Intergenic exchange maintains identity between two human lambda light chain immunoglobulin gene intron sequences

Jenny A.Udey¹ and Bonnie B.Blomberg*

Department of Microbiology and Immunology, University of Miami Medical School, PO Box 016960, Miami, FL 33101, USA

Received August 27, 1987; Revised and Accepted February 25, 1988

Accession nos: see footnote

ABSTRACT

Evidence for gene conversion or unequal double crossover in the human λ light chain immunoglobulin locus is presented. The high level of J2C2-J3C3 intron cross-hybridization, the identity of the J λ 2 and J λ 3 coding and intron sequences, the presence of multiple base differences between the C λ 2 and C λ 3 coding regions, and the presence of both the unconverted and converted alleles in the normal gene pool, suggest that a recombinational event has resulted in the conversion of the J λ 2 coding and intron sequences to those of J λ 3 and its flanking sequences. Intergenic exchanges, such as the one described here, may provide a mechanism to maintain sequence homogeneity and functionality among the duplicated members of the human lambda gene family.

INTRODUCTION

One gene segment has been shown to encode each of 3 of the four known human lambda immunoglobulin isotypes, Mcg, Kern⁻Oz⁻ and Kern⁻Oz⁺ (C λ 1, C λ 2, and C λ 3, respectively) (1). It is not known which C λ gene segment encodes the Kern⁺Oz⁻ isotype, which accounts for 6% of human lambda myelomas. Originally described from amino acid sequences of human myeloma proteins, the Ke marker is a single amino acid interchange at position 152 and the Oz marker is defined by a single interchange at position 190 and these define the four non-allelic isotypes. We have recently established that a single, functional joining region (J) gene segment lies 5' of each of the human C λ 1, C λ 2, and C λ 3 gene segments (2). Additionally, Southern hybridizations of a mouse J λ probe to our human cosmid clone (Hu λ 9), that contains the cluster of 6 genomic C λ regions, suggested the presence of at least one J-like sequence upstream of each of C λ 4, C λ 5, and C λ 6 (2). The genomic organization of the human λ locus is therefore:

J1C1-J2C2-J3C3-(J)C4-(J)C5-(J)C6 [where (J) denotes one or more J gene segment]

and is similar to the mouse λ locus in that each C λ gene segment known to be functional appears to be preceded by a single J gene segment (3-6).

MATERIALS AND METHODS

Isolation and characterization of human lambda light chain immunoglobulin clones

Screening of a human cosmid library was performed with the mouse C λ 1 probe to produce clone Hu λ 9 and its characterization has been described previously (2). Southern transfer, nick translation and hybridization were performed as before (3,7).

DNA sequencing

DNA fragments were prepared by electrophoresis, electroeluted and end-labeled with α -³²P-dNTPs and the Klenow fragment of DNA polymerase I. Se-

Figure 1. Different cross-hybridization patterns between the duplicated J-C gene segments in the Hu λ 9 clone. Southern blot analysis: a) Hu λ 9 DNA was digested with the enzymes indicated (R=EcoRI, H=HindIII, B=BglII), electrophoresed through a 0.8% agarose gel and transferred to nitrocellulose paper (4). The filter was incubated with a nick-translated 32 P-labelled 2.8 kb HindIII-EcoRI fragment which contains J λ 1-C λ 1 and flanking sequences (the origin of the probes is indicated on the map of the Hu λ 9 clone in the bottom of the figure). The filter was washed under stringent conditions (3 times for 15 min. in 2X SSC, 0.1% SDS at 48°C, followed by 3 washes in 0.1X SSC, 0.1% SDS at the same temperature, and placed under Kodak XAR film at -80°C. b) The same nitrocellulose filter used in the J λ 1-C λ 1 hybridization studies was used in Southern hybridizations to a 3.8 kb HindIII-EcoRI fragment, which contains J λ 2-C λ 2. Areas of hybridization are indicated above the map of the Hu λ 9 clone.

quencing strategy for the introns and J2/J3, as well as the complete restriction enzyme map of genes 1, 2, and 3, was described in ref. 2; briefly, the area included in the 0.4 kb PstI-HinPI fragment containing J2 or J3 was sequenced by first subcloning the 0.6 kb PstI-BamHI fragment into M13mp10, isolating a 0.6 kb HindIII-BamHI fragment (HindIII is immediately adjacent to the PstI site in the m13 multiple cloning sites), digesting with HinPI, and sequenced according to the procedure of Maxam and Gilbert (8).

RESULTS

The Human J λ 1C λ 1 and J λ 2C λ 2-J λ 3C λ 3 Units Probably Arose From Different Ancestral JC Units

The evolutionary relationship among the J1C1, J2C2, and J3C3 units was tested, using different DNA probes in Southern hybridizations to the entire Hu λ 9 cosmid clone. These hybridizations were done under the assumption that the introns, which probably are not under the same pressure to maintain particular sequences as are coding regions, would be highly homologous only between recently duplicated JC units.

When the 2.8 kb HindIII-EcoRI fragment, which contains J1C1, was used as a probe on the cosmid clone, a single band of hybridization was observed in each of the different digests of the DNA (Figure 1a). This single band corresponded to the J1C1-containing fragment. The lack of hybridization to any of the other regions of the clone suggested that this J1C1 unit was more distant evolutionarily from all of the other JC units. This conclusion was also recently supported by the data of Chang et al. (9), who have reported the existence of three λ -like genes which appear to have originated from the same set of duplications that gave rise to J1C1. Since the C λ 1 coding sequences are 95% homologous to the C λ 2 and C λ 3 coding regions (1), the lack of hybridization to the other C λ -containing fragments most likely reflects the divergence of the introns and flanking regions 5' of the J λ 1 region and approximately 1 kb 3' of the constant regions throughout evolution.

The first evidence for a gene conversion involving the J λ 2 and J λ 3 gene segments came from hybridization data using a J2C2 probe. The 3.7 kb EcoRI-HindIII fragment that contains C λ 2 and its upstream flanking sequences was used in Southern hybridizations to the same filter used for the J1C1 hybridizations. As shown in Figure 1b, this fragment hybridizes to the J2C2 region and to the intron sequences 5' of C λ 3. The J3C3 intron appears to be so homologous to the J2C2 intron, that hybridization was visible with exposure times that did not show hybridization to the 1.2 kb BglII fragment containing

```

Jλ1 HindIII PstI
Jλ2&3 TTGGGCTGCAGTGGGGCTGGGGCTAGGGGGCATCCAGGGAG-----T--A-----
                                         nonamer
                                         GGTTTTGGTCTGAGCCT

Jλ1      heptamer TyrValPheGlyThrGlyThrLysValThrValLeuGly
Jλ2&3    CACTCACTGTGTTATGTCTTCGGAAGTCAAGGTACCGTACCTAGGTAAGTGGCTC
          GTG----A----GTG--A----CGGA-----C-G-----G---
          Val      Gly      Leu

Jλ1      TCAA CCTTTCCAGCC TGTCTCACCTCTGCTGTCCCTAGAAAATCTGTTTCTGTG
Jλ2      -TCTC---C---TTC--CAC---TGGGACAATT-C-GA--GTTTTGT-TG-----A
Jλ2&3    3              G

Jλ1      TCTGGGGCTTCTCTCTGTCTCTCCAGCCTTAAGCACTGACCCTTACCTTTTCCATGG
Jλ2&3    ---T-TCTCAA-- ---GG-- -----TCT-C---CAT-CC-GGCC-GAG--A--

Jλ1      GGCCTGGAGGAGGTGCATTAGCCCCGGGTAACCGGCAGGAAGGGCTCCACAGTGGGAGC
Jλ2&3    AC-TCT-CCCTCCCTGT-C--A--TT-CTTG-CT--C-G-TCAT-A--ACCACTTCA-

Jλ1      AGCCTGG
Jλ2&3    CT-TGACCGCAGGGGCAGGGGACTAGATAGAATGACCTACTGAGCCTCGTCTGTCTGTCT

Jλ2&3    GTCTGTCTCTCTGTTTGTCTGTCTGTCTCTGTTTGTCTCTGTCTGTCTGTCTGACAGG

```

Figure 2. Comparison of the Jλ1, Jλ2, and Jλ3 coding and partial intron sequences. The fragments containing Jλ1, Jλ2, and Jλ3 were cloned into plasmid pAT153 or M13 vectors and were sequenced by the method of Maxam and Gilbert (8). The sequences were aligned and spaces inserted to give the best identity. Bases in the 2/3 sequence, identical to the 1 sequence, are indicated by "-". The recombination nonamer and heptamer sequences and the J coding regions are overlined, and the RNA splice sites underlined. The deduced amino acid sequences are indicated above the DNA sequences.

Cλ3 in the HindIII-BglII digest, which is highly homologous to Cλ2 (306/315 bases are identical in clones described by Hieter (1), and 312/315 are identical in our Huλ9 clone, see below).

The DNA sequences of Jλ2 and Jλ3 confirmed the results from the hybridization experiments. The DNA sequences of Jλ1, Jλ2, Jλ3, and their 3' flanking sequences are shown in Figure 2. For the J-coding sequence and the upstream nonamer-heptamer sequences necessary for V-J joining, Jλ1 is 78.2% homologous to Jλ2, whereas the homology of the 5' flanking sequences and the intron sequences is 41.1% for 209 bases compared. In contrast, the homology between Jλ2 and Jλ3 is 100% at the coding level and 99.8%, including the flanking and intron sequences. The DNA sequences of the 0.4 kb HindIII-HinPI fragment containing Jλ2 were found to differ from that of the 0.4 kb HindIII-HinPI fragment containing Jλ3 by a single base in the intron. Importantly, the same

Cλ2 GAACAAGGGCACCCCAACACCTCATCATATACTGAGGTCAGGGGCTCCCTAGGTGGACACTA
Cλ3

Coding Regions

Cλ2 GGACTCTGACCCCTGCCCTCATCCACCCCGCAGGTCAGCCCAAGGCTGCCCCCTCGGTCA
Cλ3

Cλ2 CTCTGTTCCCGCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGCCACACTGGTGTGTCTCATA
Cλ3 A

Cλ2 AGTGACTTCTACCCGGGAGCCGTGACAGTGGCCTGGAAGGCAGATAGCAGCCCCGTCAAGGC
Cλ3

Cλ2 GGGAGTGGAGACCACCACCCCTCCAAACAAGCAACAACAAGTACGCGGCCAGCAGCTATC
Cλ3 C

Arg
242

Cλ2 TGAGCCTGACGCCTGAGCAGTGAAGTCCACAGAAGCTACAGCTGCCAGGTACGCGATGAA
Cλ3 A

Lys

Term 323

Cλ2 GGGAGCACCGTGGAGAAGACAGTGGCCCTACAGAATGTTTCATAGTTCTCAACCCCTACCC
Cλ3 T

Cλ2 CCC-ACCA 3'
Cλ3 C

Figure 3. Evidence for multiple nucleotide differences between the Cλ2 and Cλ3 coding regions and for further intron sequence identity. The sequencing strategy is shown at the bottom of the figure. The Cλ coding regions are boxed. Arrows indicate the direction and extent of sequencing by the Maxam and Gilbert method. The start of the coding regions is indicated by an arrow. The Cλ2 and Cλ3 gene segments within the Huλ9 clone differ by 3 bases, as opposed to the 8 bases reported by Hieter et al. (1) from an independently isolated clone. "Term" refers to the translation termination signal.

change was observed in autoradiographs of gels containing DNA that had been sequenced and run on different dates. The fragment containing Jλ2 was re-isolated, re-subcloned, and re-sequenced to verify that the fragments differed only by that base. Moreover, the same, single base difference was seen in autoradiographs that contained Jλ2 and Jλ3 which had been sequenced and run in parallel. The striking homology between the Jλ2 and Jλ3 coding and intron sequences (409/410 bases) led us to believe that either: a) a JC duplication

had occurred in this clone, b) a very recent J λ 2 duplication had occurred in the population, or c) the J λ 2 and J λ 3 regions had undergone a conversion (inter-genic recombination) event.

The Identity of J λ 2 and J λ 3 is Probably Not Due to a J-C Duplication

The C λ 2 and C λ 3 gene segments have been reported to differ by 8 silent base pair changes and 1 replacement mutation, out of a total of 315 bases (the latter mutation results in the coding for the O z^- isotype by C λ 2 and the O z^+ isotype by C λ 3) (1). The presence of multiple silent mutations suggests that there has been a strong evolutionary pressure to maintain a particular amino acid sequence after many rounds of DNA replication. If the constant regions were identical in the Hu λ 9 clone (i.e., of similar sequence identity as are the J λ 2 and J λ 3 regions), the identity of the joining regions and flanking sequences could be explained by a JC duplication which had occurred only in the Hu λ 9 clone, as opposed to the human gene pool.

To study the possibility of a JC duplication in the Hu λ 9 clone, we sequenced the C λ 2 and C λ 3 gene segments according to the strategy shown at the bottom of Figure 3, which shows the DNA sequences of C λ 2 and C λ 3. These sequences verify the existence of multiple silent mutations, and of the single, replacement mutation which results in the isotypic difference between the C λ 2 and C λ 3 gene segments (at position 242). The identity of the J λ 2 and J λ 3 coding and intron sequences cannot, therefore, be explained by a JC duplication in the Hu λ 9 cosmid clone, since the C λ 2 and C λ 3 gene segments are both clearly there. Although the entire J-C introns have not been sequenced, it was surprising to find that the 75 bases immediately 5' of the C λ 2 and C λ 3 coding sequences were also identical.

J λ 3 Intron Hybridization to Human Germline DNA Samples Suggest That the J λ 2-J λ 3 Gene Conversion is Polymorphic in the Human Population

The 0.4 kb fragment containing J λ 3 and 300 bases of intron sequence was used as a probe on 11 samples of human kidney DNA to determine if the gene conversion event proposed for the J λ 2 and J λ 3 introns was present in the human gene pool. The samples were digested with the enzymes BglII and EcoRI, in order to generate fragments that distinguish the J2C2 intron (2.3 kb) from the J3C3 intron (2.9 kb) (see also Figure 1b), and from the other J-C introns in the lambda locus (9.4, 3.8, 5.4, and 6.0 kb fragments for C λ 1, C λ 4, C λ 5, and C λ 6, respectively). The digested DNA was blotted to nitrocellulose and hybridized to the 0.4 kb HindIII-HinPI fragment containing J λ 3 and intron sequences (Figure 4). Most of the DNA samples fell into one of two categories of hybridization (represented by lanes 1 and 2 in one case, and lanes 3 and 4

Figure 4. Cross-hybridization of human germline J λ 2 and J λ 3 sequences detected by Southern hybridizations. Eleven samples of human kidney DNA (15 μ g) were digested with the enzymes BglII and EcoRI. After digestion, the DNA samples were electrophoresed in 0.8% agarose gels, transferred to nitrocellulose, and hybridized to a nick-translated, 32 P-labelled 400 bp fragment containing J λ 3 and approximately 300 bp of 3' intron sequences. Hybridizations were performed under stringent conditions to identify only those sequences which were highly homologous to J λ 3. With a BglII-EcoRI digestion, J hybridization is expected in fragments of the following sizes in kb for J λ 1, J λ 2, J λ 3, J λ 4, J λ 5, and J λ 6, respectively: 9.4, 2.3, 2.9, 3.8, 5.4, and 6.0 (see also Figure 1b). A densitometric scan is immediately below each lane. Eleven DNA samples were hybridized and most fell into one of two categories, represented by lanes 1 and 2 or lanes 3 and 4. The interpretation of the two categories (lanes 1 and 2 and lanes 3 and 4) is given below the scans. Lanes 1 and 2 represent unconverted chromosomes and 3 and 4 represent chromosomes converted for the J3 region.

in the other). The intensities of hybridization to the J λ 3-intron probe differed for the J λ 2 and J λ 3 fragments in samples 1 and 2, as shown in the autoradiogram of the Southern blot and the densitometric scans below the blots. The J λ 3 fragment showed at least two-fold (2.0-3.0) higher intensity of hybridization than the J λ 2 fragment. We interpret these results to demonstrate lower homology of the J λ 3 intron for the J λ 2 intron and propose that unconverted chromosomes are present in these samples (as depicted schematically at the bottom of the figure). We propose that the 2.3 kb J λ 2 hybridization re-

presents homology between the J λ 2 and J λ 3 coding sequences and perhaps low homology in the intron but not identity as with the J λ 3-intron containing fragment (2.9 kb). Therefore, in these individuals the conversion is not present. Three of the eleven individuals had this pattern.

In lanes 3 and 4 (Figure 4), the hybridization intensities were identical in the 2.9 kb J λ 3 and 2.3 kb J λ 2 fragments. We explain this as homozygosity for the converted chromosome since the J λ 3-intron probe hybridizes identically to both fragments. Four of the eleven samples had this 1.0 ratio. The remaining four samples had an intermediate ratio (between 1.0 and 2.0) and may be accounted for by heterozygosity for the conversion. Densitometric scans were consistent when repeated. It is important to note that we are comparing the intensities of the two fragments within the same lane, which avoids having to standardize for equal amounts of DNA in different lanes.

We have determined that the human lambda gene polymorphism composed of J2C2 duplications (in an 18 kb EcoRI fragment, allele frequency 0.2) (1,11) cannot account for our results. Both categories of hybridization contained individuals with and without the 18 kb EcoRI polymorphism; therefore, there was no correlation between the presence of the duplication and the classification to a particular category (i.e. presence or absence of gene conversion).

The presence of the unconverted, as well as the converted, chromosome in the human gene pool argues against our findings being due to a simple recent duplication of J3C3 to yield J3C2, where the constant region mutated (at pos. 190) to yield the distinct protein isotype Ke⁻Oz⁺. It is unlikely to have both J3C2J3C3 and J2C2J3C3 present as allelic forms without the J3C2 being due to gene conversion or unequal double recombination. Figure 4 also demonstrates that the intergenic exchange occurred at the level of the organism rather than only in the cosmid clone Hu λ 9. However, existence of these conversions in the human gene pool can only be formally demonstrated by obtaining clones of the JC λ 2 and JC λ 3 regions from individuals that show the different hybridization patterns (to J λ 3, shown in Figure 4). Clones of both types of chromosomes (converted and unconverted) would have to be obtained and sequenced.

DISCUSSION

The human λ locus represents a complex gene family that has undergone several JC duplications throughout evolution. These duplications have given rise to both functional and nonfunctional JC units. Of the six C λ gene segments in the major cluster, three are known to be functional. Other, unlinked

λ -like genes have been previously reported by Hollis et al. (12) and by Chang et al. (9). Of these, two are known to be pseudogenes. The first of these is a pseudogene which looks as if it were derived from processed mRNA, and the other does not have a J associated with it and has accumulated multiple mutations which render it nonfunctional. Two C λ genes, recently reported by Chang, have associated J regions and open reading frames, although they do not encode for any constant region described at the protein level. Therefore, it seems that only a few human λ gene segments remain functional in a family of genes that has arisen from multiple, independent duplication events, since 94% of the human lambda isotypes can be accounted for by only three C λ gene segments (2,13). Unlike the kappa locus in mouse and man, which contains clusters of functional J gene segments 5' of a single C λ gene segment, the λ locus has duplicated each J gene segment with an associated C gene segment. The net result has been the maintenance of more than one functional JC unit. However, in the human lambda locus, mutations which render one J gene segment nonfunctional will lead to failure to produce a complete VJC gene that can be expressed. For this reason, the maintenance of functional J λ gene segments is necessary for λ gene expression. Gene conversions between immunoglobulin gene segments have been suggested to account for the maintenance of pseudogenes, to provide additional sequences that can, after recombination, be used to generate antibody diversity (14-19). Gene conversions in the major histocompatibility complex have also been proposed to explain the clustered mutations that occur in a single event contributing to the extensive polymorphism in this complex (20-22). We here propose that gene conversions can occur which use functional immunoglobulin gene segments. In this case, the gene conversion has maintained J gene segments which have both the potential for V-J recombination, and the ability to encode conserved J amino acids, next to two different C λ gene segments. Here, as in other cases of proposed gene conversion, we cannot distinguish between gene conversion and unequal double crossover to explain our results.

ACKNOWLEDGEMENTS

We thank Dr. Gregory Conner for support and assistance in the later portions of these studies and Ms. Wynn Howard for excellent manuscript preparation. These studies were supported by N.I.H. grant AI21870 to B. Blomberg.

*To whom correspondence should be addressed

¹Present address: Department of Anatomy and Cell Biology, University of Miami Medical School, PO Box 016960, Miami, FL 33101, USA

Accession nos: C2 = X06875, C3 = X06876, J1 = X06877, J2/3 = X06878

REFERENCES

1. Hieter, P.A., Hollis, G.F., Korsmeyer, S.J., Waldmann, T.A., and Leder, P. (1981) *Nature* 294, 536-540.
2. Udey, J.A. and Blomberg, B. (1987) *Immunogenetics* 25, 63-70.
3. Blomberg, B., Traunecker, A., Eisen, H.N., and Tonegawa, S. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3765-3769.
4. Blomberg, B. and Tonegawa, S. (1982) *Proc. Natl. Acad. Sci., USA* 79, 530-533.
5. Miller, J., Bothwell, A., and Storb, U. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3829-3833.
6. Selsing, E., Miller, J., and Storb, U. (1982) *Proc. Natl. Acad. Sci. USA* 79, 4681-4685.
7. Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.
8. Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
9. Chang, H., Dmitrovsky, E., Hieter, P.A., Mitchell, K., Leder, P., Turoczi, L., Kirsch, I.R., and Hollis, G.F. (1986) *J. Exp. Med.* 163, 425-435.
10. Slightom, J.L., Blechi, A.E., and Smithies, O. (1980) *Cell* 21, 627-638.
11. Taub, R.A., Hollis, G.F., Hieter, P.A., Korsmeyer, S., Waldmann, T.A., and Leder, P. (1983) *Nature* 304, 172-174.
12. Hollis, G.F., Hieter, P.A., McBride, O.W., Swan, D., and Leder, P. (1982) *Nature* 296, 321-325.
13. Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M., and Gottesman, K.S. (1987) *Sequences of proteins of immunological interest*, 4th edition, U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health.
14. Baltimore, D. (1981) *Cell* 24, 592-594.
15. Givol, D., Zakut, R., Effron, K., Rechavi, G., Ram, D., and Cohen, J.B. (1981) *Nature* 292, 426-430.
16. Dildrop, R., Bruggeman, M., Radbruch, A., Rajewsky, K., and Beyreuther, K. (1982) *EMBO J.* 1, 635-640.
17. Krawinkel, U., Zobebelein, G., Bruggemann, M., Radbruch, A., and Rajewsky, K. (1983) *Proc. Natl. Acad. Sci. USA* 80, 4997-5001.
18. Reynaud, C.-A., Anguez, V., Dahan, A., and Weill, J.-C. (1985) *Cell* 40, 283-291.
19. McCormack, W.T., Laster, S.M., Marzluff, W.F., and Roux, K.H. (1985) *Nucl. Acids Res.* 13, 7041-7054.
20. McIntyre, K.R. and Seidman, J.G. (1984) *Nature* 308, 551-553.
21. Widera, G. and Flavell, R.A. (1984) *EMBO J.* 3, 1221-1226.
22. Mengle-Gaw, L., Conner, S., McDevitt, H.O., and Fathman, C.G. (1984) *J. Exp. Med.* 160, 1184-1194.