Immunogenetics 38: 387–399, 1993

Immuno-genetics

Original articles

Physical location of the human immunoglobulin lambda-like genes, *14.1*, *16.1*, and *16.2*

Thomas R. Bauer, Jr.[1]*, Heather E. McDermid[2], Marcia L. Budarf[3], Margaret L. Van Keuren[4], Bonnie B. Blomberg[1]

[1] University of Miami School of Medicine, Department of Microbiology and Immunology, P. O. Box 016960 (R138), Miami, FL 33101, USA
[2] University of Alberta, Department of Genetics, Edmonton, Alberta T6G 2E9, Canada
[3] University of Pennsylvania School of Medicine, Department of Pediatrics and Division of Human Genetics and Molecular Biology at Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[4] University of Michigan, Howard Hughes Medical Institute, Department of Pediatrics, Ann Arbor, MI 48109, USA

Received April 14, 1993/Revised version received June 4, 1993

Abstract. The human immunoglobulin lambda-like (*IGLL*) genes, which are homologous to the human immunoglobulin lambda (*IGL*) light chain genes, are expressed only in pre-B cells and are involved in B cell development. Three *IGLL* genes, *14.1*, *16.1*, and *16.2* are present in humans as opposed to one, λ5 (*Igll*), found in the mouse. To precisely map the location of the human *IGLL* genes in relation to each other and to the human *IGL* gene locus, at 22q11.1-2, a somatic cell hybrid panel and pulsed field gel electrophoresis (PFGE) were used. Hybridization with a λ-like gene-specific DNA probe to somatic cell hybrids revealed that these genes reside on 22q11.2 between the breakpoint cluster region (*BCR*) and the Ewing sarcoma breakpoint at 22q12 and that gene *16.1* was located distal to genes *14.1* and *16.2*. Gene *14.1* was found by PFGE to be proximal to *16.2* by at least 30 kilobases (kb). A 210 kb *Not* I fragment containing genes *14.1* and *16.2* is adjacent to a 400 kb *Not* I fragment containing the *BCR* locus, which is just distal to the *IGL-C* (*IGL* constant region) genes. We have determined that the *IGLL* genes *14.1* and *16.2* are approximately 670 kb and 690 to 830 kb distal, respectively, to the 3′-most *IGL-C* gene in the *IGL* gene locus, *IGL-C7*. We thus show the first physical linkage of the *IGL* and the *IGLL* genes, *14.1* and *16.2*. We discuss the relevance of methylation patterns and CpG islands to expression, and the evolutionary significance of the *IGLL* gene duplications. Consistent with the GenBank nomenclature, these human *IGLL* genes will be referred to as *IGLL1* (*14.1*), *IGLL2* (*16.2*), and *IGLL3* (*16.1*), reflecting their position on chromosome 22, as established by this report.

* *Present address:* Seattle VA Medical Center, GMR (151), S. Columbian Way, Seattle, WA 98108, USA.

Correspondence to: B. B. Blomberg.

Introduction

The long arm of chromosome 22 (22q) has been shown to be important in the development of numerous genetic conditions, including chronic myeloid leukemia (CML; Caspersson et al. 1970), meningioma (Dumanski et al. 1987), bilateral acoustic neurofibromatosis (Rouleau et al. 1987), DiGeorge syndrome (de la Chapelle et al. 1981), and Cat Eye Syndrome (McDermid et al. 1986). In addition, 22q contains many important genes, including those genes encoding the immunoglobulin (Ig) lambda (λ) light chain, *IGL* (Goyns et al. 1984).

Immunoglobulin λ (Igλ) light chains comprise approximately 40% of the light chain usage in humans. The λ light chains may play a major role in autoimmune diseases such as Graves' disease where 17/19 patients (89%) studied used the λ light chain for the thyroid stimulating antibody (Zakarija 1983; Knight et al. 1986). The *IGL* genes are also directly involved in chromosomal translocations in approximately 10% of the cases of Burkitt's lymphomas (de la Chapelle et al. 1983). The genes encoding the Igλ light chains were first assigned to chromosome 22 (Erikson et al. 1981; McBride et al. 1982) and later were mapped proximal to the CML translocation (chromosome 9; 22) breakpoint (Goyns et al. 1984). The CML breakpoint was later shown to interrupt the breakpoint cluster region (*BCR*) locus (Heisterkamp et al. 1985).

Chang and co-workers (1986) initially identified a new family of Igλ light chain genes, the λ-like genes (*IGLL*), so called because of their high similarity (>85%) to the Igλ light chain constant (*C*) region genes (*IGL-C*). Two genes were identified, *14.1* and *16.1*, which conserved *J* and *C*λ-like coding regions and were tentatively mapped to chromosome 22. The assignment of these genes was similar to the situation in the mouse, where both the Igλ (*Igl*) genes and the λ-like (*Igll*) gene, λ5 (Sakaguchi and Melchers 1986), are

found on chromosome 16 (Kudo et al. 1987a). Recently, Mattei and co-workers (1991) used in situ hybridization to localize the human *IGLL* genes distal to the CML breakpoint.

The tissue specificity and potential function of an *Igll* gene was first described in the mouse. The protein product of the mouse λ5 (*Igll*) gene associates with a V-like protein, *Vpre-B*, which may be encoded by one of two genes, *Vpreb-1* or *Vpreb-2* (Kudo and Melchers 1987). *Vpreb-1* is found 4.6 kb 5′ or upstream of *Igll*, while *Vpreb-2* has not been located relative to *Igll*. *Vpreb-1*, *Vpreb-2*, and *Igll* have not been mapped on the mouse chromosome relative to the λ light chain locus (*IgL*). The human *Vpreb-1* analogue, *VPREB1*, has been mapped within the *Vλ* gene complex (*IGL-V*), proximal to the CML breakpoint and the *Cλ* (*IGL-C*) light chain genes (Bauer et al. 1988). Thus, in contrast to mouse *Vpreb-1* and *Igll* genes, the human *VPREB1* gene is distantly located from the human *IGLL* genes.

Recently, we identified a third *IGLL* gene, *16.2* (Bauer and Blomberg 1990), which is transcribed in pre-B cell tissues including fetal liver and fetal liver hybridomas, but not expressed by pre-B cell lines (Bauer and Blomberg, manuscript in preparation). Schiff and co-workers (1989) also described the *16.2* gene within a cDNA clone, Fλ1, which they concluded was a pseudogene because it had a one base pair (bp) deletion within the exon 3 region. The *16.2* gene was independently cloned and identified in our lab and was named *16.2* in keeping with the nomenclature first used by Chang and co-workers (1986), whereby gene *16.2* also resides on an approximately 16 kb *Eco* RI fragment which differs from that of gene *16.1*. As the in situ hybridization study by Mattei and co-workers (1991) did not differentiate between the three *IGLL* genes *14.1*, *16.1*, and *16.2*, it was unclear as to the organization and location of these three *IGLL* genes. We therefore first wished to determine the location of all three *IGLL* genes in relation to the *IGL* genes, using a panel of previously characterized somatic cell hybrids with a λ-like gene-specific probe, and to determine whether all three *IGLL* genes were located distal to the CML breakpoint (Goyns et al. 1984) and proximal to the Ewing Sarcoma (ES) breakpoint (Budarf et al. 1989). In addition, we wished to use one cell line, RAJ5BE, to map more precisely the three *IGLL* genes. This portion of our work also independently confirms the somatic cell mapping of the *IGLL* genes performed concurrently by Bossy and co-workers (1991). However, they had not determined the relative position of *14.1* and *16.2*, and the absolute distances of the *14.1* and *16.2* genes in relation to the *IGL* locus and the distance to the *16.1* gene. We wished to substantially extend the studies to map the chromosomal locations of the *14.1*, *16.2*, and *16.1* genes, using pulsed field gel electrophoresis (PFGE), probes specific for particular *IGLL* genes and polymerase chain reaction (PCR) of the PFGE fragments. In addition, we wished to determine the relative position of *14.1* and *16.2* (if *14.1* is proximal to *16.2*) as well as the precise distances of these *IGLL* genes to the *IGL-C* gene complex and the 3′-most functional *IGL-C* gene, *IGL-C7* (Bauer and Blomberg 1991), as well as their distances to the *16.1* gene. In precisely establishing the physical location of the human *IGLL* genes, we were then also able to assign gene names consistent with the GenBank nomenclature, and their location on chromosome 22: *IGLL1* (*14.1*), *IGLL2* (*16.2*), and *IGLL3* (*16.1*).

Materials and methods

Human cosmid clones Huλ17 and Huλ18. Two human cosmid clones, Huλ17 and Huλ18, were isolated in a manner similar to that previously described (Udey and Blomberg 1987). Briefly, a cosmid library had been constructed from a partial *Mbo* I digestion of DNA isolated from peripheral blood lymphocytes and cloned into the *Bam* HI-digested cosmid vector PJB8. Clones were picked based on hybridization to a mouse Cλ1 (*Igl-C1*) DNA probe and further characterized on Southern blots. Restriction maps of the cosmid clones were created through single and double digestions of cosmid DNA using various restriction enzymes.

DNA probes. The *IGLL* gene-specific DNA probe (λL) is a 0.7 kb *Nco* I fragment isolated from cosmid Huλ17 (Fig. 1 A). This DNA probe hybridizes on Southern blots specifically to 4.3 kb, 6.4 kb, and 7.8 kb *Bam* HI/*Eco* RI fragments containing genes *14.1*, *16.1*, and *16.2*, respectively. Chang and co-workers (1986) had seen a similar hybridization pattern for genomic DNA using a J16.1 DNA probe, but did not further characterize the unidentified cross-hybridizing fragment at 7.8 kb, which we know from our work contains gene *16.2*. The gene *14.1*-specific probe is a 0.8 kb *Pst* I fragment isolated from cosmid Huλ18 (Fig. 1 B). A probe specific for gene *16.2* was not produced despite extensive searching.

The *Not* I site-specific clone EN73 was isolated from a *Not* I/partial *Sau* 3AI library (McDermid et al. 1989) constructed from a hamster/human hybrid with chromosome 22 as the only intact human chromosome (Ledbetter et al. 1990). *EN73P* is a 0.9 kb *Not* I/*Xba* I fragment from an 11.8 kb phage clone and was found to map to the proximal side of a *Not* I site in the EN73 clone (McDermid et al. 1993). To obtain a probe from the distal side of the *Not* I site (*EN73D*), a phage clone which crossed the *Not* I site (EN73-8a) was isolated from a normal library. *EN73D* is a 0.3 kb *Pst* I/*Sac* I fragment from a 2.0 kb *Xba* I/*Sac* I subcloned fragment (pGEM3Z; Promega, Madison, WI) from that phage clone (Fig. 1 C).

The Cλ probe is a 2.5 kb *Bgl* II fragment from an 8 kb *Eco* RI fragment subcloned into pBR322 (pλ8; Tsujimoto and Croce 1984). It contains the *Cλ2* (*IGL-C2*) gene of the *IGL* gene locus.

The *BCR* probe is a 0.45 *Eco* RI/*Pst* I fragment from a *BCR-ABL* cDNA. The fragment is from the 5′ end and contains only *BCR* sequences (Shtivelman et al. 1985). The *BCR* locus is involved in the (9; 22) CML chromosomal translocation and the *bcr* gene was recently described to encode a GTPase-activating protein (Diekmann et al. 1991).

Description of somatic cell hybrids. Genomic DNA samples were obtained from cell hybrids GL-5, 249 4-2, AA2, 260-3-12-3, BL2-3-2-10, and BL2-1-23-16, and were generously provided by K. Huebner (Wistar Institute, Philadelphia, PA). These hybrids were previously described (Bauer et al. 1988).

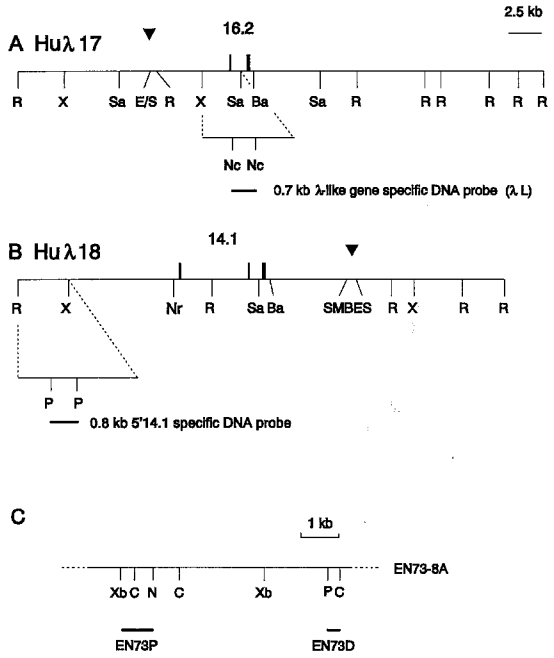


Fig. 1 A – C. Restriction maps of human cosmid clones Huλ17 and Huλ18 and DNA probes. **A** The restriction map of cosmid clone Huλ17. **B** The restriction map of cosmid clone Huλ18. **A** and **B** show the location of the 0.7 kb λL and the 0.8 kb 5′ *14.1*-specific DNA probe, respectively. **C** A partial map of phage clone EN73-8A shows the location of a subcloned 0.3 kb *Pst*/*Sac* I DNA probe (EN73D) and the 0.9 kb *Not* I/*Xba* I DNA probe (EN73P). The *downmark* (▼) indicates the location of potential CpG islands. Ba = *Bam* HI, B = *Bss* HII, C = *Sac* I/*Sst* I, E = *Eag* I, M = *Mlu* I, N = *Not* I, Nc = *Nco* I, Nr = *Nru* I, P = *Pst* I, R = *Eco* RI, S = *Sst* II, Sa = *Sal* I, X = *Xho* I, Xb = *Xba* I. **A, B, C** Only the *Bam* HI, *Nco* I, *Pst* I, and *Sal* I sites used for subcloning genes or probes are shown.

GM10888 is a human/hamster hybrid containing a normal chromosome 22 as its only human component. D6S5 is a normal human mouse fusion which spontaneously lost the distal long arm of chromosome 22. It retains the *IGL* locus but is negative for platelet-derived growth factor β (PDGFB) distal at q12.3-q13.1 (Croce et al. 1987). The human/hamster hybrid cell line c14 has a constitutional t[(11;22)(q23;q11)] retaining the der(22) as the only relevant human chromosome. Both c11-1 and A3W2-3B are independent human/hamster hybrid cell lines made from a human ES cell line which contained the t[(11;22)(q23;q11-q12)] and retains der(11) as the only relevant human chromosome. RAJ5BE is a human/hamster hybrid cell line containing the der(21) of an atypical CML translocation between chromosomes 21 and 22 (Stewart et al. 1988). RJK88 is the parental hamster cell line for some of the hybrids. Parts of this panel have been previously described (Budarf et al. 1989; McDermid et al. 1989).

Description of cell lines used in PFGE. Cell line 1788 (ATCC, Rockville, MD) is an IgMλ-secreting B cell line, reported by ATCC to be hemizygous for either chromosome 21 or 22. We believe 1788 contains only one chromosome 22, as it had only one Igλ rearrangement and no germline *Cλ1* (*IGL-C1*) fragment when hybridized with an *IGL-C1* intron probe (B. Blomberg, unpublished observations), and would be consistent with the ATCC data indicating the absence of one chromosome of either 21 or 22.

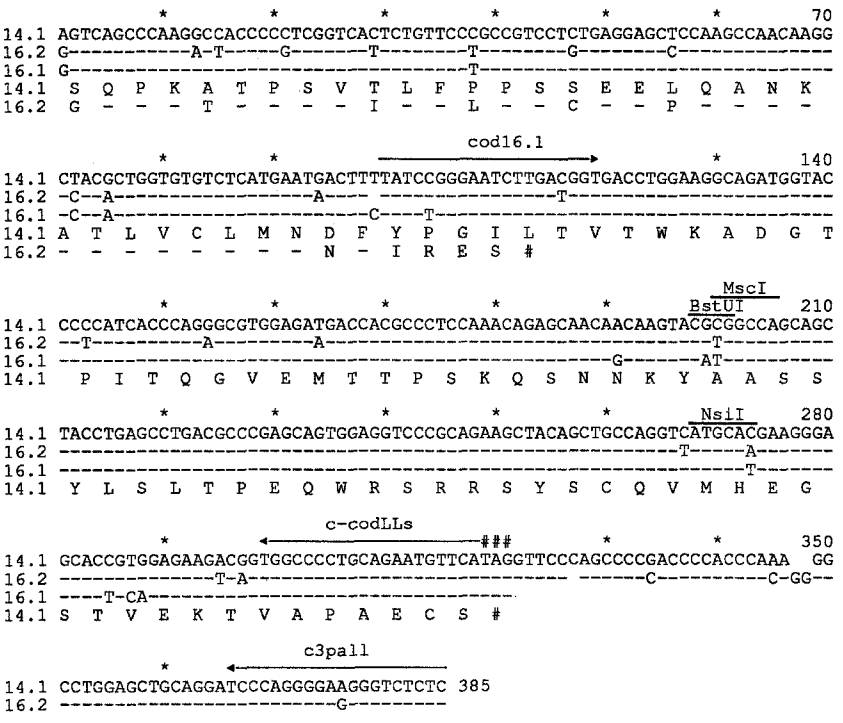
Cell lines GM5565 and GM3657 are a normal fibroblast and lymphoblastoid line, respectively (Human Genetic Mutant Cell Repository, Camden, NJ). An aliquot of GM3657 cells (LA) was treated with 2 μM 5-azacytidine for several weeks prior to harvest for DNA. The light chain produced by GM3657 is not known.

Southern blot analysis of hybrid panels. Five to 10 μg of hybrid or cell line DNA was digested to completion with *Bam* HI and *Eco* RI according to the manufacturer's directions. The digested DNA was electrophoresed on 0.8% agarose gels and blotted to nitrocellulose or nylon membrane filters according to the method of Southern (1975). The λL DNA probe was labeled by the random primer technique (Feinberg and Vogelstein 1983) with ^{32}P αdCTP (Dupont NEN Research Products, Boston, MA) to a specific activity of 5 to 10×10^8 cpm/μg DNA and hybridized to the filters for 12 to 16 h at 42° C with 50% formamide. The filter was then washed using stringent conditions (2 × standard sodium citrate (SSC), 0.1% sodium dodecyl sulfate (SDS) followed by 0.1 × SSC, 0.1% SDS at 50° C) and exposed to X-ray film at −70° C overnight or as necessary (approximately 1 – 10 days).

PFGE. DNA inserts for PFGE, with the following modifications, were prepared similarly to the technique used by Smith and co-workers (1988). Several cubic centimeters of frozen normal kidney tissue were ground to a fine powder using a chilled mortar and pestle. The material was resuspended in cold 1 × phosphate buffered saline (PBS) and filtered twice through sterile gauze. Alternatively, cultured fibroblasts and lymphoblastoid lines were used. The resulting cells were centrifuged three times at 3000 rpm for 10′. Both sources typically yielded between $1-2 \times 10^8$ cells. After the last centrifugation, the pelleted cells were resuspended to approximately $2-5 \times 10^7$ cells/ml in cold L buffer (0.1 M ethylenediaminetetraacetate (EDTA) pH 8.0, 0.01 M Tris-HCl pH 7.6, 0.02 M NaCl). The cells were then warmed to 42° C, mixed 1 : 1 with molten 1% Incert agarose (FMC Bioproducts, Rockland, ME), and pipetted into a BioRad (Richmond, CA) plug casting mold. After refrigeration for 2 h, the plugs were scooped into 30 ml of L buffer containing 0.5 M EDTA, 1 mg/ml Proteinase K (Boehringer Mannheim, Indianapolis, IN) and 1% N-Lauroyl sarcosine (Sigma, St. Louis, MO) and incubated overnight at 50° C with gentle rocking. This process was repeated in a reduced volume of 20 ml. Afterward, the plugs were washed twice in 100 ml of TE pH 7.6 containing 40 μg/ml phenylmethylsulfonylfluoride (PMSF; Research Organics, Cleveland, OH) for 1 h at 50° C. The plugs were finally washed in TE pH 7.6 overnight at 4° C. Double digests used to align the different enzyme maps were performed consecutively with TE washes in between to remove the first enzyme and its buffer.

Approximately 100 μl of plug was preincubated in 1 ml of 1 × restriction buffer for 30 min at 4° C followed by digestion in new 1 × restriction buffer (300 μl) containing 100 μg/ml bovine serum albumen (Gibco BRL, Gaithersburg, MD) and 60 units of restriction enzyme for 12 – 16 h at 37° C. Alternatively, plugs were preincubated in the presence of 10 units of enzyme at 4° C for several hours to allow diffusion. A further 10 units was then added followed by overnight digestion at the recommended temperature. Plugs were washed in 2 – 15 ml of 0.5 × tris-borate (TBE) pH 8.0 prior to electrophoresis. Electrophoresis was performed using 1% agarose and 0.5 × TBE. All gels were electrophoresed at 10 – 15° C in a CHEF DRII system (BioRad). The voltage, running times, and pulse switch times are indicated for each figure. Southern blots of PFGE gels were done essentially as performed for the hybrid panels, except that in some cases hybridization was performed without formamide at 65° C and washed in 0.2 × SSC/0.2% SDS at 65° C.

PCR analysis of PFGE bands. Eag I PFGE. DNA plugs of cell line 1788 were digested with *Eag* I overnight and loaded in two different



Enzyme	Fragment Size	Cut	Uncut
BstUI	184bp	14.1	16.1,16.2
MscI	182bp	16.1,16.2	14.1
NsiI	111bp	16.1	14.1,16.2

Fig. 2. DNA sequences of exon 3/*Cλ*-like region of the *IGLL* genes, showing primer locations and restriction enzyme sites. The locations of the PCR primers used in the assay are indicated by *arrows*. The DNA sequence of the exon 3 region of the *14.1* gene was derived from Chang and co-workers (1986) and Guglielmi and Davi (1991). The DNA sequence of gene *16.1* was derived from the work of Chang and co-workers (1986). No DNA sequence was available for the area downstream of the termination codon for gene *16.1*. The *16.2* gene sequence is our own data, GenBank accession number L02326. The 1 bp deletion at position 99 in 16.2 results in the translation stop, TGA, at position 113. The *16.2* sequence was independently published (as *Fλ1*; Bossy et al. 1991). The expected fragment size of the PCR products of genes *14.1, 16.1*, and *16.2* when amplified by PCR with the c-codLL primer is 284 bp. Sizes when digested with various restriction enzymes are shown *below* the sequences. The amino acid sequences are derived from the cDNA sequences, and the # indicates the termination codon. Restriction enzyme recognition sequences: *Bst* UI = CGCG, *Msc* I = TGGCCA, *Nsi* I = ATGCAT.

lanes on a 1.5% Fastlane agarose gel (FMC Bioproducts). The gel was electrophoresed at 200 V for 16.7 h using a 20 s switch time followed by 23.5 h using a ramped switch time of 20–50 s. After electrophoresis, the gel was cut vertically in half and one-half was blotted. After transfer and hybridization with the λL DNA probe, the resulting autoradiograph from the blotted gel was overlaid on the unblotted gel to locate the four appropriate locations of the hybridizing fragments which were then excised.

Pme I/Not I PFGE. DNA plugs of cell line 3657 were digested with *Pme* I or *Not* I and two identical lanes of each were subjected to PFGE. The gel was electrophoresed at 200 V for 24 h using a ramped switch time of 1–20 s. After electrophoresis, agarose slices were taken from one lane for each enzyme in areas corresponding approximately to where the *IGLL* genes migrated on the gel, based on the ethidium bromide-stained marker lane (λ concatamers). The gel holes were filled in with agarose, vacuum blotted to a nylon filter, and subsequently hybridized to the λL probe. Intact lanes gave the expected hybridization bands adjacent to non-hybridizing gaps in the excised lanes, indicating that excision was precise.

PCR amplification of gel fragments. PFGE fragments were separately amplified by PCR using the Gene-Amp Kit (Perkin Elmer, Norwalk, CT) with *IGLL* gene-specific primers cod16.1 [5′-TATCTGGGAATCTTGACGGT-3′] and c3pall [5′-GAGAGACCCCTCCCCTGGGA-3]′ for 5 cycles (94° C 1′35″, 50° C 2′, 72° C 2′) plus 35 cycles (94° C 1′35″, 55° C 2′, 72° C 2′) and a 10 min extension at 72° C in a Coy Thermocycler (Coy, Ann Arbor, MI). The region of the DNA sequence of primer cod16.1 was chosen because it is conserved among the *IGLL* genes and is highly distinct from the *IGL-C* genes (7, 6, 6, and 6 differences with *IGL-C1, 2, 3*, and *7*, respectively; Bauer and Blomberg 1991). The DNA sequence of primer c3pall is highly conserved among all *IGLL* and *IGL-C* genes. Preliminary experiments established that all three *IGLL* genes, *14.1, 16.2*, and *16.1*, could be amplified by the primers. The PCR products were then digested with *Bst* UI, *Msc* I, or *Nsi* I (New England Biolabs, Beverly, MA) according to the manufacturer's recommendations, which allowed identification of each IGLL gene product (see Figure 2 for primer locations and restriction enzyme sites). The digested PCR products were then electrophoresed on 1.5% agarose mini-gels, blotted to nitrocellulose, and hybridized to an *IGLL* gene primer c-codLLs [5′-TGAACATTCTGCAGGGGCCA-3′]. The c-codLLs oligonucleotide probe was endlabeled with ^{32}P γATP (ICN) using T4 polynucleotide kinase to a specific activity of 5×10^8 cpm/μg DNA. The filter was prehybridized in a solution of 0.9 M NaCl, 0.09 M Tris-HCl pH 7.4, 0.006 M EDTA pH 8.0, 0.5% NP40, 2 × Denhardts, 0.2% SDS, and 100 μg/ml denatured salmon sperm DNA for 4 h. The prehybridization solution was removed and fresh solution was added with 1×10^6 cpm/ml DNA probe. The probe was hybridized to the filter approximately 16 h at 55° C (7° C below the calculated T_D). The filter was then washed three times with 4 × SSC, 0.1% SDS at 55° C, and exposed to X-ray film at 70° C for as long as necessary (approximately 1–10 days).

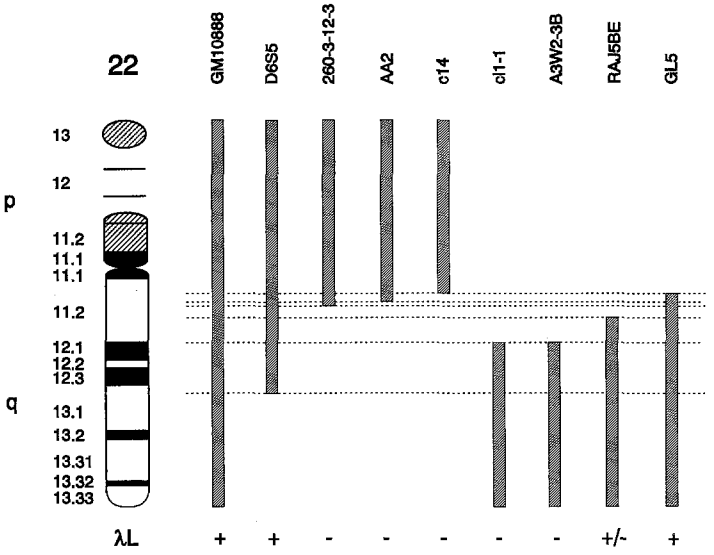


Fig. 3. Idiogram of human chromosome 22 showing hybridization to the λL DNA probe. This idiogram of human chromosome 22 shows the location of the *IGLL* genes with respect to various breakpoints. The *plus* sign indicates the presence of hybridization to all bands of the human λL DNA probe (Fig. 1) which is specific for all three of the *IGLL* genes, *14.1, 16.1,* and *16.2*, but does not hybridize to the *IGL* genes. The *minus* sign indicates the absence of all bands, and the +/– sign indicates the presence of a subset of bands (only *16.1*). *Vertical* bars indicate the region of the human chromosome 22 retained in each hybrid.

Results

Characterization of cosmid clones Huλ17 and Huλ18. Two cosmid clones, Huλ17 and Huλ18, were mapped by various restriction enzymes to determine which *IGLL* genes they contained. Huλ18 contained a 4.3 kb *Bam* HI/*Eco* RI fragment that was identical to the 4.3 kb *Bam* HI/*Eco* RI fragment (Fig. 1) previously shown by Chang and co-workers (1986) to contain gene *14.1*. Cosmid clone Huλ17 contained a 7.8 kb *Bam* HI/*Eco* RI fragment (Fig. 1) of a gene that Chang and co-workers (1986) had described on Southern blots with a J16.1 DNA probe as being distinct from *14.1* and *16.1* but not further characterized. This gene was subcloned from cosmid Huλ17 and resided on a 15.6 kb *Eco* RI fragment (Fig. 1). This gene migrated identically on Southern blots of *Eco* RI-digested genomic DNA to the *16.1* gene identified by Chang and coworkers (1986). It was thus termed *16.2*, consistent with the established nomenclature, as it was the second gene identified on a distinct, but approximately 16 kb, *Eco* RI fragment.

The DNA sequence of gene *16.2* (Fig. 2) showed more sequence identity (94%) to genes *14.1* and *16.1* than to the *IGL* genes (86%), and thus *16.2* was concluded to be a new *IGLL* gene. The restriction maps of these cosmids also revealed that genes *14.1* and *16.2* were separated by more than 30 kb of DNA, based on the more than 15 kb flanking sequence contained in each cosmid clone, which had distinct restriction enzyme sites. The distances separating the *IGLL* genes are in contrast to those of the *IGL* gene complex, where each *IGL-JC* gene is separated from another *IGL-JC* gene by typically no more than 5 kb (Hieter et al. 1981; Udey and Blomberg 1987; Vasicek and Leder 1990; Bauer and Blomberg 1991; Combriato and Klobeck 1991). Mapping of the cosmid clones also revealed the presence of possible CpG-rich islands. CpG-rich islands are associated with the transcription start of housekeeping genes and have been recently found to be associated 40% of the time with the transcription start of tissue-specific genes (Larsen et al. 1992). In addition, Larsen and co-workers (1992) found these islands to be associated with promoter regions and some enhancers 3′ of expressed genes. A potential CpG-rich island containing *Eag* I and *Sst* II sites within 200 base pairs (bp) of one another was found 5′ of gene *16.2* in Huλ17, and a cluster of five CpG-rich enzyme sites within a 500 bp stretch of DNA was found 3′ of gene *14.1* in cosmid Huλ18 and may also represent a CpG-rich island (see Figures 1 A, B). We are currently investigating whether these sites may be potential enhancer sites and involved in gene expression.

Chromosome mapping of IGLL genes. A series of experiments were undertaken to determine the chromosomal location and distances separating the *IGLL* genes in order to gain insight into the potential function, regulation of expression, and evolutionary origin of these genes. Chang and co-workers (1986) first proposed that the human *IGLL* genes resided on chromosome 22. They re-examined hybrid panels used by McBride and co-workers (1982) to localize the *IGL* genes to chromosome 22 and found no discordance between the *IGL* genes and the *IGLL* genes (Chang et al. 1986). However, Chang and co-workers (1986) were also unable to detect discordance for the processed lambda pseudogene, *λψ1*, which has 79% identity to *IGL-C3* and was previously shown not to reside on chromosome 22 (Hollis et al. 1982). Furthermore, Mattei and co-workers (1991) used in situ hybridization to reveal that at least some of the *IGLL* genes were distal to the *IGL* genes and mapped to 22q11.2–22q12.3. However, since the DNA probe used in their experiments can

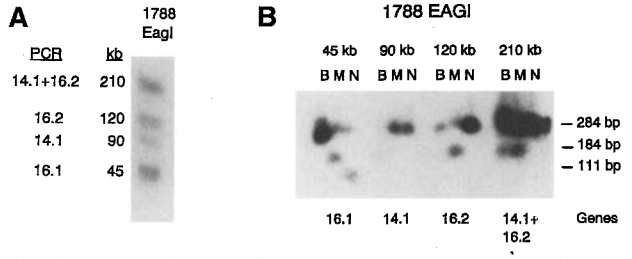


Fig. 4 A, B. Localization of *IGLL* genes to *Eag* I fragments in cell line 1788. **A** DNA plugs of cell line 1788 were digested 16 h with *Eag* I. The digested plugs were subject to PFGE at 200 V for 21.1 h using a ramped switch time of 10–50 s. After electrophoresis, the gel was blotted to a Zetaprobe GT membrane filter (BioRad), baked, and hybridized to the λL DNA probe. The 45 kb, 90 kb, and 120 kb bands in the *Eag* I lane were later determined by PCR to correspond to genes *16.1, 14.1*, and *16.2*, respectively. The 210 kb band is partially digested and contains both genes *14.1* and *16.2* by PCR. **B** Identification of each *IGLL* gene in 1788 *Eag* I fragments by PCR. DNA plugs from fragments of a 1788 *Eag* I PFGE were melted in a boiling water bath and 5 μl were used in a PCR reaction to amplify any *IGLL* gene present. Digestions of the PCR products from the various plug sizes were performed with *Bst* UI (B), *Msc* I (M) or *Nsi* I (N; Fig. 1) to determine the gene present. A band for the 90 kb *Bst* UI digest is present but faint at 184 bp, and thus indicates the presence of *14.1* only in this PCR product. The 210 kb band digested with both *Bst* UI and *Msc* I, but not *Nsi* I, indicating the presence of both genes *14.1* and *16.2*.

hybridize to three different *IGLL* genes, it was unclear whether all three genes were located at this area of hybridization.

To address more precisely the location of these genes, to see whether they were clustered, and to use an additional approach to the in situ studies, Southern blots of genomic DNA isolated from several somatic cell hybrids were performed using the λL DNA probe (see Figure 1 for probe location). Localization of all three *IGLL* genes to chromosome 22 was confirmed by hybridization to DNA of cell line GM10888 (data not shown) which contains chromosome 22 as the only human component. Genes *16.2, 16.1*, and *14.1* were present within *Bam* HI/*Eco* RI-digested DNA fragments of 7.8, 6.4, and 4.3 kb, respectively. Analysis of the hybrid panels revealed that all of the *IGLL* genes map distal to the *IGL* genes and proximal to the ES breakpoint at the 22q11.2/22q12.1 boundary. This conclusion was based on the absence of a hybridization signal to somatic cell hybrid 260-3-12-3, which retains the proximal portion of chromosome 22 above the CML breakpoint as well as to hybrids cl1-1 and A3W2-3B, which contain the portion of chromosome 22 distal to the ES breakpoint.

Finer mapping of the three genes was provided by the somatic cell hybrid line RAJ5BE, which contained only the *16.1* gene, as demonstrated by a 6.4 kb hybridizing band. The RAJ5BE breakpoint is located between the *BCR* (CML) and ES breakpoints. A summary of the somatic cell hybrid mapping data is shown in Figure 3. Thus, all three *IGLL* genes map to 22q11.2 and are ordered as follows: *BCR* – 14.1 and 16.2 – RAJ breakpoint – 16.1 – ES breakpoint.

PCR analysis of cell line 1788 PFGE Eag I fragments reveals that genes 14.1 and 16.2 reside together on a 210 kb fragment. To determine the location of the *IGLL* genes in relation to one another, DNA plugs from cell line 1788 were digested with *Eag* I, *Mlu* I, *Not* I, and *Sst* II, and subjected to PFGE. The *Eag* I digestion pattern seen for 1788 in Figure 4 A was intriguing in that four bands were visible and suggested the possibility of a fourth *IGLL* gene. To determine whether this was the case, gel slices corresponding to the hybridizing fragments in the 1788 *Eag* I digests were isolated and subjected to amplification by PCR using an *IGLL* gene-specific primer (see Materials and methods). Each of the three lower *Eag* I bands at 120 kb, 90 kb, and 45 kb were found to contain a single *IGLL* gene, either gene *16.2, 14.1*, or *16.1*, respectively (Fig. 4 B). The top band at 210 kb, however, gave a PCR digestion pattern indicative of the presence of both genes *14.1* and *16.2*. Therefore, this 210 kb fragment contains an *Eag* I site which partially digests to produce the lower fragment sizes of approximately 120 kb and 90 kb containing genes *16.2* and *14.1*, respectively. Five different PFGE blots with *Eag* I digestions all showed the four bands and confirmed the partial digestion by *Eag* I. These experiments then definitively map the *14.1* and *16.2* genes to a 210 kb *Eag* I fragment. Hybridization of the same blot with the gene *14.1*-specific DNA probe confirmed the localization of this gene to the 210 and 90 kb bands (data not shown).

Eag I digestion of DNA from a fibroblastoid line (GM5565) and a second lymphoblastoid line (GM3657) showed fragment sizes different from the 1788 cell line hybridizing with the λL probe (data not shown). The *IGLL* gene-specific probe revealed bands of 95, 120, and 145 kb, with a weak partial digestion band at 180 kb. The gene *14.1*-specific probe hybridized to the 145 and 180 kb bands. This discrepancy is most likely due to methylation differences in cell line 1788, because when GM3657 is grown in the demethylating agent, 5-azacytidine, the patterns of hybridization with the *λL*- and *14.1*-specific probes resemble more closely that of 1788 (results not shown). Nevertheless, the PCR analysis with line 1788 shows that genes *14.1* and *16.2* are no more than 210 kb from each other.

Long range mapping of Cλ, BCR, and 14.1. A long-range restriction map was produced showing the relationship of *IGL-C* and *BCR* loci to the *IGLL* gene *14.1*, for which a gene-specific probe was prepared. The *Not* I site-specific probes EN73P and EN73D hybridized to contiguous *Not* I fragments of 400 kb and 210 kb, respectively. The *14.1*-specific probe also hybridized to the 210 kb *Not* I fragment, as did the λL probe (data not

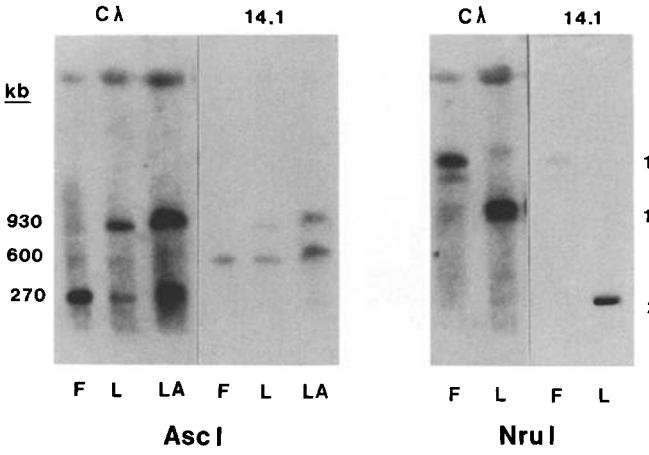


Fig. 5. *Cλ* and *14.1* are located on contiguous *Asc* I and *Nru* I fragments. DNA plugs from normal GM5565 fibroblasts (F), GM3657 lymphoblastoid cells (L), and GM3657 lymphoblastoid cells grown in 5-azacytidine (LA) were digested with either *Asc* I or *Nru* I (NEB). PFGE was performed at 75 V using a switch time of 600 s for 145 h. After transfer to Gene Screen Plus (Dupont NEN), the filter was probed with Cλ, then stripped of probe and rehybridized to *14.1*. Sizes are averages determined from several independent gels. The sizes given here are approximate and averages determined from multiple independent gels are finalized in Figure 6 A. With both enzymes, *Cλ* and *14.1* hybridize to contiguous fragments, as judged by common hybridization to partial bands.

shown). PCR analysis of the excised 210 kb fragment again revealed the presence of both the *14.1* and *16.2* genes (data not shown). Because the recognition sequence of *Eag* I is contained within that of *Not* I, the 210 kb *Eag* I band described above and the 210 kb *Not* I band are presumably identical fragments. The *BCR* probe hybridized to the 400 kb *Not* I fragment. All three probes identified the same 600 kb *Asc* I fragment, as well as a partial fragment of 930 kb in the lymphoblastoid cells (L; Figure 5). The Cλ probe hybridized to an *Asc* I fragment of 270 kb, plus the same lymphoblastoid-specific 930 kb partial fragment (Figure 5; confirmed on two independent blots). In this case, 5-azacytidine (lane LA) did not change the methylation pattern of the region. Thus, the approximately 930 kb partial fragment must represent the addition of the 270 kb and 600 kb fragments, making them contiguous on the chromosome. Therefore, we have physically linked the *IGL-C* genes and the *IGLL* genes *14.1* and *16.2*. The internal *Asc* I site on the 930 kb fragment is fully digested in the fibroblast but only partially digested in the lymphoblastoid line.

DNA digested with *Nru* I also revealed the physical linkage of *IGL-C, BCR*, and *14.1*. Lymphoblastoid DNA (GM3657) gave a 1030 kb band when probed with Cλ, whereas fibroblast DNA (GM5565) gave a major band of 1300 kb (Fig. 5). Previous experiments have shown that this difference may be due to tissue-specific methylation, as GM3658 fibroblast DNA (from the same individual as lymphoblastoid GM3657) shows the 1300 kb band (data not shown). BCR and EN73P both gave an identical hybridization pattern to Cλ. Probe 14.1, however, hybridized to a 250 kb fragment with lymphoblastoid DNA and a 1300 kb fragment with fibroblast DNA. This indicates that the 1300 kb fragment represents the addition of the 250+1030 kb fragments, with the middle site unmethylated and therefore available to enzyme digestion in lymphoblastoid cells but not fibroblasts. A faint band below the 1300–1350 kb fragment in the fibroblast lane probed with Cλ (about 1250 kb) represents a second, partially digested, *Nru* I site (at the dot without a letter) within the large *Nru* I fragment (Fig. 6 A).

Analysis with Pme I indicates that gene 14.1 maps proximal to gene 16.2. Long-range mapping and PCR analysis indicated that *EN73D, 14.1*, and *16.2* all map to a common 210 kb *Not* I fragment, with *EN73D* being the most proximal of the three. To determine the order of 14.1 and 16.2 within the *Not* I fragment, fibroblast (GM5565) and lymphoblastoid (GM3657) DNA were digested with a number of rare-cutting enzymes that tend to give fragments below 200 kb. Three independent Southern blots of the digests were then probed with *EN73D, 14.1*, and *λL*. Hybridization patterns were compared to detect overlap between the probes. For enzymes *Eag* I, *Bss* HII, *Pac* I, *Sma* I, *Sfi* I, and *Cla* I, the bands produced by *EN73D* were different from those produced by the *IGLL* gene probes. Using *Pme* I, however, there was overlap: *EN73D* hybridized to a 100 kb band as well as a faint 255 kb partial fragment in lymphoblastoid cells (Fig. 7). The *14.1*-specific probe showed a similar pattern. The presence of the lymphoblastoid-specific partial fragment in each case makes comigration of two independent bands an unlikely explanation for these results. The λL probe also revealed these bands as well as additional bands at 150 kb and 70 kb, which PCR analysis confirmed as representing the *16.2* and *16.1* genes, respectively (data not shown, but performed as previously in Figure 4 B). Digestion of the PCR product from the 150 kb band occurred only with *Msc* I, revealing only gene *16.2*. The 100 kb *Pme* I PCR product completely digested with *Bst* UI, indicating only *14.1*. Digestion by *Msc* I and *Nsi* I of the 70 kb *Pme* I PCR product indicated the presence of only gene *16.1*. These data confirmed the hybridization data shown with the *14.1* probe (Fig. 7). As the *14.1* gene mapped to the same *Pme* I band as *EN73D*, while the *16.2* gene maps to a contiguous fragment, the order of

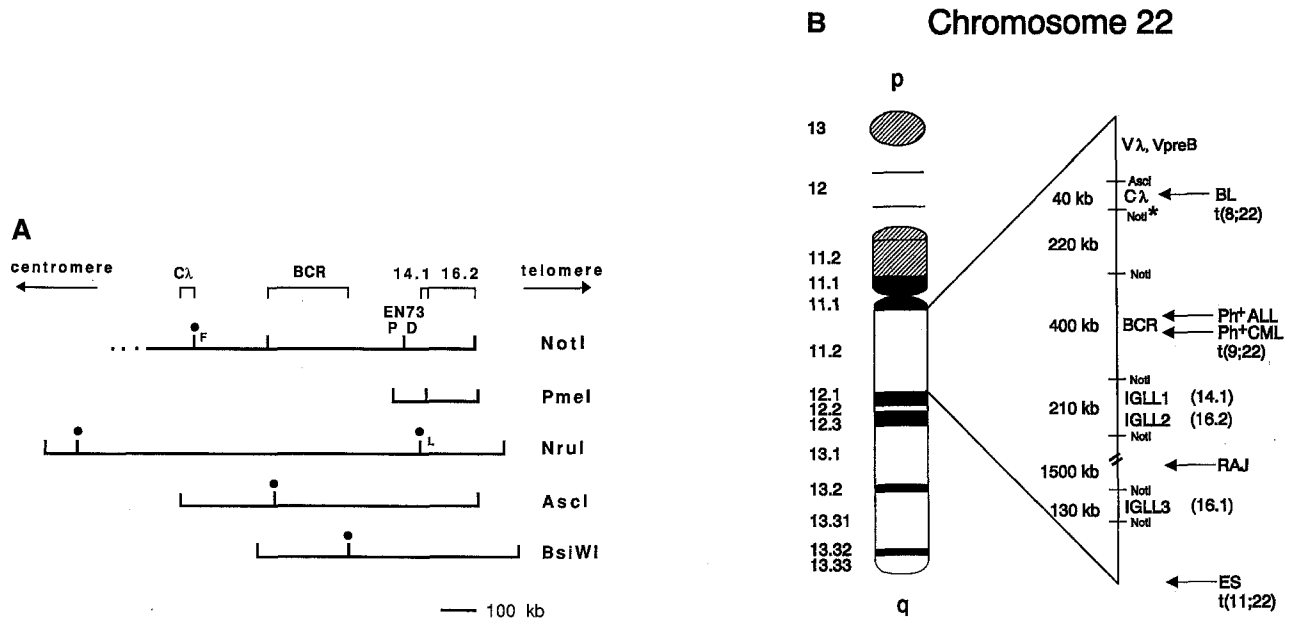


Fig. 6 A, B. A long range map of the *IGLL* genes in relation to the *IGL-C* and *BCR* genes. **A** Sites for enzymes *Not* I, *Pme* I, *Nru* I, *Asc* I, and *Bsi* WI are given. This Figure represents the best estimate of the fragment sizes averaged over many independent blots. *Dots* indicate sites which show partial digestion patterns, possibly due to variable methylation. The presence of an F or L beside the *dot* indicates that the site only cuts (is unmethylated) in fibroblasts or lymphoblasts, respectively. Single enzyme maps were aligned by analyzing double digests. The locations of Cλ, BCR, EN73P and D, and 14.1 are given above the map. **B** Summary of chromosomal locations of the *IGLL* genes *14.1, 16.2*, and *16.1* on chromosome 22. The * indicates partial methylation of the *Not* I site. The 16.1 data are taken from McDermid and co-workers (manuscript submitted).

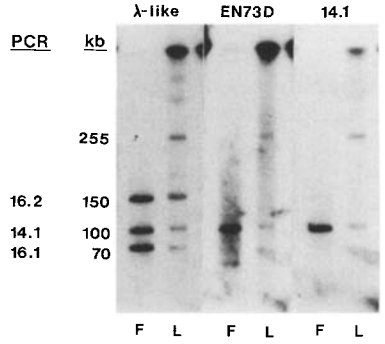


Fig. 7. PFGE analysis with *Pme* I reveals that *14.1* maps proximal to 16.2. DNA plugs from normal fibroblast line GM5565 (F) and normal lymphoblastoid line GM3657 (L) were digested with *Pme* I (NEB). PFGE was at 200 V using a ramped switch time of 5–35 s for 24 h. After blotting to Gene Screen Plus, the membrane was consecutively hybridized and stripped, using the probes as shown. Sizes are averages determined from three independent gels. EN73D and *14.1* hybridize to a 100 kb band as well as the 255 kb partial band. PCR data indicate that gene *16.2* is present in the 150 kb band. Therefore, *14.1* maps closer to EN73 than 16.2. The *IGLL* 70 kb band represents gene *16.1*. The faint 55 kb band of EN73D probably represents cross-hybridization to a locus located elsewhere. The λL DNA probe is designated λ-like in this Figure.

the probes must be *EN73D – 14.1 – 16.2*, moving from proximal to distal.

Figure 6A shows the long range map of the area. The map can be oriented on the chromosome, since *Cλ* is known to be proximal to BCR (Emanuel et al. 1986). Gene *14.1* can be localized within the approximately 20 kb overlap between the 250 kb *Nru* I fragment and the 100 kb *Pme* I fragment. It is thus located immediately distal to the *Nru* I site digested in the lymphoblastoid cells but not the fibroblasts. We have localized an *Nru* I site within our cosmid clone of gene *14.1*, Huλ18, within 0.2 kb 5′ of exon 1 (M. Donohoe and B. Blomberg, work in progress). As the *14.1* probe hybridizes upstream (5′) of this *Nru* I site, our placement of the *Nru* I site also determines the orientation of the *14.1* gene on chromosome 22 as 3′ proximal (exon 3) and 5′ distal (exon 1). The *Pme* I site could not be detected within Huλ18 and therefore the precise distance between the *Nru* I and *Pme* I sites cannot be determined at this time, but it is at least 14 kb.

The map indicates that *14.1* is 670 kb distal to a sometimes methylated *Not* I site adjacent to the *IGL-C* locus. This *Not* I site digests in fibroblasts and 5-azacytidine-treated lymphoblastoid cells, but not untreated lymphoblastoid cells. We have identified a *Not* I site within our cosmid clone Huλ 12.1 which contains *IGL-C7*, a transcriptional enhancer and flanking sequences (Blomberg et al. 1991; B. Blomberg, unpublished da-

ta). This *Not* I site in Huλ 12.1 lies about 2 kb 3′ of *IGL-C7* and 9 kb 5′ of the enhancer and is presumably the *Not* I site adjacent to the *IGL-C* region on the map (Fig. 6A). Thus we have precisely mapped the distance between the most 3′ *IGL-C* gene, *IGL-C7*, and gene *14.1*. Gene *14.1* is 670 kb distal to *IGL-C7* and in the opposite orientation and gene *16.2* is 690 to 830 kb distal to *IGL-C7*.

A summary of the locations of the *IGL* genes on chromosome 22 is given in Figure 6B. Long-range mapping of the region between the *IGL* locus and the ES breakpoint has located pseudogene *16.1* as being 2350–2480 kb distal to Cλ7, or 1660–1810 kb distal to *14.1* (McDermid et al. 1993). The *14.1*, *16.2*, and *16.1* genes have been designated *IGLL1*, *2*, and *3*, respectively, to denote their proximal to distal locations on chromosome 22 and in agreement with the Genome Data Base nomenclature. The *IGLL* genes had also been referred to previously as *IGLL* (Emanuel et al. 1991).

Discussion

The human Igλ light chain constant regions are encoded by a cluster of seven genes (Hieter et al. 1981; Udey and Blomberg 1987; Vasicek and Leder 1990; Bauer and Blomberg 1991; Combriato and Klobeck 1991). Three of these light chains, Mcg, Ke$^-$Oz$^-$, and Ke$^-$Oz$^+$, were determined to be encoded by *IGL-C1*, *IGL-C2*, and *IGL-C3*, respectively (Hieter et al. 1981), and were determined to have a *JC-JC-JC* organization (Udey and Blomberg 1987) as previously determined for the mouse *Igl* locus (Blomberg et al. 1981; Selsing et al. 1982). More recently, we have discovered a fourth, transcriptionally active, *JCλ7* or *IGL-JC7* gene and have determined that *JCλ6* or *IGL-JC6* is a pseudogene (Bauer and Blomberg 1991), as are *JCλ4* and *5* (Vasicek and Leder 1990). The λ light chain gene cluster was identified by Erickson and co-workers (1981) and McBride and co-workers (1982) as residing on chromosome 22 through the use of somatic cell hybrid panels that retained characteristic *Eco* RI fragments containing these genes. Chang and co-workers (1986) found two additional genes related to the *IGL* genes with open reading frames for the *J* and *C* exons, which were then named the λ-like (now *IGLL*) genes, *14.1* and *16.1*, but no further information on their expression or function was reported at that time. Another gene described by Chang and co-workers (1986), *ψ18.1*, had only 58% identity to *IGL-C1* with several large deletions and was not considered further. The *ψ18.1* gene is a pseudogene and distantly related to the *IGLL* genes *14.1*, *16.2* and *16.1*. Gene ψ 18.1has 61% identity to *14.1*, whereas *14.1* and *16.1* are 97% identical in DNA sequence within the exon 3 region. Gene *14.1* was subsequently found by Hollis and co-workers (1989) and Schiff and co-workers (1990) to be expressed only in pre-B cells. We recently found a third *IGLL* gene, *16.2*, and obtained evidence that *14.1* and *16.2* may be expressed in pre-B cells from different tissues (Bauer and Blomberg, manuscript in preparation). Through re-examination of the somatic cell hybrids of McBride and co-workers (1982), Chang and co-workers (1986) had proposed that these *IGLL* genes also resided on chromosome 22. This proposition was further supported by the work of Mattei and co-workers (1991) who used in situ hybridization to identify an area 3′ of the CML breakpoint – a breakpoint distal to the *IGL* genes – that hybridized to their *IGLL* gene-specific probe. However, these results gave neither the relative order of the *IGLL* genes nor their precise locations.

To map more precisely the location of all three *IGLL* genes, *14.1*, *16.1* (Chang et al. 1986), and *16.2* (Bauer and Blomberg 1990), we examined DNA samples obtained from previously characterized somatic cell hybrids for the presence or absence of these genes. Hybrids containing acute lymphocytic leukemia and CML breakpoints in the BCR cluster allowed mapping of these genes distal to these breakpoints. Similarly, hybrids containing ES-derived chromosomes showed these genes to be located proximal to this breakpoint. Additionally, hybrid RAJ5BE which contains a breakpoint between the BCR and ES breakpoints (Delattre et al. 1991) split the *IGLL* genes. The three breakpoints allowed an ordering of the *IGLL* genes, which from the centromere to the telomere on chromosome 22q11 is: BCR > (*14.1*, *16.2*) > RAJ > *16.1* > ES. This part of our study was preformed concurrently by Bossy and co-workers (1991) who also used somatic cell hybrid panels to map the *IGLL* genes and those results, independently performed, are consistent with the results summarized by us. However, Bossy and co-workers (1991) assigned neither the relative order of *14.1* and *16.2* nor the distance given by us below. Our studies have extended those findings by extensive PFGE and PCR of selective fragments. We have determined that *14.1* is proximal to *16.2* on chromosome 22, and have determined the distance of *14.1* to the *IGL-C* locus at 670 kb and the distance of *IGL-C* to *16.2* at 690 to 830 kb. The approximate distance of *14.1* to *16.1* has been determined at 1620–1760 kb (McDermid et al. 1993).

We have also identified CpG islands which may be important in the gene activity of one or both of the *14.1* and *16.2* loci, as these islands are characteristic of various housekeeping genes and tissue-specific genes (Lindsay and Bird 1987). The putative *14.1* CpG island is located about 7 kb 3′ of exon 3. This island may be involved in transcriptional regulation, since a homologous region in mouse λ5, *Igll*, contains several DNAse hypersensitive sites (Yang and Blomberg, unpublished results). The first enhancer found for the

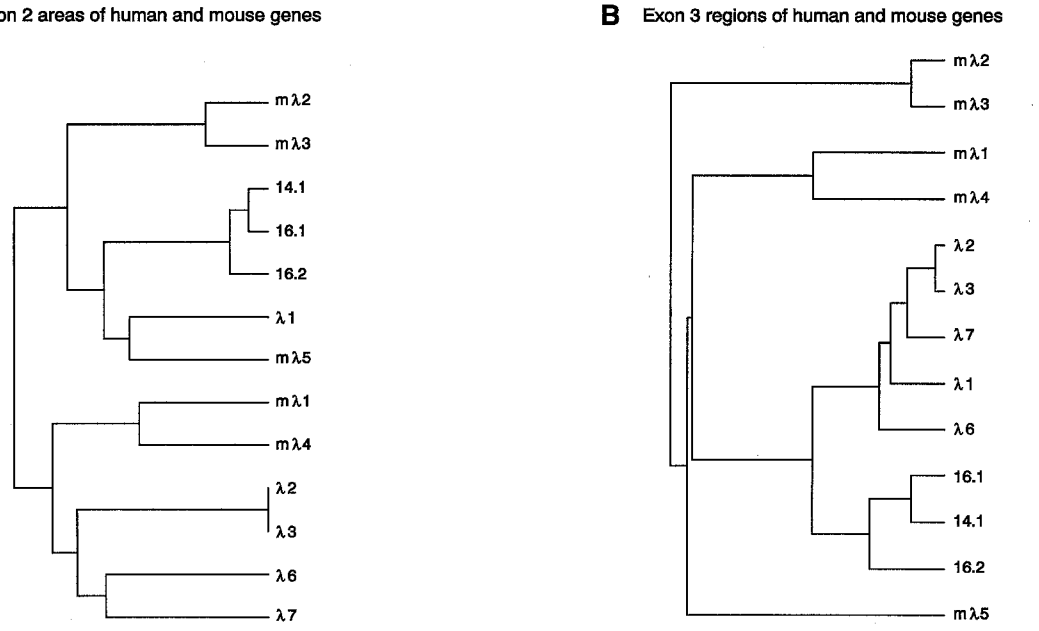


Fig. 8 A, B. DNA sequence comparison for *IGL* and *IGLL* exons 2 and 3. **A** Tree of exon 2 region: trees were obtained using the output function of the PILEUP program of the GCG package. DNA sequences were first optimally aligned pairwise using the GAP program and the gapped sequences were used as input to the PILEUP program. **B** Tree of exon 3 (*Cλ*) region: trees were obtained using the output function of the PILEUP program of the GCG package, using a gap penalty of 5 and a gap length weight penalty of 0.3, as in **A**. DNA sequences for exons 2 and 3 were compared using the GAP program of the University of Wisconsin GCG package with a gap penalty of 2 and a gap length weight penalty of 0.1. After comparisons, the squences were aligned in sequential order, with gaps inserted, using the LINEUP program. The output file was then used for the input file of the DISTANCES program. The exon 2 regions of the *IGL* genes were based on areas of maximum similarity to the exon 2 regions of the *IGLL* genes. The exon 3 regions were based on the exon 3 regions of the *IGLL* genes and the *C* regions of the *IGL-C* genes. The following sequences were used for the exon 2 and exon 3 regions for trees: the *IGLL* genes *14.1* and *16.1* (pseudogene) exons 2 and 3 (Schiff et al. 1990); 16.2 (*Fλ1*, probably a pseudogene) exon 2 (Schiff et al. 1990), exon 3, this paper. The human *IGL-C* genes: *λ1*, exons 2 and 3 (Vasicek and Leder 1990); *λ2*, exon 2 (Vasicek and Leder 1990), exon 3 (Udey and Blomberg 1988); *λ3*, exon 2 (Vasicek and Leder 1990), exon 3 (Udey and Blomberg 1988); *λ6* (pseudogene), exon 2 (Poul et al. 1991), exon 3 (Bauer and Blomberg 1991); *λ7*, exon 2 (Combriato and Klobeck 1991), exon 3 (Bauer and Blomberg 1991), mouse (m) *λ5*, exons 2 and 3 (Kudo et al. 1987 b). The mouse *Igl-C* genes: *mλ1* exon 2 (Bernard et al. 1978; Arp et al. 1982), *mλ2,3,4* (*mλ4* is a pseudogene), exon 2 (Blomberg and Tonegawa 1982), *mλ1,2,3,4*, exon 3 (Selsing et al. 1982). The human pseudogenes *ψ18.1*, *λ4*, and *λ5* have major deletions within exons 2 or 3, which would affect tree similarity calculations and therefore were not used in this analysis.

human *IGL* is also 3′ (11.7 kb) of the 3′-most gene, *IGL-C7* (Blomberg et al. 1991). We have found that the *Not* I site, immediately 3′ of *IGL-C7*, and 9 kb 5′ of the *IGL-C* transcription enhancer (11.7 kb 3′ of *IGL-C7*; Blomberg et al. 1991) is probably methylated in B lymphoblastoid cells and not methylated in fibroblasts. If the *Not* I were within 1 kb of the enhancer we might have predicted the site to be hypomethylated in lymphocytes, but the methylation of this site in the lymphoblastoid cells is probably not significant for the transcriptional enhancer activity of the *IGL* locus. If partial digestion is also indicative of variable methylation at the *Nru* I site near gene *14.1*, it is much less methylated in B lymphoblasts than in fibroblasts (Fig. 5). We have precisely localized this *Nru* I site within our cosmid clone of gene *14.1* to the area including the promoter (and a DNaseI hypersensitive site; Donohoe and Blomberg, work in progress) which is consistent with pre-B lymphocytes and some early B cells which express gene *14.1*, being less methylated at this site. Thus, the methylation status may be useful to measure the transcriptional availability of these genes, as has been reported previously for other immunoglobulin genes (Storb and Arp 1983).

Isolation of *IGLL* gene-containing cosmids and PFGE mapping allowed an estimate of the distance separating *14.1* and *16.2* and the distance to several genes in the area. The central location of *14.1* and *16.2* in their individual non-overlapping cosmids gives a minimal separation of approximately 30 kb. Localization of both genes to 210 kb *Eag* I and *Not* I fragments gives the maximal separation. Localizing *14.1* to the 100 kb *Pme* I and both *14.1* and *16.2* to the 250 kb *Nru* I places *14.1* between the *Nru* I and *Pme* I sites at the 3′-most side of the 100 kb *Pme* I, and therefore *14.1* and *16.2* are most likely within 150 kb of each other (Fig. 6 A). Further analysis produced a map which physically links *IGL-C*, *BCR*, and the two *IGLL* genes, *14.1* and *16.2*, and showed that gene *14.1* resides more proximal to the *IGL-C* locus than gene *16.2*.

The mapping studies of the human *IGL* gene locus and *IGLL* genes presented here allowed comparisons between humans and mouse. In humans, the *IGL-C* region genes and the *IGLL* genes reside more than 600 kb apart. Similarly, PFGE studies in mice by Storb and co-workers (1989) revealed that the *Igll* gene, λ*5*, resides more than 175 kb distant from the *Igl-C* gene segments, although it is unclear from their report whether the λ*5* gene is proximal or distal to *Igl-C*. In addition, one of the mouse *VpreB* genes, *Vpreb-1*, coordinately expressed with λ*5* in mouse pre-B cells, resides only 4.6 kb upstream of λ*5* and may utilize the same promoter for *Vpreb-1* and λ*5* (Kudo et al. 1989), whereas the human homologue, *VPREB1*, resides more than 420 kb proximal (upstream) to *IGL-C*. The other expressed mouse *VpreB* gene, *Vpreb-2*, has neither been mapped nor further characterized (Kudo and Melchers 1987). Thus, the regulation of the coordinated expression of *VPREB1* and *IGLL* genes is likely to differ between mouse and humans.

The human *IGLL* genes, previously referred to as *14.1, 16.2* (*Fλ1*), and *16.1*, have been given the designations *IGLL1, 2*, and *3*, respectively, in accordance with the Gemone Data Base and their relative position on chromosome 22. Although genes *IGLL1* and *IGLL2* were determined as residing within about 150 kb of each other, gene *IGLL3* was determined as mapping more than 1600 kb distal to gene *IGLL1* (McDermid and co-workers, manuscript submitted). Gene *IGLL2*, although physically closer to gene *IGLL1*, has less DNA sequence identity to *IGLL1* than does gene *IGLL3* and, therefore, most likely represents an earlier duplication event (Fig. 2). At the DNA level, gene *IGLL2* has 94.1% sequence similarity within its exon 3 region to *IGLL1*, whereas gene *IGLL3* has 96.3% similarity to gene *IGLL1*. The *IGL* and *IGLL* genes were compared for their sequence identities within exon 2 and exon 3 separately and are shown as tree diagrams in Figure 8. Most of the *IGL-C* (exon 3) regions were previously compared (Hayzer 1990; Vasicek and Leder 1990). Our comparison here includes exon 2 as well as exon 3 of human and mouse *IGL* and *IGLL* genes. Exon 3, which includes the *C* portion of *IGL* genes, had high similarity for both gene sets and probably reflects the conservation necessary for the light chain (or surrogate light chain) to bind the heavy chain, i.e., there is no differential evolutionary pressure on the *IGL* and *IGLL* Ig-like domain. The evolution of exon 2 shows striking differences from that of exon 3. The *IGL* genes use only part of exon 2, expressed as the *J* (joining) portion, whereas the *IGLL* genes have a larger exon 2, including an additional sequence approximately 90 bp upstream of the *IGL J* region. When comparing the exon 2 sequences, it is clear that the *IGLL* genes, [mouse λ*5*, (*mλ5*) *IGLL1, 2*, and *3*] and the exon 2 upstream of *IGL-C1* are most closely related, perhaps reflecting a selection in the *IGL-C* genes (*IGL* other than human *IGL-C1*) away from the conserved sequence seen in exon 2 for the *IGLL* genes. This observation may be explained, as *IGL*, but not *IGLL* genes, undergo DNA rearrangement during B cell development by maintaining consensus nonamer and heptamer recombination sequences 5′ of the *J* sequence, whereas the *IGLL* genes do not.

In comparing the *mλ5* with human *IGLL* genes, exons 2 and 3 show similar sequence identity: exon 2 shows 71%–73% identity to the *IGLL*, whereas exon 3 has 72%–74% identity. Therefore, no apparently different evolutionary pressures were operative on exons 2 and 3 for the *IGLL* genes during the time of evolution, at least from mouse to humans. These data suggest that there was a common evolutionary ancestor for the human and mouse *IGLL* genes. In contrast, interesting differences can be seen between exons 2 and 3 in comparing *mλ5* with human *IGL-C1* and *IGL-C7*, where exon 2 shows 78% identity to *IGL-C1* but only 59% identity to *IGL-C7*, while exon 3 shows 74% identity to both *IGL-C1* and *IGL-C7*. This may indicate conservation of function associated with *IGLL* exon 2 in *IGL-C1* but not *ILG-C7*, and is consistent with *IGL-C1*, but not *IGL-C7*, being transcribed with an *IGLL*-like exon 1. A similar mechanism may be operating in the mouse, where *mλ5* exon 2 is more similar to *Igl-C2* and *Igl-C3* but *mλ5* exon 3 is more similar to *Igl-C1*. In summary, exons 2 and 3 of *IGL* and *IGLL* genes *separately* appear to be evolutionarily conserved between man and mouse within their respective exons, as is seen with *mλ5* and the human *IGLL*. However, these exons have diverged within their exon 2 regions as opposed to their exon 3 regions, reflecting possible different functions for these regions in the *IGL* and *IGLL* genes. The seemingly different selection operative on exons 2 and 3 may be related either to the ontogenetic transcription regulation of the 5′ genes, dependent on the exon 2 or flanking region, and/or on the necessity of the *IGL* genes to use nonamer/heptamer sequences 5′ of the *J* region for DNA recombination to the variable region.

Acknowledgments. We thank Dr. Kay Huebner (Wistar Institute of Anatomy and Biology) for providing DNA samples from various somatic cell hybrids for use in our mapping studies, Ms. Sherri Selby and Mr. Eugene Chomey for technical assistance, and Ms. Michele Glozak for help with figures. This work was supported by NIH grant AI21870 to B. B., CA39926 and HG00425 to M. B., and the Natural Sciences and Engineering Research Council of Canada and the Alberta Heritage Foundation for Medical Research to H. M.

References

Arp, B., McMullen, M., and Storb, U. Sequences of immunoglobulin λ1 genes in a λ1 defective mouse strain. *Nature 298:* 184–187, 1982

Bauer, S. R., Huebner, K., Budarf, M., Finan, J., Erikson, J., Emanuel, B. S., Nowell, P. C., Croce, C. M., and Melchers, F. The human V_{preB} gene is located on chromosome 22 near a

cluster of $V_{\lambda 1}$ gene segments. *Immunogenetics 28:* 328–333, 1988

Bauer, Jr., T. R. and Blomberg, B. A new λ-like light chain gene, 16.2, is related to genes 14.1 and 16.1 expressed in pre-B cells. *J Cell Biochem Suppl 14D:* 215 (M202), 1990

Bauer, Jr., T. R. and Blomberg, B. The human λ L chain Ig locus: Recharacterization of JCλ6 and identification of a functional JCλ7. *J Immunol 146:* 2813–2820, 1991

Bernard, P., Hozumi, N., and Tonegawa, S. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell 15:* 1135–1144, 1978

Blomberg, B., Traunecker, A., Eisen, H., and Tonegawa, S. Organization of four mouse λ light chain immunoglobulin genes. *Proc Natl Acad Sci USA 78:* 3765–3769, 1981

Blomberg, B. and Tonegawa, S. DNA sequences of the joining regions of mouse λ light chain immunoglobulin genes. *Proc Natl Acad Sci USA 79:* 530–533, 1982

Blomberg, B. B., Rudin, C. M., and Storb, U. Identification and localization of an enhancer for the human λL chain Ig gene complex. *J Immunol 147:* 2354–2358, 1991

Bossy, D., Milili, M., Zucman, J., Thomas, G., Fougereau, M., and Schiff, C. Organization and expression of the λ-like genes that contribute to the μ-ψ light chain complex in human pre-B cells, *Int Immunol 3:* 1081–1090, 1991

Budarf, M., Sellinger, B., Griffin, C., and Emanuel, B. S. Comparative mapping of the constitutional and tumor-associated 11;22 translocations. *Am J Hum Genet 45:* 128–139, 1989

Caspersson, T., Gahrton, G., Lindsten, J., and Zech, L. Identification of the Philadelphia chromosome as a number 22 by quinacrine mustard fluorescence analysis. *Exp Cell Res 63:* 238–240, 1970

Chang, H., Dmitrovsky, E., Hieter, P. A., Mitchell, K., Leder, P., Turoczi, L., Kirsch, I. R., and Hollis, G. F. Identification of three new Ig λ-like genes in man. *J Exp Med 163:* 425–435, 1986

Combriato, G. and Klobeck, H.-G. V_λ and J_λ-C_λ gene segments of the human immunoglobulin λ light chain locus are separated by 14 kb and rearrange by a deletion mechanism. *Eur J Immunol 21:* 1513–1522, 1991

Croce, C. M., Huebner, K., Isobe, M., Fainstain, E., Lifshitz, B., Shtivelman, E., and Canaani, E. Mapping of four distinct *BCR*-related loci to chromosome region 22q11: order of *BCR* loci relative to chronic myelogenous leukemia and acute lymphoblastic leukemia breakpoints. *Proc Natl Acad Sci USA 84:* 7174–7178, 1987

De la Chapelle, A., Herva, R., Koivisto, M., and Aula, P. A deletion in chromosome 22 can cause DiGeorge syndrome. *Hum Genet 57:* 253–256, 1981

De la Chapelle, A., Lenoir, G., Boué, J., Boué, A., Gallano, P., Huerre, C., Szajnert, M.-F., Jeanpierre, M., Lalouel, J.-M., and Kaplan, J.-C. Lambda Ig constant region genes are translocated to chromosome 8 in Burkitt's lymphoma with t(8;22). *Nucleic Acids Res 11:* 1133–1142, 1983

Delattre, O., Azambuja, C. J., Aurias, A., Zucman, J., Peter, M., Zhang, F., Hors-Cayla, M. C., Rouleau, G., and Thomas, G. Mapping of human chromosome 22 with a panel of somatic cell hybrids. *Genomics 9:* 721–727, 1991

Diekmann, D., Brill, S., Garrett, M. D., Totty, N., Hsuan, J., Monfries, C., Hall, C., Lim, L., and Hall A. *BCR* encodes a GTPase-activating protein for $p21^{rac}$. *Nature 351:* 400–402, 1991

Dumanski, J. P., Carlbom, E., Collins, V. P., and Nordenskjöld, M. Deletion mapping of a locus on human chromosome 22 involved in the oncogenesis of meningioma. *Proc Natl Acad Sci USA 84:* 9275–9279, 1987

Emanuel, B. S., Nowell, P. C., McKeon, C., Croce, C. M., and Israel, M. A. Translocation breakpoint mapping: molecular and cytogenetic studies of chromosome 22. *Cancer Genet Cytogenet 19:* 81–92, 1986

Erikson, J., Martinis, J., and Croce, C. M. Assignment of the genes for human λ immunoglobulin to chromosome 22. *Nature 294:* 173–175, 1981

Feinberg, A. P. and Vogelstein, B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem 132:* 6–13, 1983

Goyns, M. H., Young, B. D., Geurts van Kessel, A., de Klein, A., Grosveld, G., Bartram, C. R., and Bootsma, D. Regional mapping of the human immunoglobulin lambda light chain to the Philadelphia chromosome in chronic myeloid leukaemia. *Leuk Res 8:* 547–553, 1984

Gugliemi, P. and Davi, F. Expression of a novel type of immunoglobulin Cλ transcripts in human mature B lymphocytes producing κ light chains. *Eur J Immunol 21:* 501–508, 1991

Hayzer, D. J. Immunoglobulin lambda light chain evolution: Ig1 and Ig1-like sequences form three major groups. *Immunogenetics 32:* 157–174, 1990

Heisterkamp, N., Stam, K., Groffen, J., de Klein, A., and Grosveld, G. Structural organization of the *bcr* gene and its role in the Ph′ translocation. *Nature 315:* 758–761, 1985

Hieter, P. A., Hollis, G. F., Korsmeyer, S. J., Waldmann, T. A., and Leder, P. Clustered arrangement of immunoglobulin λ constant region genes in man. *Nature 294:* 536–540, 1981

Hollis, G. F., Hieter, P. A., McBride, O. W., Swan, D., and Leder, P. Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. *Nature 296:* 321–325, 1982

Hollis, G. F., Evans, R. J., Stafford-Hollis, J. M., Korsmeyer, S. J., and McKearn, J. P. Immunoglobulin λ light-chain-related genes 14.1 and 16.1 are expressed in pre-B cells and may encode the human immunoglobulin ω light-chain protein. *Proc Natl Acad Sci USA 86:* 5552–5556, 1989

Knight, J., Laing, P., Knight, A., Adams, D., and Ling, N. Thyroid-stimulating autoantibodies usually contain only λ-light chains: evidence for the "forbidden clone" theory. *J Clin Endocrinol Metab 62:* 342–347, 1986

Kudo, A., Bauer, S., and Melchers, F. Structure, control of expression and putative function of the preB cell-specific genes V_{preB} and λ_5. *In* F. Melchers, E. D. Albert, H. von Boehmer, M. P. Dierich, L. Du Pasquier, K. Eichmann, D. Gemsa, O. Götze, J. R. Kalden, S. H. E. Kaufmann, H. Kirchner, K. Resch, G. Riethmüller, A. Schimpl, C. Sorg, M. Steinmetz, H. Wagner, and H. G. Zachau (eds): *Progress in Immunology*, pp. 339–347, Springer, Berlin Heidelberg New York, 1989

Kudo, A. and Melchers, F. A second gene, V_{preB} in the λ_5 locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. *EMBO J 6:* 2267–2272, 1987

Kudo, A., Pravtcheva, D., Sakaguchi, N., Ruddle, F. H., and Melchers, F. Localization of the murine λ_5 gene on chromosome 16. *Genomics 1:* 277–279, 1987 a

Kudo, A., Sakaguchi, N., and Melchers, F. Organization of the murine-Ig-related λ_5 gene transcribed selectively in pre-B lymphocytes. *EMBO J 6*: 103–107, 1987; Corrigendum: *EMBO J 6:* 4242, 1987 b

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. CpG islands as gene markers in the human genome. *Genomics 13:* 1095–1107, 1992

Ledbetter, S. A., Garcia-Heras, J., and Ledbetter, D. H. "PCR-karyotype" of human chromosomes in somatic cell hybrids. *Genomics 8:* 614–622, 1990

Lindsay, S. and Bird, A. P. Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature 327:* 336–338, 1987

Mattei, M.-G., Fumoux, F., Roeckel, N., Fougereau, M., and Schiff, C. The human pre-B-specific λ-like cluster is located in the 22q11.2-22q12.3 region, distal to the IgCλ locus. *Genomics 9:* 544–546, 1991

McBride, O. W., Hieter, P. A., Hollis, G. F., Swan, D., Otey, M. C., and Leder, P. Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J Exp Med 155:* 1480–1490, 1982

McDermid, H. E., Duncan, A. M. V., Brasch, K. R., Holden, J. J. A., Magenis, E., Sheehy, R., Burn, J., Kardon, N., Noel, B., Schinzel, A., Teshima, I., and White, B. N. Characterization of the supernumerary chromosome in Cat Eye syndrome. *Science 232:* 646–648, 1986

McDermid, H. E., Budarf, M. L., and Emanuel, B. S. Toward a long-range map of human chromosome band 22q11. *Genomics 5: 1*–8, 1989

McDermid, H. E., Budarf, M. L., and Emanuel, B. S. Long-range restriction map of human chromosome 22q11–22q12 between the lambda immunoglobulin locus and the Ewing sarcoma breakpoint. *Genomics,* in press, 1993

Poul, M.-A., Zhang, X.-M., Ducret, F., and Lefranc, M. P. The IGLJ6 joining segment as a STS in the human immunoglobulin lambda light chain constant region gene locus (located at 22q11). *Nucleic Acids Res 19:* 4785, 1991

Rouleau, G. A., Wertelecki, W., Haines, J. L., Hobbs, W. J., Trofatter, J. A., Seizinger, B. R., Martuza, R. L., Superneau, D. W., Conneally, P. M., and Gusella, J. F. Genetic linkage of bilateral acoustic neurofibromatosis to a DNA marker on chromosome 22. *Nature 329:* 246–248, 1987

Sakaguchi, N. and Melchers, F. λ5, a new light-chain-related locus selectively expressed in pre-B lymphocytes. *Nature 324:* 579–582, 1986

Schiff, C., Milili, M., and Fougereau, M. Isolation of early immunoglobulin λ-like gene transcripts in human fetal liver. *Eur J Immunol 19:* 1873–1878, 1989

Schiff, C., Bensmana, M., Guglielmi, P., Milili, M., Lefranc, M.-P., and Fougereau, M. The immunoglobulin λ-like gene cluster (14.1, 16.1, and Fλ1) contains gene(s) selectively expressed in pre-B cells and is the human counterpart of the mouse λ5 gene. *Int Immunol 2:* 201–207, 1990

Selsing, E., Miller, J., Wilson, R., and Storb, U. Evolution of mouse immunoglobulin λ genes. *Proc Natl Acad Sci USA 79:* 4681–4685, 1982

Shtivelman, E., Lifshitz, B., Gale, R. P., and Canaani, E. Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukemia. *Nature 315:* 550–554, 1985

Smith, C. L., Klco, S. R., and Cantor, C. R. Pulsed field gel electrophoresis and the technology of large DNA molecules. *In* K. E. Davies (ed.): *Genome Analysis: A Practical Approach,* pp. 41–60, IRL press, Washington, DC, 1988

Southern, E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol 98:* 503–517, 1975

Stewart, G. D., Tanzi, R. E., Kishimoto, T. K., Buraczynska, M., Haines, J. L., Drabkin, H., Gusella, J. F., Springer, T. A., Kurnit, D. M., and Van Keuren, M. The CD18 gene maps to distal chromosome 21q22.3: RFLPs create a highly informative terminal haplotype. *Am J Hum Genet 43: 3 (Suppl A160),* 1988

Storb, U. and Arp, B. Methylation patterns of immunoglobulin genes in lymphoid cells: correlation of expression and differentiation with undermethylation. *Proc Natl Acad Sci USA 80:* 6642–6646, 1983

Storb, U., Haasch, D., Arp, B., Sanchez, P., Cazenave, P.-A., and Miller, J. Physical linkage of mouse λ genes by pulsed field gel electrophoresis suggests that the rearrangement process favors proximate target sequences. *Mol Cell Biol 9:* 711–718, 1989

Tsujimoto, Y. and Croce, C. M. Molecular cloning of human immunoglobulin λ chain variable sequence. *Nucleic Acids Res 12:* 8407–8414, 1984

Udey, J. A. and Blomberg, B. Human λ light chain locus: organization and DNA sequences of three genomic *J* regions. *Immunogenetics 25:* 63–70, 1987

Udey, J. A. and Blomberg, B. B. Intergenic exchange maintains identity between two human lambda light chain immunoglobulin gene intron sequences. *Nucleic Acids Res 16:* 2959–2969, 1988

Vasicek, T. J. and Leder, P. Structure and expression of the human immunoglobulin λ genes. *J Exp Med 172:* 609–620, 1990

Zakarija, M. Immunochemical characterization of the thyroid-stimulating antibody (TSAb) of Graves' disease: evidence for restricted heterogeneity. *J Clin Lab Immunol 10:* 77–85, 1983

Note added in proof.

The mouse *Igll* genes have recently been mapped to chromosome 16 to within one centimorgan proximal of the *Igl* genes. (Miller, R. D., Ozaki, J. J., and Riblet, R. The mouse *severe combined immune deficiency (scid)* mutation is closely linked to the B-cell-specific developmental genes *VpreB* and *λ5. Genomics 16:* 740–744, 1993)