The Immunoglobulin Heavy Chain Locus of the Duck

GENOMIC ORGANIZATION AND EXPRESSION OF D, J, AND C REGION GENES*

Received for publication, July 5, 2001, and in revised form, October 5, 2001
Published, JBC Papers in Press, October 9, 2001, DOI 10.1074/jbc.M106221200

Mats L. Lundqvist‡, Darlene L. Middleton‡, Starr Hazard§, and Gregory W. Warr¶

From the ‡Department of Biochemistry and Molecular Biology and §Biomolecular Computing Resource, Medical University of South Carolina, Charleston, South Carolina 29425

The region of the duck IgH locus extending from upstream of the proximal diversity (D) segment to downstream of the constant gene cluster has been cloned and mapped. A sequence contig of 48,796 base pairs established that the organization of the genes is D-JH-μ-α-v. No evidence for a functional homologue (or remnant) of a δ gene was found. The α gene is in inverted transcriptional orientation; class switch to IgA expression thus requires inversion of the ~27-kilobase pair region that includes both μ and α genes. The secreted forms of duck α and μ are each encoded by 4 constant region exons, and the hydrophobic C-terminal regions of the membrane receptor forms of α and μ are encoded by one and two transmembrane exons, respectively. Putative switch (S) regions were identified for duck μ and ν by comparison with chicken Sμ and Sν sequences and for duck α by comparison with mouse Sα. The duck IgH locus is rich in complex variable number tandem repeats, which occupy ~60% of the sequenced region, and occur at a much higher frequency in the IgH locus than in other sequenced regions of the duck genome.

Antibodies, found only in the vertebrates (1), are molecules that show enormous diversity in structure. The greatest diversity is associated with their almost limitless repertoire of antigen-binding sites, which are encoded by the V, (D), and J genes. However, the constant (C) regions of the Ig molecules also show diversity within an individual. The antibody heavy chains specify the functions associated with the different classes of antibody, such as complement activation, recognition by phagocytic cells, and secretion across mucous membranes. Within the vertebrates, the V, (D), J, and C genes are arranged in many different patterns. In cartilaginous fishes, the genes encoding both heavy and light chains are arranged in multiple small clusters, each of which contains a single V, one or two D elements (in the case of the heavy chain), a single J, and a single C gene. In contrast, in bony fishes, amphibia, and mammals the IgH loci typically show the so-called translocon arrangement, in which groups of V, D, J, and C genes occur sequentially, from 5′ to 3′ within the locus (reviewed in Ref. 2).

The birds are, apart from the mammals, the most highly evolved vertebrate lineage, and their Ig genes show some of the most unusual arrangements and forms of expression. The IgL and IgH loci of chicken each possess only a single functional V and J segment but have up to 100 V region pseudogenes located upstream of the functional V gene (3, 4). The major mechanism creating the large and effective repertoire of the chicken antibody molecule is gene conversion, from the upstream pseudogene segments into the functional V gene. Gene conversion occurs both before B cells encounter antigen and during antigen-induced diversification of the binding site, a process in which point mutation events are also involved (5, 6). Birds possess homologues of IgM and IgA (7–9) and a third class of antibody, IgY (10–12). Avian IgY (sometimes termed IgG) shows homologies with both IgG and IgE of mammals. Extant avian IgY is likely descended from the evolutionary precursor of both IgG and IgE (10, 13).

Ducks and their relatives have unusual antibody structure and expression, with consequences for their function. Their immune responses are often ineffectual (14), a feature that is explained, in part, by the properties of their antibodies. Although ducks produce a typical avian IgY, they can also generate large amounts of a truncated IgY, termed IgY(ΔFc) because it is missing the two C-terminal domains of its H (e) chains (11, 12). This structural abnormality of the duck IgY(ΔFc) would result in the loss of biological effector functions (such as complement activation) associated with the Fc region. The IgA-dependent mucosal immune response of ducks is also problematic, being delayed in its development following hatching (9, 15), as compared with the chicken, in which IgA secretion develops more rapidly (16).

The search for a genetic basis to the inept antibody response of the duck has been informative. The IgY(ΔFc) molecule results from the utilization (by alternative pathways of RNA processing) of a novel, small terminal exon within the ν gene (11, 12). Furthermore, the ν and α genes have been shown to be in head-to-head configuration within the IgH locus (17). This arrangement would require that class switching to produce either IgA or IgY in the duck involves the inversion of a segment of the IgH locus, rather than the usual deletional mechanism of class switching (18) seen in mammals. Detailed knowledge of the structure of the duck IgH locus will be required if a full understanding of the unusual features of its
The donor site of the JH-segment is indicated by an unamplified recombinant duck genomic library (12), constructed in IgM and IgA, and the genetic basis for class switching. The mechanism of expression of the membrane receptor forms was identified by comparison with cDNA sequences. The exon encoding the constant domains of the secreted and transmembrane (TM) forms were sequenced part of the contig, indicated by a dotted line, has a length of ~48 kb and starts 5' of the D segment and ends 3' of the short terminal exon of the \( v \) gene. The coding exons are shown as black boxes. Transcriptional orientation is shown by the arrows above the genes. The sequence has been deposited in EMBL under accession number AJ314754.

**D segment**

\[
\begin{array}{cccccccccccccccccccc}
D & G & Y & T & G & C & G & Y & A & T & Y \\
\alpha & V & I & L & V & V & V & M & L & V & L \\
\end{array}
\]

(206)

GAGTGTGGTCCACCTCTGTTGACTGTG

GGTTATATCTGTGTTGTGTTATGCTTGTACTTATCACCTGTTGATCAT

339

CCCATGACAAAAAACCATCC

**JH segment**

\[
\begin{array}{cccccccccccccccccccc}
JH & T & V & G & S & I & D & L & W & G & H & T & E & V \\
\alpha & T & V & S & S \\
T & 1552
\
CAGGCTCTCTCGGCTGAGCCACACGGAGAGCT
\end{array}
\]

Mapping and Sequencing of Recombinant Duck Genomic \( \lambda \) Clones—DNA was isolated from plaque-purified recombinant genomic \( \lambda \) clones containing IgC region genes (12, 17) and subjected to restriction enzyme digestion. Fragments were separated on agarose gel by electrophoresis and subcloned into pBluescript (Stratagene, La Jolla, CA). The sequencing strategy involved cloning overlapping restriction fragments from the phage inserts. Subclones were completely sequenced (Biotech-ology Resource Laboratory, Medical University of South Carolina) on both strands using a combination of double-stranded nested deletions (Nested Deletion Kit, Amersham Pharmacia Biotech) and transposon-mediated sequencing (Primer Island System, Applied Biosystems, Inc., Foster City, CA). Sequences were assembled into a contig using the SeqMan program (DNastar Inc., Madison, WI). The exons encoding the constant domains of the secreted and transmembrane (TM) forms were identified by comparison with cDNA sequences. The exon encoding the \( \alpha \)TM region was identified through a blastx search (www.ncbi.nlm.nih.gov/blast/blast.cgi) of the region of the locus 3' of the C4d exon and confirmed by RT-PCR analyses.

**Genomic PCR**—The C4d to \( \mu \)TM1 intron was amplified from genomic DNA (prepared by the erythrocytes of Super M strain ducks by Dr. David Higgins, Hong Kong University (12) using specific primers (G-1411. C4d forward, 5'-CAGCTCAACGCCACACGAGA-3', and for \( \mu \)TM1 reverse, G-1530, 5'-CTTGTACAAGTGACCGTGTTG-3') with the Advantage2 GC Genomic PCR kit (CLONTECH, Palo Alto, CA). The C4d to TM intron was amplified from genomic DNA using the forward primer G-1283, 5'-GTTGCCCTCGCGGAGGTTGACCAA-3', and the reverse primer G1284, 5'-GAGGGACAAACAGTGTCGACAAAG-3'. The PCRs were performed using 10 ng of genomic duck DNA as template, with initial denaturation at 94 °C for 5 min, 30 cycles of 94 °C for 25 s, 65 °C for 1 min, 68 °C for 8 min, and a final step of 68 °C for 15 min. The amplified fragments were gel-purified and directly sequenced (C2 to TM intron) or ligated into pGent6-Easy vector (Cp4 to TM1 intron) (Promega, Madison, WI) and subjected to sequencing as described above.

**Reverse Transcription and PCR Analysis**—Total RNA from the spleen and duodenum of two 6-week-old Super M ducks was prepared (9) by Dr. David Higgins, Hong Kong University. To detect the duck \( \lambda \)m message, the RNA was reverse-transcribed using the SMART IV oligonucleotide and PowerScript TM (CLONTECH, Palo Alto, CA). Amplification of the first strand cDNA was then carried out using a two-step protocol (20 cycles of 94 °C, 1 min, 68 °C 6 min, following an initial denaturation of 94 °C for 3 min) using the 5'-PCR and CDSIII/3'-PCR primers (CLONTECH, Palo Alto, CA). A PCR to specifically amplify regions of the \( \lambda \)m sequence was then carried out on the amplified PCR product utilizing an initial denaturation at 94 °C for 3 min, followed by 25 cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 1 min. One reaction used primers G-778 (forward, 5'-GTGACTTGACCTGTGAGACACC-3') and G-1752 (reverse 5'-CTTGATCAAGGTGACGGTGG-3') RACE (20) was performed. The primer used for reverse transcription was G-411, 5'-CCCTGTGGAGACACC-3'. Amplification of the \( \lambda \)m sequence, 3'-RACE (20) was performed. The primer used for reverse transcription was G-413 (5'-TCTGTGTTCTCGAGATCTGCTACTTATCACCTGTTGATCAT-3'), and the anchor primer was G-414 (5'-TCTGTGTTCTCGAGATCTGCTACTTATCACCTGTTGATCAT-3'), and the gene-specific primer was G-1411 (5'-CAAGCTCAACGCCACACGAGA-3'). PCR products were separated by electrophoresis in a 1.2% agarose gel, purified using the Nucleospin kit (CLONTECH, Palo Alto, CA), and directly sequenced.
Analysis of Sequences and Prediction of Class Switch Regions—An initial analysis of repeat sequences within the duck IgH locus and in the duck /H92542-crystallin gene (Ref. 21, GenBank™ accession number U06050) and the adjacent duck S-acyl fatty-acid synthase thioesterase and acyl-CoA-binding protein genes (Refs. 22 and 23, GenBank™ accession numbers M21635 and S73733) was carried out using DotPlot within the Megalign (DNAstar™) program. In addition, the duck IgH sequence was compared with the switch regions of chicken /H9262 and /H9271 genes (Ref. 24, GenBank™ accession numbers AB029075 and AB029077, respectively) and with the switch region of the mouse /H9251 gene (Ref. 25, GenBank™ accession number D11468). The nature of repeat sequences within the putative duck switch regions was also analyzed using a local copy of the EMBOSS program etandem (www.uk.embnet.org/Software/EMBOSS/). The significance of differences in the degree of allelic polymorphism in the duck /H9251 gene and in other immune-related and nonimmune-related genes was assessed as follows. The coding sequences of alleles of duck /H9251, /H9262, /H9271, interferon-/H9253, Mx, S-acyl fatty-acid synthase thioesterase and serum amyloid A type B (GenBank™ accession numbers AJ314754, U27222, U27213, X65218, X65219, X78355, X78356, AF087134, AF100929, Z21549, Z21550, M12101, M21635, U59909, and U64985) were aligned using Megalign (DNAstar™). The aligned sequences for each gene were then randomly subsampled by the SeqBoot program in PHYLIP (Ref. 26, evolution.genetics.washington.edu/phylip.html) to generate 1000 bootstrap replicates of the original alignments. The data were then analyzed for DNA distances using the DNADIST program in PHYLIP. The Kimura two-parameter model was employed with a transition/transversion ratio of 2.0. The resulting 1000 distance estimates were then used to compare the average distance between the /H alleles and those of other duck genes by paired t tests.

RESULTS

Organization of D, J\textsubscript{H} and Constant Region Genes—The physical map of the IgH locus of the duck extending from, at the 5’ end, the proximal D segment to downstream of the v gene is shown in (Fig. 1). Nucleotide sequences of recombinant /H9261 clones 13.1, 3.2, 2.1, and 12.2 formed a contig that included the /H9271 and /H9251 genes and the 3\textsubscript{/H11032} region of the /H9262 gene but which did not overlap with the sequence of /H9261 clone 5.1, which included the D and J\textsubscript{H} segments and the 4 exons encoding the secreted form of /H9262. The 3.8-kb fragment 00-106, generated by genomic PCR between exons /H9262 4 and /H9262 TM1, filled the gap and linked these sequences (Fig. 1). The sequenced region spans more than 48 kb of the duck IgH locus and includes, in addition to a D and a J\textsubscript{H} segment, the /H9262 gene, an inverted /H9251 gene, and the exons that encode the /H004 Fc form of the /H9271 chain. The region downstream of the terminal (T) exon of the /H9271 gene showed frequent recombinations and deletions upon attempted subcloning, and a reliable sequence could not be determined. The intron between the Cu2 and T exons was also unstable upon cloning, and its sequence was confirmed by direct sequencing of an 850-bp fragment derived by genomic PCR between exons /H4 and /H1, filled the gap and linked these sequences (Fig. 1). The sequenced region spans more than 48 kb of the duck IgH locus and includes, in addition to a D and a J\textsubscript{H} segment, the /H gene, an inverted /H gene, and the exons that encode the ΔFc form of the v chain. The region downstream of the terminal (T) exon of the v gene showed frequent recombinations and deletions upon attempted subcloning, and a reliable sequence could not be determined. The intron between the Cu2 and T exons was also unstable upon cloning, and its sequence was confirmed by direct sequencing of an 850-bp fragment derived by genomic PCR. The 36-nt-long D segment is open in all three reading frames (Fig. 2) and is flanked on both sides by nearly canonical recombination signal sequences (RSS) with 12-nt spacers. The RSS 5’ of the J\textsubscript{H} segment is less conserved and has a 23-nt spacer. The duck D and J\textsubscript{H}...
RSS thus follow the 12/23 rule that ensures correct V\textsubscript{H}-D-J\textsubscript{H} recombination. The splice boundaries of the exons of the secreted forms of duck \(\mu\) (A) and \(\alpha\) (B) heavy chain. Primers used in the amplification of the fragments are underlined and identified. The borders between the C4 and TM regions are indicated by vertical bars. The cleavage/polyadenylation signal of \(\mu\) TM is shown in lowercase. Amino acid residues forming the CART motif are shown in boldface.

Identification of Switch Regions—The switch from expression of IgM antibodies to the production of IgY or IgA involves chromosomal recombination at switch (S) regions typically characterized by long (several kb) regions of complex VNTR-like sequences. A DotPlot analysis of the duck IgH locus gave a surprising result (Fig. 5A); the locus is very rich in repeats of the VNTR type, which accounts for \(-60\%\) of the sequence even as assessed at relatively high stringency (\(\geq 90\%\) identity). The repeats also show strong local clustering. The expected location of functional switch regions would be in the I\(\mu\)/C\(\mu\)1 intron for S\(\mu\) and in the C1/C21 intergenic region for both S\(\alpha\) and (because of the reverse orientation of the \(\alpha\) gene) S\(\alpha\). The three largest blocks of VNTRs in the locus occur immediately upstream of exons C\(\mu\)1, C\(\alpha\)1, and C\(\alpha\)2. The functional S\(\mu\) and S\(\alpha\) regions have been identified in the chicken (24), and DotPlot comparison of the duck IgH sequence with the S\(\mu\) and S\(\alpha\) sequences of the chicken (Fig. 6, A and B, respectively) strongly suggests that the large blocks of VNTRs immediately upstream of S\(\mu\) and S\(\alpha\) in the duck (Fig. 6, A and B) are candidates for product detection. Open reading frames that could encode an \(\alpha\) TM segment were then sought by analysis of the genomic sequence 3’ of C\(\alpha\). A candidate sequence was identified between bases 22,680 and 22,852. To determine whether this sequence was expressed, RT-PCR was performed on duck duodenal mRNA using a forward primer specific for C\(\alpha\)4 (G-778) and a reverse primer within the putative TM exon (G-1752, Fig. 4B). A product of \(-350\) bp was amplified and sequenced. This sequence confirmed that the genomic region putatively identified as encoding the \(\alpha\) TM exon was expressed and identified the cryptic donor splice site within C\(\alpha\)4 (Fig. 3). It was then possible to amplify the 3’ end of the \(\alpha\) TM message in RT-PCR by using a forward primer overlapping the C\(\alpha\)4/TM splice site (G-1805) and a reverse primer 3’ of the termination codon (G-1819), confirming that the duck \(\alpha\) TM segment is encoded by a single exon (Fig. 4B). The duck \(\alpha\) TM exon is shorter than the \(\alpha\) TM exons of mouse and human in both the cytoplasmic and extracellular regions. However, the highly conserved residues of the CART motif are present in the membrane-spanning region (Fig. 4B).
functional S regions. The Sa of chicken is not known. Although DotPlot comparisons of mouse Sa with the duck IgH sequence (Fig. 6C) did not yield clear results, the heaviest density of S-acyl-CoA-binding protein (ACBP) genes (GenBank™ accession numbers are as follows: chicken Sa, AB029075; chicken Sγ, AB029077; and mouse Sa, D11468. The proposed Sp, Sa, and Sv have been projected (ovals) onto a map of the duck locus (D) and are found S′ of each corresponding gene. The organization of the repeated consensus motifs, identified with the EMBOSSTandem program and listed in Table I, is shown in transcriptional orientation under the corresponding putative S region. The arrows above the genes indicate transcriptional orientations.

Allelic Variations—Comparisons of the previously published cDNA sequences of duck μ, α, and ν clones and the genomic sequence permitted an analysis of sites of allelic polymorphisms (Table II). This analysis showed an unexpectedly high concentration of polymorphic sites in the α gene: 37 sites of substitution were observed in the α gene, as compared with 10 and 7 in the μ and ν genes, respectively (Table II). The substitutions in the α gene are found in all 5 exons but are concentrated in the first part of the Ca2 exon, where 11 of the 37 substitutions are found within a 30-bp region. The overall rate of allelic polymorphism observed in the duck α gene (2.6%) was determined (as described under “Experimental Procedures”) to be significantly greater (p < 0.001) than that calculated for all duck Ig genes (0.67%) or for the duck non-Ig sequences (0.29%) that have been described to date. Thus, whereas Ig genes are known to evolve relatively rapidly (28), the duck α gene seems to be evolving at a particularly accelerated rate.

DISCUSSION

The IgH locus can be subject to four processes that modify its coding sequence in the course of B cell development and the immune response as follows: site-specific recombination (of V/D-J), point mutations, gene conversions (in some species), and region-specific recombination (of C region genes). Knowledge of the structure and expression of the IgH locus in ducks sheds light on the genetic basis of their poorly functional antibody response and on the evolution of this complex locus in the vertebrates. The observations on the duck IgH locus made in this study include the following: 1) the unusual organization (μ-α-ν) of the C region genes, 2) the inversion that has accompanied the apparent transposition of the α gene in the locus, 3) the absence of a duck δ gene, and 4) the remarkable prevalence of VNTR sequences in the locus. These observations, in total, point to a unique structure for
The results presented here show, similarly, a single TM exon (35). The membrane receptor form of IgA has been studied that splice into a cryptic site in the terminal secreted C region of the receptor form of Igs is encoded by 2 exons (TM1 and TM2) of antigen on B cells. Typically, the hydrophobic transmembrane tail of the functional redundancy of this class of antibody (34). The absence of an IgD from ducks provides further evidence for the case of the mammalian lineages that would give rise to birds and mammals. In observation that mammalian 2.5 kb) intron separating the TM exon from C/H11011 (36). The inverted transcriptional orientation of the duck a gene, while definitively shown here, was the only logical interpretation of previous mapping analyses (17) and, interestingly, may be widespread in the birds, as PCR-based approaches have indicated a similar arrangement in the chicken (29). The current position and orientation of the duck a gene is most readily explained by an ancient translocation event that also inverted the a gene as it was inserted into its present position. This follows from the observation that mammalian α genes are found as the 3′-most C region genes and in the same transcriptional orientation as the other C region genes (30). The translocation and inversion of α must have occurred in a common ancestor of chickens and ducks, but whether this feature is restricted to the gallinaceous lineage (31) or shared by all birds is unknown. The position immediately downstream of the μ gene is the site in which all known δ genes are found (32, 33). The insertion of the α gene in this position downstream of μ may have disrupted the δ gene and accounted for the apparent absence of δ from the present day IgH locus of the duck. No evidence has been found to support the presence of a δ gene, or the discernible remnant of one, in the duck IgH locus. The absence of an IgD from ducks provides further evidence for the functional redundancy of this class of antibody (34).

All lgs can be expressed, by alternative RNA processing, in either secreted forms or as membrane-bound receptors for antigen on B cells. Typically, the hydrophilic transmembrane tail of the receptor form of lgs is encoded by 2 exons (TM1 and TM2) that splice into a cryptic site in the terminal secreted C region exon (35). The membrane receptor form of IgA has been studied previously only in mammals, where the transmembrane region has been shown, uniquely, to be encoded by only a single exon (36). The results presented here show, similarly, a single TM exon in the duck α gene. Thus, the single TM exon in the vertebrate α gene must have developed prior to the divergence of the lineages that would give rise to birds and mammals. In the case of the mammalian α gene, the low frequency of the om message has been suggested to reflect, at least in part, the long (~2.5 kb) intron separating the TM exon from Co3 (37). The homologous Co4TM intron in the duck is close to 5 kb long and may, by the same reasoning, be responsible for the low frequency of om message and in part explain the difficulty in detecting it, even in RT-PCR. However, undefined cis-acting elements present in the mammalian Co3/TM intron also appear to influence the regulation of a mRNA processing (38). Thus, the principal difference between the IgA of mammals and that of birds (only ducks and chickens have been examined in detail) is that in mammals the α chain is shorter by one C region domain, reflecting the loss of the original Co2 exon (39), which has been replaced by a flexible hinge region.

The information presented here allows further examination of the likelihood that the structure of the IgH locus is the cause of the “inert” antibody response in the duck. In one instance it is clear that the structure of the IgH locus leads to expression of a deficient antibody. The IgY(ΔFc) antibody, which lacks the functionally important Fc region, results from an alternative pathway of processing of the primary transcript from the v gene, in which the small terminal exon between the exons encoding v2 and v3 (Fig. 1) is used (12). In a second instance, the inverted position of the a gene in the duck raises the possibility that its orientation is related to the delayed production of IgA observed in ducks (9, 15). This is because the expression of α requires, of necessity, an inversion mechanism of class switching, as opposed to the typical deletional rearrangement. The frequency of inversions during Ig class switching in the IgH locus has been shown, in a mouse cell line, to be lower than that of deletion events (40). Inversions occurred in ~23% of the rearrangements, indicating that deletions are apparently, in mammals, the favored outcome of class switch events at the IgH locus. However, the simple correlation of an inverted a gene with a delayed switch to IgA production is not supported by the evidence from the chicken. The a gene in chickens also appears, from indirect evidence based on PCR approaches (29), to be inverted. However, IgA production in chickens develops rapidly after hatching (16), indicating that an inverted a gene is not, per se, linked to inefficiencies of expression. Thus, delayed IgA expression in the duckling must result from other causes, such as the cytokine control of the mechanisms driving class switching to IgA.

Birds are considered to have a condensed genome, about one-third the size of that of mammals (41). Although the IgH locus in ducks is shorter overall than in mammals, each duck C region gene is considerably larger than its mammalian homologue. For example, the mouse μ gene covers ~4 kb (42), whereas the duck μ gene measures close to 10 kb in length, a difference that is attributable to differences in intron length. In the case of the α gene, lengths are ~6 kb in the mouse versus 11 kb in the duck, a difference attributable both to longer introns

### TABLE I

| Motif | Consensus* | No. repeats | Nucleotide positions |
|-------|------------|-------------|---------------------|
| Sμ    | ggacgtggggggctcg | 3 | 8361–8421 |
| B     | ggacgtat    | 24 | 8931–9020; 9082–9181; 9233–9282 |
| C     | acetccgatgcccaat | 18 | 9342–9421; 9541–9620; 9875–9974; 10,153–10,252 |
| D     | acctccag    | 17 | 9719–9818; 10,071–10,140 |
| Sα    | cagattttccaccaaat | 39 | 39,985–39,926; 39,731–39,592; 39,549–39,390; 39,378–39,219; 39,207–39,009 |
| B     | ggtggeaccaggggtctat | 10 | 38,942–38,883; 38,856–38,797; 38,727–38,645 |

* The consensus motifs are summarized from the list of repeats delivered by the EMBOSS tandem program. The positions of the repeats are numbered according to GenBank™ accession number AJ314754. Consensus sequences are written in transcriptional orientation.
and to the presence of an additional exon in the duck gene. Whereas the overall condensation of the avian genome is generally considered to have been accompanied by a loss of repetitive DNA (43, 44), the data presented here show that the duck IgH locus is, unexpectedly, very rich in VNTRs, which account for ~60% of its sequence (Fig. 5A), a much higher value than seen in mammalian IgH loci. Whereas other sequenced regions of the duck genome (Fig. 5B) contain much lower numbers of VNTRs than the IgH locus, the interpretation of these comparisons is complicated by the fact that VNTRs are asymmetricaly distributed on chromosomes. For example, a telomeric bias in VNTR distribution has been reported on human chromosome 22 and chromosome 1 of Caenorhabditis elegans, but a centromeric bias is present in chromosome 4 of Arabidopsis thaliana (45). As only a small proportion of the duck genome has been sequenced, it is not possible to conclude definitively that a high content of VNTRs is unique to the IgH locus. However, VNTRs are often associated with sites of recombination (46), and the VNTRs in the duck IgH locus may, in addition to a role in the Ig class switch mechanism, have facilitated the inversion and translocation of the α gene that is a prominent feature of this locus.

Acknowledgments—We thank Dr. Kathy Magor and Dr. Ellen Hsu for critical reading of the manuscript and helpful suggestions. We also thank Dr. Robert Chapman for generously performing the analysis of allelic variation.

REFERENCES

1. Du Pasquier, L. (2000) Curr. Top. Microbiol. Immunol. 248, 159–185
2. Bengtén, E., Wilson, M., Miller, N., Clem, L. W., Pilström, L., and Warr, G. W. (2000) Curr. Top. Microbiol. Immunol. 248, 189–219
3. McCormack, W. T., Tjoelker, L. W., and Thompson, C. B. (1991) Annu. Rev. Immunol. 9, 219–241
4. Reynaud, C.-A., Bertocci, B., Dahan, A., and Weill, J.-C. (1994) Adv. Immunol. 57, 353–377
5. Arakawa, H., Furusawa, S., Ekino, S., and Yamagishi, H. (1996) EMBO J. 15, 2540–2546
6. Arakawa, H., Kuma, K., Yasuda, M., Furusawa, S., Ekino, S., and Yamagishi, H. (1998) J. Immunol. 160, 4232–4241
7. Mansikkia, A. (1992) J. Immunol. 149, 855–861
8. Dahan, A., Reynaud, C.-A., and Weill, J.-C. (1983) Nucleic Acids Res. 11, 5381–5389
9. Major, K. E., Warr, G. W., Bando, Y., Middleton, D. L., and Higgins, D. A. (1998) Eur. J. Immunol. 28, 1063–1068
10. Parvari, R., Avivi, A., Lenzner, F., Ziv, E., Tel-Or, S., Burstein, Y., and Schedeicher, I. (1988) EMBO J. 7, 739–744
11. Major, K. E., Warr, G. W., Bando, Y., Wilson, M. R., and Higgins, D. A. (1992) J. Immunol. 149, 2627–2633
12. Major, K. E., Higgins, D. A., Bando, Y., and Warr, G. W. (1994) J. Immunol. 153, 5549–5555
13. Warr, G. W., Major, K. E., and Higgins, D. A. (1995) Immunol. Today 16, 392–398
14. Higgins, D. A., and Warr, G. W. (1993) Avian Pathol. 22, 211–236
15. Ng, P. L. K., and Higgins, D. A. (1986) Immunology 58, 325–327
16. Leslie, G. A., Stankus, R. P., and Martin, L. N. (1976) Int. Arch. Allergy Appl. Immunol. 51, 175–185
17. Major, K. E., Higgins, D. A., Bando, Y., and Warr, G. W. (1999)
18. Stavnezer, J. (1996) Adv. Immunol. 61, 79–146
19. Feinberg, P. A., and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13
20. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8998–9002
21. Li, X., Wistow, G. J., and Piatigorsky, J. (1995) Biochim. Biophys. Acta 1261, 25–34
22. Sasaki, G. C., Chessbrough, V., and Kolattukudy, P. E. (1988) DNA (New York) 7, 449–457
23. Rose, T. M., Schultz, E. R., Sasaki, G. C., Kolattukudy, P. E., and Shoyab, M. (1994) Int. Immunol. 12, 959–968
24. Arakawa, H., Kuma, K., Yamagishi, H., Nakamura, N., Furusawa, S., Matsuda, H., Yasuda, M., Ekino, S., and Shimizu, A. (2000) Int. Immunol. 12, 959–968
25. Davis, M. M., Kim, S. K., and Hood, L. E. (1980) Science 209, 1360–1365
26. Felsenstein, J. (1989) Cladistics 5, 164–168
27. Campbell, K. S., Bickstrom, B. T., Tienfenthaler, G., and Palmer, R. (1994) Semin. Immunol. 6, 393–410
28. Barker, W. C., McLaughlin, P. J., and Dhyoff, M. O. (1972) Atlas of Protein Sequence and Structure, Vol. 5, pp. 31–39, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D. C.
29. Zhao, Y., Rabban, H., Shimizu, A., and Hammarström, L. (2000) Immunology 101, 348–353
30. Flanagan, J. G., and Rabbits, T. H. (1982) Nature 300, 709–713
31. van Tuinen, M., Sibley, C. G., and Hedges, S. B. (2000) Mol. Biol. Evol. 17, 451–457
32. Wilson, M., Bengten, E., Miller, N. W., Clem, L. W., Du Pasquier, L., and Warr, G. W. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 4593–4597
33. White, M. B., Shen, A. L., Word, C., Tucker, P. W., and Blattner, F. R. (1985) Science 228, 735–737
34. Roes, J., and Rajewsky, K. (1993) J. Exp. Med. 177, 45–55
35. Ross, D. A., Wilson, M. R., Miller, N. W., Clem, L. W., and Warr, G. W. (1998) Immunol. Rev. 166, 145–151
36. Yu, L., Peng, C., Starnes, S. M., Liou, R. S., and Chang, T. W. (1990) J. Immunol. 145, 3932–3936
37. Seipel, R. L., and Peterson, M. L. (1995) Mol. Immunol. 32, 277–285
38. Coyle, J. H., and Lebman, D. A. (2000) J. Immunol. 164, 3659–3665
39. Tucker, P. W., Slightom, J. L., and Blattner, F. R. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 7684–7688
40. Jack, H. M., McDowell, M., Steinberg, C. M., and Wahl, M. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 1581–1585
41. Wachtel, S. S., and Tiersch, T. R. (1993) Comp. Biochem. Physiol. B 104, 207–213
42. Bruggemann, M. (1987) Molecular Genetics of Immunoglobulin (Calabi, F., and Neuberger, M. S., eds) pp. 51–80, Elsevier Science Publications B.V., Amsterdam
43. Longmire, J. L., Hahn, D. C., and Roach, J. L. (1999) J. Hered. 90, 574–578
44. Primmer, C. R., Raudsepp, T., Chowdhary, B. P., Möller, A. P., and Ellegren, H. (1997) Genome Res. 7, 471–482
45. Vergnaud, G., and Denoeud, F. (2000) Genome Res. 10, 899–907
46. Brusco, A., Saviozier, S., Cinque, F., Bottaro, A., and DeMarchi, M. (1999) J. Immunol. 163, 4392–4398
Access the most updated version of this article at doi: 10.1074/jbc.M106221200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 45 references, 16 of which can be accessed free at http://www.jbc.org/content/276/50/46729.full.html#ref-list-1