHC. Because many of the genes duplicated in the allopolyploidization event are still retained and expressed in *X. laevis* (reviewed in Kobel and Du Pasquier 1986), the diploidization of the MHC suggests that a selective pressure is at work against an increase in the number of functional MHC genes. Second, the tissue distribution of MHC class I and class II molecules changes dramatically before and after metamorphosis (Flajnik et al. 1986). In particular, class I molecules start to appear ubiquitously first at metamorphosis, raising the possibility that the emergence of class I molecules might serve as a signal for the tissue destruction and reorganization that accompany metamorphosis.

To understand the molecular basis of the above-mentioned unique features of the *Xenopus* MHC and to gain further insights into the evolution of the MHC, we have initiated the molecular genetic characterization of the *Xenopus* MHC. In our preceeding papers, we described isolation of cDNA clones encoding MHC class I α and class II β chains of *X. laevis* (Flajnik et al. 1991, 1993; Sato et al. 1993; Shum et al. 1993). These studies showed that the MHC of *X. laevis* contains only one class I α and most likely three class II β chain genes per haplotype. In the present study, we describe the isolation and characterization of genomic clones encoding class II β chains of *X. laevis*. In

accordance with a proposal by Klein and co-workers (1990), we named the three class II β chain genes described in this paper *Xela-DAB, Xela-DBB*, and *Xela-DCB*.

Materials and methods

Isolation and characterization of genomic clones encoding X. laevis MHC class II β chains

Two genomic libraries were used to isolate clones encoding *X. laevis* MHC class II β chains. One was a lambda GEM-11 library made from *X. laevis* homozygous for the MHC *f* haplotype. This library was constructed using the *Xho* I half-site arms cloning system (Promega, Madison, WI) following the instructions of the manufacturer. The other was an EMBL3 genomic library prepared from MHC *f/f* homozygous *X. laevis* (Schwager et al. 1988). Screening of the libraries was performed according to the standard procedure (Sambrook et al. 1989) at 42 °C in a solution containing 30% formamide, 1 M NaCl, 10 × Denhardt's solution, 50 mM Tris-HCl (pH 7.5), 1% $Na_4P_2O_7$, 1% sodium dodecyl sulfate, and 150 μg/ml of sheared and denatured salmon sperm DNA. The probes used for screening were full-length cDNA clones encoding *X. laevis* class II β chains (Sato et al. 1993). Restriction maps of the positive λ clones were constructed using the method of Rackwitz and co-workers (1984).

Isolation of Xela-DAB cDNA by polymerase chain reaction (PCR)

Total spleen RNA isolated from *X. laevis* homozygous for the MHC *f* haplotype was reverse-transcribed using a (dT)17 adapter primer (5′-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT-3′) as previously described (Kasahara et al. 1991). The cDNA thus obtained was subjected to PCR according to the standard procedure (Saiki et al. 1988) using Ampli*Taq* DNA polymerase (Perkin Elmer Cetus, Norwalk, CT). The coding region of the *Xela-DAB* cDNA was obtained using a primer (5′-GTGTGTGCTACTGACCCTGG-3′) corresponding to the sense strand of the exon encoding the signal peptide (nucleotides 89–108 in Figure 1) and a primer (5′-CCAACTCCTGTTGTTAATA-3′), the sequence of which was inversely complementary to that of the exon encoding the 3′ untranslated (UT) region (nucleotides 865–883 in Fig. 1). The remaining 3′-portion of the *Xela-DAB* cDNA was amplified by the 3′-RACE (rapid amplification of the 3′-end) technique (Frohman et al. 1988) using an adapter primer (5′-GACTC-GAGTCGACATCG-3′) and an oligonucleotide (5′-CTATGGCT-CACTTTTCATCT-3′) specific for the cytoplasmic region of the *Xela-DAB* gene (nucleotides 802–821 in Fig. 1). The PCR reaction mixtures (50 μl) contained 3 μl of the cDNA, 200 μM dNTPs, 0.4 μM of primers, 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 2.5 mM $MgCl_2$, gelatin (0.2 mg/ml), and 2.5 units of Ampli*Taq* DNA polymerase. The condition of PCR was 30 cycles of 1 min denaturation at 94 °C, 2 min annealing at 54 °C, and 2 min extension at 72 °C, followed by 10 min incubation at 72 °C.

Determination of transcription initiation sites

Transcription initiation sites of the *Xela-DAB* gene were determined by the 5′-RACE (rapid amplification of the 5′ end) technique (Frohman et al. 1988) as described by Kandil and co-workers (1995), with minor modifications. Briefly, total spleen RNA (10 μg) isolated from *X. laevis* homozygous for the MHC *f* haplotype was reverse-transcribed using a primer (5′-CAATCCCATTCTTCAGCCA-3′) inversely complementary to nucleotides 513–531 of the β2 domain (Fig. 1). After tailing with dATP, the cDNA was amplified by PCR using the primer mentioned above, the (dT)17 adapter primer, and the adapter primer (30 cycles of 1 min at 94 °C, 2 min at 52 °C, and 2 min at 72 °C). The second round of PCR was performed with 1/25 of the material from the first PCR using the adapter primer and a nested *Xela-DAB*-specific primer (5′-ATCGGTTTCCTCCAGGTTATAGTAGTG-3′) that was inversely complementary to nucleotides 210–236 of the β1 domain (Fig. 1). The condition of PCR was the same as in the first round of PCR.

Inverse PCR

Exon 4 and the adjacent introns of the *Xela-DAB* gene were cloned by inverse PCR as described by Ochman and co-workers (1990). Briefly, genomic DNA isolated from *X. laevis* homozygous for the MHC *f* haplotype was digested with *Taq* I and circularized by self-ligation. PCR was performed using 0.5 μg of the circularized DNA as a template. The sequences of the primers were 5′-ATCA-TAGTTGGCCTCGTAGTC-3′ (sense) and 5′-GATG-GATCCAAGCACAAACCCA-3′ (antisense). These primers were designed to anneal to a putative exon 4 sequence in the opposite, outward orientations (nucleotides 762–782 and 737–758 in Figure 1). The condition of PCR was 50 cycles of 1 min denaturation at 94 °C, 2 min annealing at 52 °C, and 2 min extension at 72 °C, followed by 10 min incubation at 72 °C.

DNA sequencing

Double-stranded plasmid DNA was sequenced by the chain-termination method (Sanger et al. 1977) using the Sequenase Version 2.0 sequencing kit (US Biochemicals, Cleveland, OH).

Northern blot hybridization

Northern blot hybridization was performed as previously described (Kasahara et al. 1987).

Results

Isolation of genomic clones encoding X. laevis MHC class II β chains

We previously isolated two distinct class II β chain cDNA clones, designated F3 and F8, from *X. laevis* homozygous for the MHC *f* haplotype (Sato et al. 1993). In addition, a class II β chain cDNA clone, designated T4, was identified from a thymus/spleen library constructed from outbred *X. laevis* (Sato et al. 1993). Sequence comparison of these clones suggested that T4 was derived from a gene distinct from that corresponding to F3 or F8. We therefore screened the *X. laevis f/f* genomic libraries using a mixture of cDNA clones F3, F8, and T4 as probes. Screening of the libraries was performed under reduced stringency (30% formamide at 42 °C) so that one could isolate all class II β chain genes with sufficient sequence similarity to the probes. Seven positive clones were identified (Fig. 2). Restriction mapping of these clones showed that only clones λ2-4 and λ2-1 had an overlap. Clone λ1-1 contained exons encoding the signal peptide and the β1 domain, and clone λ2-6 contained an exon encoding the β2 domain. The sequences of the exons in clones λ1-1 and λ2-6 were identical to those of the corresponding regions of clone F3. Thus, we concluded that

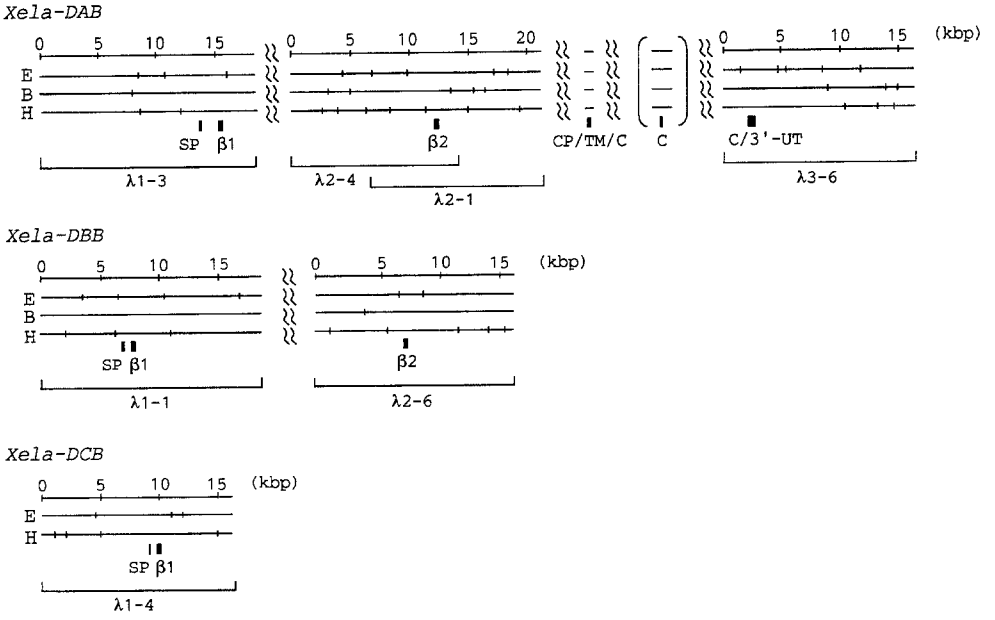

Fig. 2 Exon-intron organization of *X. laevis* MHC class II β chain genes. Exons are shown as *solid boxes*. *SP* Signal peptide; *β1* β1 domain; *β2* β2 domain; *CP* Connecting peptide; *TM* Transmembrane region; *C* Cytoplasmic region; *3′UT* 3′UT region. Restriction enzymes used for the construction of the map are *Eco* RI, *Bam* HI, *Hin* dIII (abbreviated as E, B, and H, respectively). Individual λ clones are indicated by *horizontal bars* beneath the restriction map. The exon of the *Xela-DAB* gene encoding the connecting peptide, the transmembrane region, and a part of the cytoplasmic region was isolated by inverse PCR. The genomic fragment corresponding to the 33 bp cytoplasmic region of the *Xela-DAB* cDNA could not be cloned. Assuming that it is encoded by a single exon, it is shown in parentheses. The exon encoding the 3′UT region of the *Xela-DAB* gene is drawn in length corresponding to the shorter transcript ending at nucleotide position 1466 (Fig. 1)

λ1-1 and λ2-6 were derived from a single gene, and designated this gene *Xela-DBB*. Clone λ1-4 contained exons encoding the signal peptide and the β1 domain (Fig. 2). The sequences of these exons matched those of the corresponding regions of clone F8. This gene was named *Xela-DCB*. The sequences of the exons found in clones λ1-3, λ2-4 (λ2-1), and λ3-6 (Fig. 1) were most closely related to those of the corresponding regions of clone T4. This result suggested that these λ clones might be derived from a single gene, presumably allelic to the gene from which the cDNA clone T4 was derived.

To test this possibility, we synthesized two PCR primers based on the exonic sequences of λ1-3 and λ3-6. One of the primers (5′-GTGTGTGCTACTGACCCTGG-3′) had a sequence corresponding to the sense strand of the exon encoding the signal peptide (nucleotides 89–108 in Figure 1). The sequence of the other primer (5′-CCAACTCCTGTTGTTAATA-3′) was inversely complementary to that of the exon encoding the 3′UT region (nucleotides 865–883 in Figure 1). PCR was then performed using cDNA synthesized from the spleen RNA of an MHC *f/f* individual as a template. A single DNA fragment of 795 base pairs (bp) was amplified. The sequence of this fragment was identical to the exonic sequences of λ1-3, λ2-4 (λ2-1), and λ3-6, indicating that these λ clones are derived from a single gene. We named this class II β chain gene *Xela-DAB*.

Isolation of Xela-DAB^f cDNA clones

Xela-DBB and *Xela-DCB* correspond to the previously characterized cDNA clones F3 and F8, respectively. In contrast, *Xela-DAB* is a third class II β chain gene of the MHC *f* haplotype, for which no corresponding cDNA clones were identified in our previous study (Sato et al. 1993). We therefore decided to determine the complete cDNA sequence and the structure of the *Xela-DAB^f* gene.

The 5′ and 3′ ends of the *Xela-DAB^f* cDNA not covered by the 795 bp fragment mentioned above were cloned by 5′- and 3′-RACE, respectively, using spleen cDNA prepared from *X. laevis* homozygous for the *f* haplotype as a template. Two cDNA fragments of ~680 and ~920 bp were obtained by 3′-RACE using the adapter primer and the oligonucleotide (5′-CTATGGCTCACTTTTCATCT-3′) specific for the cytoplasmic region of the *Xela-DAB^f* gene. Sequence analysis of these fragments showed that the *Xela-DAB^f* gene has two transcription termination sites, one located at nucleotide 1466, and the other at 1720 (indicated by asterisks in Figure 1). In both cases, putative polyadenylation signals were found ~25 bp upstream from the transcription termination sites. The 5′ end of the *Xela-DAB^f* cDNA was cloned and sequenced as described below (see the section on the determination of transcription initiation sites). The composite, full-length cDNA sequence thus obtained was identical to the exonic sequence of the *Xela-DAB^f* gene shown in Figure 1.

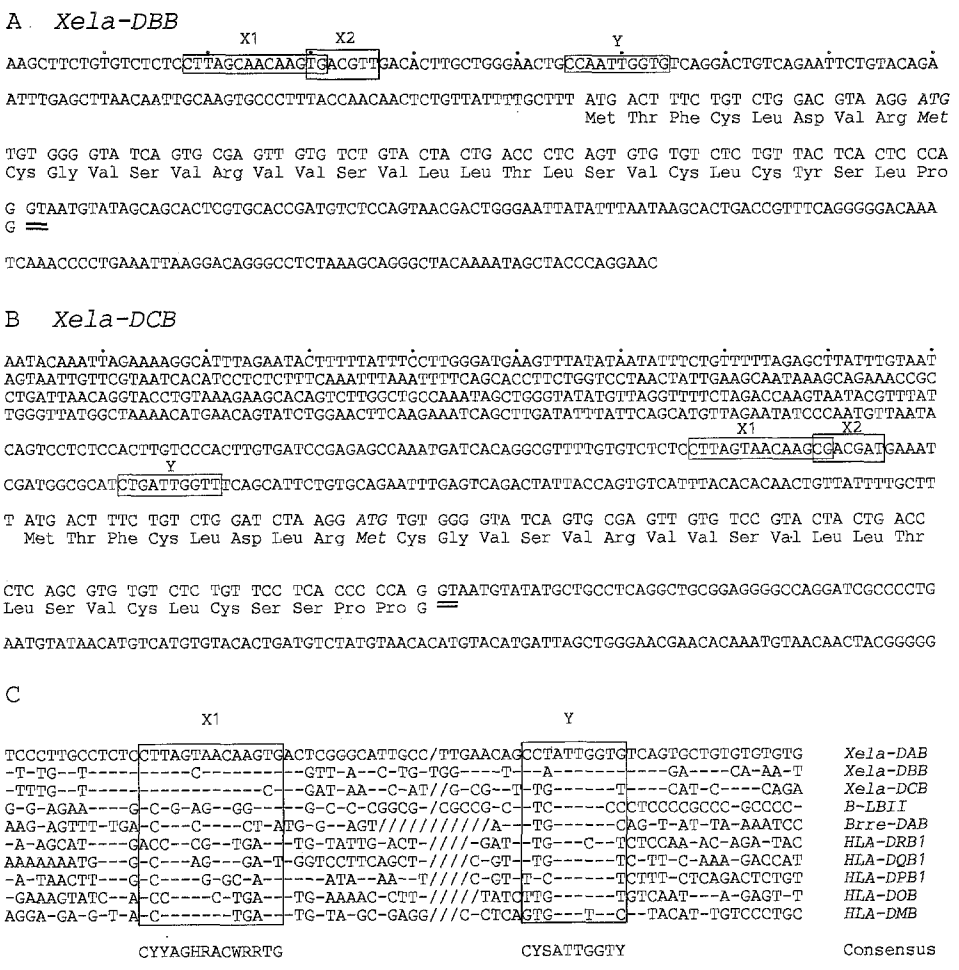

Fig. 3A–C Nucleotide and deduced amino acid sequences of the 5′ end regions of the **A, B** *Xela-DBB^f* and *Xela-DCB^f* genes and the alignment of the promoter sequences of the MHC class II β chain genes isolated from **C** *X. laevis*, chicken, zebrafish, and humans. The GSDB/DDBJ/EMBL/NCBI accession numbers for the sequences shown in **A** and **B** are D50043 and D50040, respectively. The second methionine residue, which might be the actual translation initiation site, is *italicized*. The GT dinucleotide at the splice donor site is *doubly underlined*. The sources of the sequences in **C** were as follows: *Xela-DAB*, *Xela-DBB*, and *Xela-DCB* from *X. laevis* (this work); *B-LBII* from the chicken (Zoorob et al. 1990); *Brre-DAB* from the zebrafish (Sültmann et al. 1994); *HLA-DRB1* (Andersson et al. 1987), *HLA-DQB1* (Jonsson et al. 1987), *HLA-DPB1* (Kelly and Trowsdale 1985), *HLA-DOB* (Servenius et al. 1987), and *HLA-DMB* (Radley et al. 1994) from the human. The *dash* (-) and the *oblique stroke* (/) indicate identity to the top sequence and absence of residues, respectively

Exon-intron organization of the Xela-DAB^f gene

Because neither clone λ2-1 nor λ3-6 contained sequences corresponding to the connecting peptide, transmembrane region, or the cytoplasmic region (Fig. 2), inverse PCR was used to isolate these sequences. A pair of primers described in Materials and methods led to amplification of the 114 bp exonic sequence (exon 4) and the adjacent intronic sequence (Fig. 1). However, repeated attempts to isolate the 33 bp sequence encoding the cytoplasmic region were unsuccessful. Thus, it was not possible to rule out the possibility that this 33 bp sequence was split into more than two exons. The precise exon-intron boundaries of the *Xela-DAB^f* gene were established by comparing the cDNA and genomic sequences (Fig. 1). This analysis revealed that the exons of the *Xela-DAB^f* gene encode the following domains or regions: exon 1, the 5′UT region, the signal peptide, and the first four amino acids of the β1 domain; exon 2, the β1 domain; exon 3, the β2 domain; exon 4, the connecting peptide, the transmembrane region, and the first six amino acids of the cytoplasmic region; exon 5, the cytoplasmic region; and the last exon, the last three amino acids of the cytoplasmic region and the entire 3′UT region. Thus, the overall exon-intron organization of the *Xela-DAB* gene was similar to that of a typical mammalian class II β chain gene (Radley et al. 1994; Trowsdale 1995). Furthermore, as in mammalian MHC class I and class II genes (Hood et al. 1983), RNA splicing always took place between the first and second bases of the junctional codon (Fig. 1).

Determination of transcription initiation sites of the Xela-DAB gene

To determine the transcription initiation sites of the *Xela-DAB* gene, the spleen cDNA prepared from *X. laevis* homozygous for the *f* haplotype was subjected to 5′-RACE using the adapter primer and the *Xela-DAB*-specific primer. This resulted in amplification of 250∼270 bp fragments. These fragments were cloned into the pBluescript SKII$^+$ vector, and the sequences of ten independent clones were determined. Three of the clones contained the inserts that started at the adenosine residue located 62 bp upstream from the translation initiation site (indicated by # in Figure 1). This site was therefore judged to be the major transcription initiation site. The remaining seven clones had inserts with various sizes and started at nucleotides –21, –11, –9, +2, +11, +43, and +44.

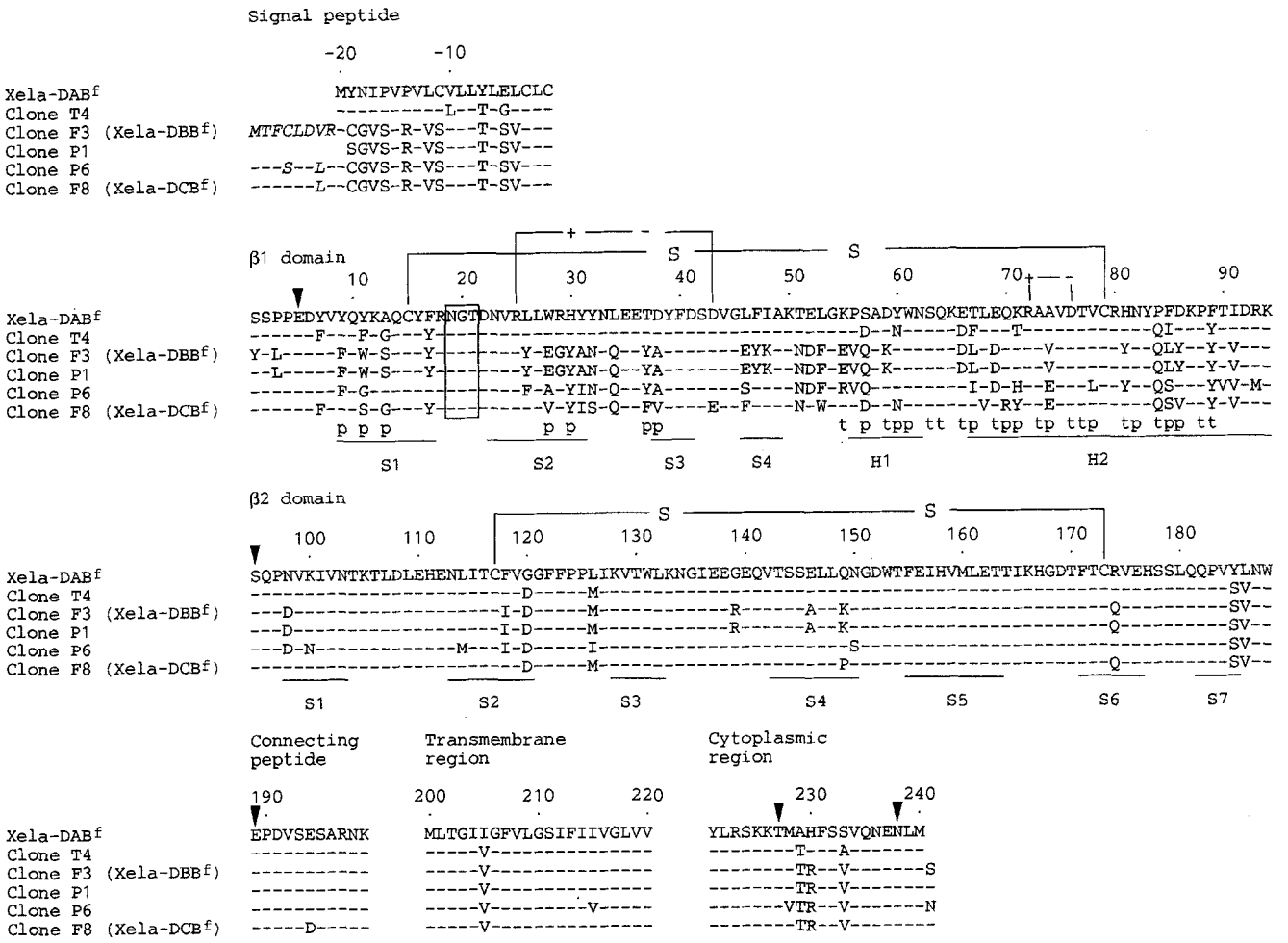

Fig. 4 Amino acid sequence comparison of the Xela-DAB^f molecule with previously reported *X. laevis* class II β chains. The sequences of clone F3 (Xela-DBB^f), clone F8 (Xela-DCB^f), and clones T4, P1, and P6 (cDNA clones isolated from outbred *X. laevis*) were taken from Sato and co-workers (1993). Sequence alignment was performed manually to maximize sequence similarity and to avoid gaps in the predicted β strand regions. A *dash* (-) indicates identity to the top sequence. The predicted location of β strands (S1–S4 in the β1 domain and S1–S7 in the β2 domain) and α-helices is based on the crystal structure of HLA-DR1 molecules (Brown et a. 1993). Residues thought to interact with peptides and the T-cell receptor are indicated by "p" and "t", respectively. A potential *N*-linked glycosylation site is *boxed*. Exon-intron boundaries of the *Xela-DAB^f* gene are indicated by *arrowheads*. The first eight residues of the signal peptide (–28 to –21), which might not be translated, are shown in *italics*

Structural features of the promoter region of the X. laevis MHC class II β chain genes

The promoter region of the *Xela-DAB^f* gene did not contain typical TATA or CCAAT boxes at expected positions. However, the regulatory sequences known as the X1, X2, and Y boxes (Benoist and Mathis 1990) were found 94, 89, and 62 bp upstream from the major transcription initiation site, respectively (Fig. 1). A sequence related to the X1 box was also found 200 bp upstream from the major transcription initiation site. The *Xela-DAB^f* gene had a pyrimidine-rich sequence immediately upstream from the X1 box and a sequence with weak similarity to the S box. The tissue distribution of class II molecules changes markedly at metamorphosis in which thyroid hormone plays an indispensable role (Flajnik and Du Pasquier 1990). However, the thyroid hormone responsive element (TCAGGTCANNNTGACCTGA; Beato 1989) was not present in the promoter region of the *Xela-DAB^f* gene. Analysis of the promoter regions of the *Xela-DBB* and *Xela-DCB* genes revealed that they also have X1, X2 and Y boxes (Fig. 3). With the exception of these boxes, the promoter regions of the three *X. laevis* class II β chain genes displayed only a low level of sequence similarity. Interspecies comparison of the promoter sequences showed that the X1 and Y boxes, in particular the Y box, are highly conserved in the class II β chain genes (Fig. 3).

Amino acid sequence comparison of X. laevis MHC class II β chains

The deduced amino acid sequence of the Xela-DAB molecule (Fig. 4) isolated from *X. laevis* with the MHC *f* haplotype (Xela-DAB^f) was then compared with those of the previously reported class II β chains (Sato et al. 1993). The Xela-DAB^f molecule was most similar to the T4

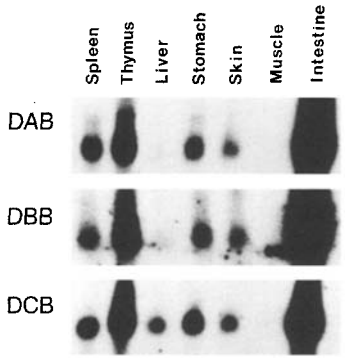

Fig. 5 Tissue distribution of *Xela-DAB, -DBB*, and *-DCB* transcripts. Total cellular RNA (10 μg) isolated from the indicated organs of *X. laevis* homozygous for the *f* haplotype was subjected to northern blot analysis. Hybridization was performed in a solution containing 50% formamide using the 3′UT region segments of *Xela-DAB*^{f}, *-DBB*^{f}, and *-DCB*^{f} cDNAs as probes. The amount of RNA loaded on each lane was similar as judged by the intensity of ethidium bromide staining (data not shown). Note that the RNAs were extracted from whole organs without enrichment for any particular populations of cells. The approximate sizes of the transcripts were 1700 ~ 1900 bases for *Xela-DAB*, 1800 bases for *Xela-DBB*, and 1600 bases for *Xela-DCB*

molecule isolated from outbred *X. laevis* (91.5% overall amino acid sequence identity). This was followed by F8 (DCB^f; 81.9%), P6 (78.8%), P1 (78.1%), and F3 (DBB^f; 76.9%). This result indicates that the T4 molecule is most likely encoded by an allele of the *Xela-DAB* locus. The β1 domains of the Xela-DAB^f and T4 molecules differ by 12 amino acids (Fig. 4). Seven of the substitutions are located at residues presumed to interact with peptides, and three at residues presumed to interact with the T-cell receptor. This distribution pattern of polymorphic residues suggests stronlgy that the *Xela-DAB* gene encodes a functional class II β chain that plays a major role in antigen presentation. Our previous analysis of the T4, F3, P1, P6, and F8 cDNA sequences showed that nonsynonymous substitutions occur more frequently than synonymous substitutions at presumed peptide-binding sites (Sato et al. 1993). Thus, the three class II β chain genes of *X. laevis* are equipped with essential features of functional MHC genes.

It is notable that, unlike F3, P6, and F8, the Xela-DAB^f molecule does not have the ATG codon at position –28 (Figs. 1, 4). Instead, Xela-DAB^f has the ATG codon at position –21 (Fig. 1). However, this ATG codon is absent in T4 (Sato et al. 1993) and is not embedded in Kozak's consensus sequence (Kozak 1984). Therefore, the translation most likely starts at position –20 in the Xela-DAB^f molecule. The methionine at position –20, rather than at position –28, might also serve as a translation initiation site in other class II β chains, because the former methionine is shared in all class II β chains.

Expression of Xela-DAB, Xela-DBB, and Xela-DCB genes in X. laevis tissues

Northern blot analysis was performed to examine expression of the *Xela-DAB, Xela-DBB*, and *Xela-DCB* genes in *Xenopus* tissues (Fig. 5). To monitor the expression of individual genes separately, hybridization was performed under stringent conditions using the 3′UT region probes. Figure 5 shows that the three class II β chain genes have essentially similar expression patterns. The transcripts were most abundant in the intestine and barely detectable in the muscle. In the liver, the *Xela-DCB* transcript was more abundant than the *Xela-DAB* or *Xela-DBB* transcript.

Discussion

Accumulated evidence indicates that the number of polymorphic, abundantly expressed, and hence functionally important MHC genes shows little variation in phylogenetically diverse species ranging from humans to fish (reviewed in Kasahara et al. 1995). Thus, humans, mice, chickens (Zoorob et al. 1993), and zebrafish (Sültmann et al. 1994) have one to three polymorphic class II β chain genes (per haplotype) thought to play a major role in antigen presentation. Because an individual MHC molecule can bind only a limited set of peptides with certain structural motifs (Rammensee et al. 1993), one might think that acquisition of a large number of MHC molecules with distinct peptide binding capabilities is advantageous to the organism. However, an increase in the number of MHC molecules enlarges not only the repertoire of foreign peptides presented to T cells, but also that of self peptides, and consequently, the number of T cells deleted or inactivated in the thymus. Because a significant proportion of T cells positively selected by one MHC molecule is negatively selected by another MHC molecule, the presence of too many functional MHC molecules has a net effect of decreasing the T-cell repertoire size (Lawlor et al. 1990; Takahata 1995). Presumably, the optimal number of functional MHC class II β chain genes that maximizes the size of the class II-restricted T-cell repertoire lies in the range of one to three per haplotype.

The amphibian *Xenopus* offers a unique opportunity to illustrate the constraint imposed on the number of functionally important MHC genes. Because the genus *Xenopus* speciates by allopolyploidization, there exist polyploid species with *4n, 8n*, and *12n* chromosomes. Evidence obtained from electrophoretic analysis of proteins and gene cloning indicates that the polyploid species of *Xenopus* including *X. laevis* (*4n* chromosomes) retain and express many of the duplicated genes (reviewed in Kobel and Du Pasquier 1986). In sharp contrast, they express only one diploid set of MHC genes, with the exception of *X. ruwenzoriensis* (*12n* chromosomes) that is thought to be a recent polyploid (Du Pasquier et al. 1977). Because laboratory-made polyploids express MHC genes of each constituting species codominantly (Du Pasquier et al. 1977), diploidization of the MHC in polyploid species of *Xenopus* is a process accomplished over evolutionary time. It is therefore reasonable to assume that diploidization of the *Xenopus* MHC is a consequence of the selection pressure that acts against an increase in the number of functional

MHC genes. We previously showed by Southern blot analysis that the genome of *X. laevis* contains only three class II β chain genes per haplotype which cross-hybridize with the cDNA clone F3 and that the three genes segregate with serologically defined MHC haplotypes (Sato et al. 1993). This result suggested that the diploidization of the MHC in *X. laevis* was achieved at the genomic rather than the transcriptional level. Our present study confirms this suggestion and demonstrates that the genome of *X. laevis* indeed contains the three predicted class II β chain genes. Failure to find more than three genes indicates that the second, inactivated set of MHC class II β chain genes was either deleted or that their sequences diverged too far to be detected by DNA hybridization with the expressed class II β chain genes.

The overall exon-intron organization of the *Xela-DAB* gene (Fig. 1) is remarkably similar to that of typical mammalian class II β chain genes (reviewed in Trowsdale 1995). Comparison of the *Xela-DAB* gene and the *HLA-DRB1* gene (Andersson et al. 1987) indicates that the relative positions of introns are perfectly conserved. Furthermore, the promoter regions of the *Xela-DAB, -DBB*, and *-DCB* genes contain the X1, X2, and Y boxes (Figs. 1, 3). Similar observations have been made for the class II β chain genes of chickens (Zoorob et al. 1990), cichlid fishes (Ono et al. 1993a), and zebrafish (Sültmann et al. 1994). Therefore, the present study confirms and extends the previous conclusion that the exon-intron organizations as well as the regulatory sequences of MHC genes are highly conserved in evolution (Klein et al. 1993). The X1 and Y boxes of the *X. laevis* class II β chain genes are 21–23 bp apart (Fig. 1). This distance, which nearly corresponds to two helical turns of DNA, is in good agreement with that (18–22 bp) observed in the mammalian and avian class II β chain genes (Fig. 3). A notable exception is the zebrafish class II β chain gene *Brre-DAB* (Sültmann et al. 1994), in which the distance is only 12 bp. However, this spacing almost corresponds to one helical turn. It is therefore conceivable that, as in the other class II β chain genes, the X1 and Y boxes of the *Brre-DAB* gene can form the specific stereospecific alignments required to interact with transcription factors (Vilen et al. 1990).

A notable feature of the *X. laevis* class II β chain genes is that they have long introns (Fig. 2). Thus, intron 2 is at least 12.7 kilobases (kb) long in the *Xela-DAB* gene, and at least 17.0 kb long in the *Xela-DBB* gene. Furthermore, both genes have intron 3 that exceeds 9.0 kb. In contrast, the entire length of most mammalian MHC genes is less than 10 kb (Radley et al. 1994). To estimate the intron sizes of the *Xela-DAB* gene more accurately and to clone intron sequences between unlinked exons, we performed long-distance PCR (Barnes 1994) using primers located on adjacent exons. However, no specific amplification products were obtained (F. Kobari and M. Kasahara, unpublished data). It therefore appears that all introns other than intron 1 are quite long in the *Xela-DAB* gene. The haploid genome size of *X. laevis* is similar to that of humans (Thiébaud and Fischberg 1977). Also, other *X. laevis* genes, such as α and β globin genes (Patient et al. 1980), have introns with sizes comparable to those of their mammalian counterparts. These observations suggest that the presence of long introns is unlikely to be a general feature of *X. laevis* genes. As shown in Figure 2, none of the three class II β chain genes could be cloned in its entirety. One obvious reason for this is that the introns are unusually long. A small exon such as that encoding the transmembrane or cytoplasmic region is unlikely to give strong hybridization signals if it is present alone in a λ clone. Another plausible explanation is that the *Xenopus* genome might contain many unclonable sequence stretches. This possibility is suggested by the fact that, when cloned into cosmid vectors, *Xenopus* genomic DNA is quite susceptible to rearrangements and deletions (B. P. Shum and M. F. Flajnik, unpublished data). The *Xenopus* genome might contain unclonable sequences on the average every 50 kb.

Our results indicate that the MHC of *X. laevis* contains three almost equally transcribed class II β chain genes (Fig. 5), and no pseudogenes with sufficient sequence similarity to the *Xela-DAB, -DBB*, or *-DCB* gene (Fig. 2). This is in contrast to the fact that the MHCs of humans, mice, and zebrafish harbor class II pseudogenes. Our previous study showed that the genome of *X. laevis* contains a large number of non-MHC-linked class Ib genes and pseudogenes (Flajnik et al. 1993). However, we were unable to detect any class I pseudogenes encoded by the MHC (Shum et al. 1993). The significance of the apparent absence of class I α and class II β chain pseudogenes in the *Xenopus* MHC is not clear.

The three class II β chain genes of *X. laevis* are transcribed most abundantly in the intestine (Fig. 5). Although class II genes are expressed also in the intestine of mammals (reviewed in Klein 1986) and fish (Ono et al. 1993b; Sültmann et al. 1994), the expression level is not as high as in *X. laevis*. Preliminary results indicate that the organ in which the class Ia gene of *X. laevis* is transcribed at the highest level is the intestine (L. Salter-Cid and M. F. Flajnik, unpublished data). These observations suggest that the mucosal immune system might play a particularly important role in *X. laevis*.

In adult *X. laevis*, class II molecules are expressed on the surface of all resting T and B cells. In contrast, tadpoles express class II molecules on the surface of B, but not T, cells (Flajnik and Du Pasquier 1990). The genomic clones described in this study should facilitate the identification and isolation of the DNA-binding proteins responsible for differential expression of class II molecules in tadpoles and adult *X. laevis*.

Acknowledgments We thank Ms. Keiko Kohda for her technical assistance. This work was supported by Grants-in-Aid for Scientific Research (05670291, 06044008) from The Ministry of Education, Science, and Culture of Japan (to M. K.), and by Grant AI27877 from The National Institutes of Health, Bethesda, Maryland, USA (to M. F. F.). The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche AG, Basel, Switzerland. The authors are willing to share the genomic clones described in this work with interested investigators.

References

Andersson, G., Larhammar, D., Widmark, E., Servenius, B., Peterson, P. A., and Rask, L. Class II genes of the human major histocompatibility complex. Organization and evolutionary relationships of DRβ genes. *J Biol Chem 262:* 8748–8758, 1987

Barnes, W. M. PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proc Natl Acad Scie USA 91:* 2216–2220, 1994

Beato, M. Gene regulation by steroid hormones. *Cell 56:* 335–344, 1989

Benoist, C. and Mathis, D. Regulation of major histocompatibility complex class-II genes: X, Y, and other letters of the alphabet. *Annu Rev Immunol 8:* 681–715, 1990

Bisbee, C. A., Baker, M. A., Wilson, A. C., Hadji-Azimi, I., and Fischberg, M. Albumin phylogeny for clawed frogs (*Xenopus*). *Science 195:* 785–787, 1977

Brown, J. H., Jardetzky, T. S., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature 364:* 33–39, 1993

Du Pasquier, L., Miggiano, V. C., Kobel, H. R., and Fischberg, M. The genetic control of histocompatibility reactions in natural and laboratory-made polyploid individuals of the clawed toad *Xenopus*. *Immunogenetics 5:* 129–141, 1977

Du Pasquier, L., Schwager, J., and Flajnik, M. F. The immune system of *Xenopus*. *Annu Rev Immunol 7:* 251–275, 1989

Flajnik, M. F., Kaufman, J. F., Riegert, P., and Du Pasquier, L. Identification of class I major histocompatibility complex encoded molecules in the amphibian *Xenopus*. *Immunogenetics 20:* 433–442, 1984

Flajnik, M. F., Kaufman, J. F., Hsu, E., Manes, M., Parisot, R., and Du Pasquier, L. Major histocompatibility complex-encoded class I molecules are absent in immunologically competent *Xenopus* before metamorphosis. *J Immunol 137:* 3891–3899, 1986

Flajnik, M. F. and Du Pasquier, L. The major histocompatibility complex of frogs. *Immunol Rev 113:* 47–63, 1990

Flajnik, M. F., Canel, C., Kramer, J., and Kasahara, M. Evolution of the major histocompatibility complex: Molecular cloning of major histocompatibility complex class I from the amphibian *Xenopus*. *Proc Natl Acad Sci USA 88:* 537–541, 1991

Flajnik, M. F., Kasahara, M., Shum, B. P., Salter-Cid, L., Taylor, E., and Du Pasquier, L. A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian *Xenopus*. *EMBO J 12:* 4385–4396, 1993

Frohman, M. A., Dush, M. K., and Martin, G. R. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA 85:* 8998–9002, 1988

Hood, L., Steinmetz, M., and Malissen, B. Genes of the major histocompatibility complex of the mouse. *Annu Rev Immunol 1:* 529–568, 1983

Jonsson, A.-K., Hyldig-Nielsen, J.-J., Servenius, B., Larhammar, D., Andersson, G., Jörgensen, F., Peterson, P. A., and Rask, L. Class II genes of the human major histocompatibility complex. Comparisons of the DQ and DX α and β genes. *J Biol Chem 262:* 8767–8777, 1987

Kandil, E., Noguchi, M., Ishibashi, T., and Kasahara, M. Structural and phylogenetic analysis of the MHC class I-like Fc receptor gene. *J Immunol 154:* 5907–5918, 1995

Kasahara, M., Figueroa, F., and Klein, J. Random cloning of genes from mouse chromosome 17. *Proc Natl Acad Sci USA 84:* 3325–3328, 1987

Kasahara, M., Canel, C., McKinney, E. C., and Flajnik, M. F. Molecular cloning of nurse shark cDNAs with high sequence similarity to nucleoside diphosphate kinase genes. *In* J. Klein and D. Klein (eds.): *Molecular Evolution of the Major Histocompatibility Complex*, pp. 491–499, Springer, Berlin, 1991

Kasahara, M., Flajnik, M. F., Ishibashi, T., and Natori, T. Evolution of the major histocompatibility complex: a current overview. *Transplant Immunol 3:* 1–20, 1995

Kaufman, J. F., Flajnik, M. F., Du Pasquier, P., and Riegert, P. *Xenopus* MHC class II molecules. I. Identification and structural characterization. *J Immunol 134:* 3248–3257, 1985a

Kaufman, J. F., Flajnik, M. F., and Du Pasquier, L. *Xenopus* MHC class II molecules. II. Polymorphism as determined by two-dimensional gel electrophoresis. *J Immunol 134:* 3258–3264, 1985b

Kelly, A. and Trowsdale, J. Complete nucleotide sequence of a functional HLA-DPβ gene and the region between the DPβ1 and DPα1 genes: comparison of the 5′ ends of HLA class II genes. *Nucleic Acids Res 13:* 1607–1621, 1985

Klein, J. *Natural History of the Major Histocompatibility Complex*, John Wiley and Sons, New York, 1986

Klein, J., Bontrop R. E., Dawkins, R. L., Erlich, H. A., Gyllensten, U. B., Heise, E. R., Jones, P. P., Parham, P., Wakeland, E. K., and Watkins, D. I. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics 31:* 217–219, 1990

Klein, J., Ono, H., Klein, D., and Oh'Uigin, C. The accordion model of *Mhc* evolution. *In* J. Gergely and G. Petranyi (eds.): *Progress in Immunology, Vol 8,* pp 137–143, Springer, Heidelberg, 1993

Kobel, H. R. and Du Pasquier, L. Genetics of polyploid *Xenopus*. *Trends Genet 2:* 310–315, 1986

Kozak, M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res 12:* 857–872, 1984

Lawlor, D. A., Zemmour, J., Ennis, P. D., and Parham, P. Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. *Annu Rev Immunol 8:* 23–63, 1990

Mount, S. M. A catalogue of splice junction sequences. *Nucleic Acids Res 10:* 459–472, 1982

Ochman, H., Medhora, M. M., Garza, D., and Hartl, D. L. Amplification of flanking sequences by inverse PCR. *In* M. A. Innis, D. H. Gefand, J. J. Sninsky, and T. J. White (eds.): *PCR Protocols. A Guide to Methods and Applications*, pp. 219–227, Academic Press, San Diego, 1990

Ono, H., O'hUigin, C., Vincek, V., and Klein, J. Exon-intron organization of fish major histocompatibility complex class II *B* genes. *Immunogenetics 38:* 223–234, 1993a

Ono, H., O'hUigin, C., Vincek, V., Stet, R. J. M., Figueroa, F., and Klein, J. New β chain-encoding *Mhc* class II genes in the carp. *Immunogenetics 38:* 146–149, 1993b

Patient, R. K., Elkington, J. A., Kay, R. M., and Williams, J. G. Internal organization of the major adult α- and β-globin genes of *X. laevis*. *Cell 21:* 565–573, 1980

Rackwitz, H. R., Zehetner, G., Frischauf, A.-M., and Lehrach, H. Rapid restriction mapping of DNA cloned in phage vectors. *Gene 30:* 195–200, 1984

Radley, E., Alderton, R. P., Kelly, A., Trowsdale, J., and Beck, S. Genomic organization of HLA-DMA and HLA-DMB. Comparison of the gene organization of all six class II families in the human major histocompatibility complex. *J Biol Chem 269:* 18834–18838, 1994

Rammensee, H.-G., Falk, K., and Rötzschke, O. MHC molecules as peptide receptors. *Curr Opin Immunol 5:* 35–44, 1993

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science 239:* 487–491, 1988

Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, 1989

Sanger, F., Nicklen, S., and Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA 74:* 5463–5467, 1977

Sato, K., Flajnik, M. F., Du Pasquier, L., Katagiri, M., and Kasahara, M. Evolution of the MHC: Isolation of class II β-chain cDNA clones from the amphibian *Xenopus laevis*. *J Immunol 150:* 2831–2843, 1993

Schwager, J., Grossberger, D., and Du Pasquier, L. Organization and rearrangement of immunoglobulin M genes in the amphibian *Xenopus*. *EMBO J 7:* 2409–2415, 1988

Servenius, B., Rask, L., and Peterson, P. A. Class II genes of the human major histocompatibility complex. The DOβ gene is a divergent member of the class II β gene family. *J Biol Chem 262:* 8759–8766, 1987

Shum, B. P., Avila, D., Du Pasquier, L., Kasahara, M., and Flajnik, M. F. Isolation of a classical MHC class I cDNA from an amphibian: evidence for only one class I locus in the *Xenopus* MHC. *J Immunol 151:* 5376–5386, 1993

Sültmann, H., Mayer, W. E., Figueroa, F., O'hUigin, C., and Klein, J. Organization of *Mhc* class II *B* genes in the zebrafish (*Brachydanio rerio*). *Genomics 23:* 1–14, 1994

Takahata, N. MHC diversity and selection. *Immunol Rev 143:* 221–247, 1995

Thiébaud, C. H. and Fischberg, M. DNA content in the genus *Xenopus*. *Chromosoma (Berl.) 59:* 253–257, 1977

Trowsdale, J. "Both man & bird & beast": comparative organization of MHC genes. *Immunogenetics 41:* 1–17, 1995

Vilen, B. J., Coggswell, J. P., and Ting, J. P.-Y. Requirement for stereospecific alignment of the X and Y elements in MHC class II DRA function. *Mol Cell Biol 11:* 2406–2415, 1990

Zoorob, R., Béhar, G., Kroemer, G., and Auffray, C. Organization of a functional chicken class II *B* gene. *Immunogenetics 31:* 179–187, 1990

Zoorob, R., Bernot, A., Renoir, D. M., Choukri, F., and Auffray, C. Chicken major histocompatibility complex class II B genes: analysis of interallelic and interlocus sequence variance. *Eur J Immunol 23:* 1139–1145, 1993