Structure and Expression of Human Germline V\text{H} Transcripts

By Jeffrey E. Berman,* Caroline G. Humphries,† Jeremy Barth,* Frederick W. Alt,* and Philip W. Tucker†

From the *Howard Hughes Medical Institute and the Departments of Biochemistry and Molecular Biophysics, and Microbiology, College of Physicians and Surgeons of Columbia University, New York, New York 10032; and the †Department of Microbiology, University of Texas Health Science Center, Dallas, Texas 75235

Summary

The human V\text{H} family consists of two functional genes and one pseudogene. We have found a novel 1.2-kb V\text{H} gene transcript in normal fetal liver and cord blood and in transformed B lineage cells. V\text{H}-positive cDNA clones were isolated from precursor B acute lymphoblastic leukemia, B chronic lymphoblastic leukemia, Epstein-Barr Virus-transformed B cell lines, and cord blood, and were identified as transcripts of unrearranged V\text{H} genes (germline transcripts). The cDNA clones were derived from both functional and pseudo-V\text{H} genes. Most germline transcripts appear to initiate at the normal V\text{H} promoter and are cleaved and polyadenylated at sites several hundred bases downstream of the V\text{H} coding region. Correct splicing of the leader intron was observed in all clones. In functional and pseudo-V\text{H} cDNAs, an open translational reading frame extends from the leader to a termination codon in the nonamer. Only limited polymorphisms were observed in the coding as well as flanking regions of the V\text{H} transcripts. Functional and pseudo-V\text{H} cDNAs are also highly homologous throughout. The many similarities between human germline V\text{H} transcripts and previously identified murine germline V\text{J}558 transcripts are discussed.

The variable regions of immunoglobulin (Ig) heavy (H) chains are assembled from three groups of germline gene segments, V\text{H}, D, and J\text{H}. During precursor B cell differentiation in fetal liver and adult bone marrow, these segments are joined into a complete V\text{H}-D-J\text{H} variable region by an ordered assembly process. First, D to J\text{H} segments are joined, followed by V\text{H} to D-J\text{H} rearrangement (reviewed in reference 1). In humans, there are 100–200 V\text{H} gene segments, which have been classified into six families (V\text{H}1 to V\text{H}6) on the basis of nucleic acid sequence homology. V\text{H} families range in size from the single-member V\text{H}6 family to the large V\text{H}1 and V\text{H}3 families, each containing >25–30 members (reviewed in reference 2). The small V\text{H}5 family consists of two functional genes and one pseudogene (3–5). One of the functional V\text{H}5 genes is not present in all individuals (4).

In general, most V\text{H} segments are not expressed at detectable steady-state levels in unrearranged, germline configuration. V\text{H} to D-J\text{H} rearrangement brings the V\text{H} segment and its associated promoter elements into proximity of an enhancer located within the J\text{H}-C\text{H} intron, thus allowing for high-level transcription of the V\text{H}-D-J\text{H} rearrangement initiating at the V\text{H} promoter and continuing into the downstream constant region (reviewed in reference 6). In mice, at least some V\text{H} segments belonging to the large J\text{H}-distal J558 family are transcribed in unrearranged configuration outside the influence of the H chain enhancer (7). Germline V\text{H}J558 genes are specifically transcribed at the pre-B stage of B cell development. The function of germline V\text{H} transcription is unknown, although the specific expression of germline V\text{H}J558 transcripts in pre-B cells, the cells in which V\text{H} to D-J\text{H} rearrangement occurs, led to the proposal that germline V\text{H} transcription may be related to accessibility of the V\text{H} locus to recombinase enzymes (7).

The large size of the murine J558 family (up to 500–1,000 members; reference 8) has complicated more detailed analyses of germline V\text{H} transcription. Thus, although it is clear that germline V\text{H} transcripts are derived from multiple V\text{H}J558 genes (7), it is difficult to ascertain the exact number and nature of the V\text{H} segments that are expressed. Potential germline transcripts derived from the human V\text{H}5 family have been detected in B chronic lymphoblastic leukemia (B-CLL)1 tumors by Northern analysis (4). In contrast to the murine V\text{H}J558 family, the small size of the human V\text{H}5 family simplifies studies of germline V\text{H} expression. We have

1 Abbreviations used in this paper: ALL, acute lymphoblastic leukemia; BM, bone marrow; CLL, chronic lymphoblastic leukemia; ORF, open reading frame.
therefore further characterized human germline V<sub>H</sub> transcripts in normal and transformed B lineage cells and compared them with murine germline V<sub>H</sub>558 transcripts.

**Materials and Methods**

**Cells and Tissues.** Acutelymphoblastic leukemia (ALL) samples were classified as immature B lineage malignancies (precursor ALL) by standard morphologic and cytochemical criteria (9); three were cell lines (NALL, NALM, and LAZZ) and the remainder were PBMC or bone marrow (BM) samples (>65% lymphoblasts) from ALL patients. B-CLL samples (>98% CD5<sup>+</sup>) were obtained as described (3, 4). EBV-transformed lines were derived from adult PBMC or from fetal BM as described (10). EBV-21 and EBV-321 were transformants of normal CD5<sup>+</sup> cells sorted from adult PBMC (11). Both lines were polyclonal as confirmed by sequencing functionally rearranged V<sub>H</sub>5-C<sub>H</sub><sub>5</sub> transcripts (P.W. Tucker and C. Humphries, unpublished results).

Fetal livers were obtained with permission and Internal Review Board approval from patients undergoing spontaneous or medically indicated therapeutic abortions. Fetal age was estimated by last menstrual period and by fetal foot length. T cell–depleted PBMC were prepared from healthy donors as described (10). Adult spleens were obtained from patients undergoing splenectomy for abdominal trauma. Newborn cord blood was enriched for B cells as described (10).

**RNA Blotting and Preparation.** Preparation of total and poly(A)<sup>+</sup> RNA, fractionation of RNA by formaldehyde/agarose gel electrophoresis, and blotting was done as described (7).

**Preparation of cDNA Libraries.** cDNA, made essentially as described (5). Genomic V<sub>H</sub>5 probes were prepared from 5-1R1 (5) or from V<sub>H</sub>251 (4).

**Preparation of DNA Sequencing.** cDNA clones were sequenced on both strands by the dideoxy chain termination method using M13 universal and reverse oligonucleotide primers (U.S. Biochemical Corp., Indianapolis, IN). Sequences were analyzed using the MicroGenie (Beckman Instruments, Inc., Palo Alto, CA) and DNAStar programs.

**Results**

**Transformed B Lineage Cells Express Novel-sized V<sub>H</sub>5 Gene Transcripts.** Pre-B ALLs have H chain rearrangements but lack slg and thus appear to represent human pre-B cells (9). 10 pre-B ALLs were examined by Northern analysis for expression of the six human V<sub>H</sub> gene families (V<sub>H</sub>1–V<sub>H</sub>6). The ALLs expressed mature V<sub>H</sub>5-C<sub>H</sub>5 H chain mRNA, with individual samples expressing V<sub>H</sub>5 segments from a particular family. In addition, we detected a novel 1.2-kb transcript that hybridized to a V<sub>H</sub>5 probe. Novel V<sub>H</sub>5 transcripts were found in all precursor B ALLs examined, including three cell lines (Fig. 1 A, lanes 1–3) and seven PBMC or BM samples from ALL patients (Fig. 1 A, lanes 4–10). One ALL line (Fig. 1 A, lane 2) also expressed a novel V<sub>H</sub>5 mRNA, slightly

Figure 1. Northern analyses of V<sub>H</sub>5 expression in transformed B lineage cells. (A) Expression of a 1.2-kb germ-line V<sub>H</sub>5 transcript in pre-B ALL cell lines (lanes 1–3) and PBMC or BM from ALL patients (lanes 4–10). (B) Germ-line V<sub>H</sub>5 transcript in B-CLLs. (C and D) Germ-line V<sub>H</sub>5 expression in EBV-transformed B cell lines. B-CLL-12 and LAZZ pre-B ALL RNAs are included for comparison. Mature C<sub>H</sub>5 H chain transcripts (V<sub>H</sub>5-D<sub>H</sub>-J<sub>H</sub>-C<sub>H</sub>5) derived from rearranged V<sub>H</sub>5 genes were detected in some samples; e.g., EBV-21. EBV-21 and EBV-321 are transformants of sorted CDS<sup>+</sup> cells. FB-EBV is a fetal BM line, and Sp-EBV and Br-EBV are PBMC lines (10). Amounts of RNA per lane were: A1-10, 10–20 µg total RNA; B1-4, 5 µg poly(A)<sup>+</sup>; C1-4, 5 µg poly(A)<sup>+</sup>; D1-4, 4 µg poly(A)<sup>+</sup>; D2-4, 5–10 µg total RNA.
greater than 1.2-kb, that is apparently derived from a partially assembled V_{\mu}5-D rearrangement (J.E. Berman and F.W. Alt, unpublished results). Occasionally, transcripts larger than mature H chain mRNA were detected with V_{\mu} family probes other than V_{\mu}5. However, they were not consistently observed in all ALLs and may represent precursor mRNAs (data not shown).

Murine germline V_{\mu} transcripts are expressed in pre-B lines and are not observed in more mature B lineage tumors such as B cell lymphomas and myelomas (7). In human slg + CLLs, V_{\mu}5 transcripts similar in size to the 1.2-kb species found in pre-B ALLs were detected by Northern blotting (4). Therefore, we examined transformed B lineage cells other than pre-B ALLs for V_{\mu}5 expression. The novel 1.2-kb V_{\mu}5 transcript was observed in some B-CLLs and EBV-transformed B lymphoblastoid cell lines, but in contrast to the pre-B ALLs, it was not found in all such lines (representative data are shown in Fig. 1, B, C, and D). As with the ALLs, germline transcripts of other V_{\mu} families were not observed (data not shown).

The Novel-sized 1.2-kb V_{\mu}5 Species Are Germline Transcripts. cDNA libraries were constructed from several ALL, CLL, and EBV-transformed B cell samples tested above, V_{\mu}5-hybridizing clones were randomly isolated, and complete nucleotide sequences of two of the longest cDNAs (L3-4, L2-9) from a pre-B ALL library (LAZZ; Fig. 1 A) were determined. We also sequenced three V_{\mu}5 + cDNAs randomly isolated from two B CLL samples (CLL12, CLL27) and from an EBV-transformed B cell line (EBV-21) that also expressed the 1.2-kb V_{\mu}5 transcript (Fig. 1, B and C). Comparison of

---

**Figure 2.** Nucleotide sequences of germline cDNAs derived from the functional V_{\mu}5 gene. Genomic sequences of unrearranged (V_{\mu}251; 4) and rearranged (5-1R1; 5) functional V_{\mu}5 genes are provided for comparison. Dashes indicate identity with the V_{\mu}251 sequence; differences are indicated. cDNA sequences are condensed into a single line in regions of complete identity (1, L3-4; 2, CB-4; 3, EBV-21; 4, CLL-12). The leader is spliced in the cDNAs so the genomic intron sequences are not shown. Corrections have been made in the originally reported V_{\mu}251 sequence (4) immediately 3' of the nonamer. The sequence of V_{\mu}251 further downstream of the nonamer is previously unpublished. Heptamer and nonamer recombination sequences are boxed. The internal "V replacement" heptamer at the end of framework 3 is underlined. The asterisk and line above the nonamer indicate the end of the ORE Polyadenylation signals are underlined in the untranslated region 3' of the nonamer.

---

1531 Berman et al.
these nucleotide sequences to genomic germline V_{H}5 sequences (4, 5) confirms that they represent germline transcripts of unarranged V_{H}5 genes (Figs. 2 and 3). The cDNAs extend from 5 to 39 bp upstream of the leader, through the coding regions, and, in contrast to rearranged V_{H} gene transcripts, include the normally deleted recombination signals and 3' flanking regions.

**Germline Transcripts Are Expressed from both Functional and Pseudo-V_{H}5 Genes.** As expected, some cDNAs were derived from the functional V_{H}5 gene (V_{H}5(21), Fig. 2), but surprisingly, others were transcripts of the V_{H}5 pseudogene (V_{H}5(15/ V_{H}5(1-V; Fig. 3)). To determine the relative steady-state levels of these two transcripts and to search for other germline V_{H}5 mRNA species, 10 V_{H}5* clones from the LAZZ pre-B ALL (12; Fig. 1 A, lane 3) cDNA library were randomly characterized. Half of the clones were derived from the functional V_{H}5 gene and the others were transcripts of the V_{H}5 pseudogene. 9 of the 10 cDNA clones were from similar regions of the V_{H}5 genes; however, one pseudogene-derived cDNA, L6-1, extended further upstream to ~420 bp 5' of the initiation codon and ended ~0.1 kb 3' of the nonamer (data not shown).

**Structure of Germline V_{H}5 Transcripts.** Comparison of the V_{H}5 cDNAs (including L6-1) and germline genomic V_{H}5 sequences (V_{H}5(251) coding region, reference 4; newly determined V_{H}251 3' flank, Fig. 2; 1-V, reference 5) indicated that the leader intron was appropriately spliced but no additional splicing occurred in the 3' flanking regions. The germline V_{H}5 transcripts are polyadenylated, although there is apparently slight variation in the site of pre-RNA 3' end cleavage and polyadenylation since some cDNAs such as L3-4 and L2-9 extend beyond the beginning of the poly(A) tails present in CLL12 and CLL27 (Figs. 2 and 3). This variation is not surprising because there are multiple consensus polyadenylation/cleavage signals (reviewed in reference 13) at the 3' ends of both the functional and pseudo-V_{H}5 genes (underlined in Figs. 2 and 3).

Together, the functional germline V_{H}5 cDNAs encompass 967 bp of sequence and the pseudo-V_{H}5 cDNAs span a 1,001-bp region. Assuming an ~250-base poly(A) tail (13), the cDNA clones correspond to the full-length 1.2-kb transcript. This suggests that transcriptional initiation of the majority of V_{H}5 germline mRNA sequences occurs within 100 bp upstream of the leader, probably from the normal V_{H} promoter; although the L6-1 cDNA indicates that there is also some transcriptional initiation upstream of the V_{H} promoter, at least for the pseudo-V_{H}5 gene.

Although from different individuals, the coding region sequences of the four functional germline V_{H}5 cDNAs and the corresponding germline gene (V_{H}251) are highly conserved.

---

**Figure 3.** Nucleotide sequences of germline cDNAs derived from the V_{H}5 pseudogene. Two versions of the genomic sequence are provided for comparison (1-V, 5, V_{H}15, 4). cDNA sequences are condensed into a single line in regions of complete identity (1, L2-9; 2, CLL27). An error in the published 1-V sequence has been corrected (indicated by the asterisk at the beginning of FR1) and now yields an ORF throughout the coding region. The nonamer recombination sequence is boxed. The internal heptamer is underlined. The asterisk and line above the nonamer indicate the end of the ORF. Polyadenylation signals in the 3' untranscribed region are underlined. The sequence data in Figs. 2 and 3 will be available from EMBL/GenBank/DDBJ under accession numbers X58397-X58402.
in the 3' flanking regions of the functional VH5 sequences, unexpectedly, a high degree of conservation is also found in limited VH251 coding region polymorphism (14). This is consistent with previous findings of limited VH251 coding region polymorphism (14).

Discussion

Limited Polymorphism among Germline VH5 cDNAs. We have demonstrated that unrearranged functional and pseudo-VH5 gene segments are transcribed in normal and transformed B lineage cells. Although germline transcripts of the poly-

Figure 4. Germline transcripts of functional and pseudo-VH5 genes are highly related. The coding regions of functional (L3-4) and pseudo-VH5 (L2-9) cDNAs (isolated from LAZZ, see text) are >98% similar (only FR3 is shown), and the 3' flanking regions are >98% similar throughout both the coding and flanking regions (Fig. 3).

Open Reading Frames (ORF) Are Found in the Germline VH5 cDNAs. In the germline cDNAs derived from the functional VH5 gene, an ORF extends throughout the VH coding region and ends in a termination codon within the nonamer recombination sequence (see asterisk in Fig. 2). Beyond the nonamer there are multiple stop codons in all reading frames. A search of GenBank did not reveal significant homologies between this untranslated region and other human or murine genes. The germline cDNAs potentially encode a VH region the same as that of a rearranged gene, except for 11 COOH-terminal residues (His-Thr-Val-Arg-Pro-Val) identical to the heptamer/23-bp spacer/nonamer region of all the functional VH5 cDNAs (Fig. 2). A similar ORF was also found in the pseudo-VH5 cDNAs, again terminating within the nonamer, but only three additional amino acids (Gly-Thr-Ile) are encoded beyond FR3 due to a deletion within the recombination sequences (Figs. 3 and 4; see Discussion).

Identical Germline VH5 Transcripts Are Expressed in Normal Tissues. Characteristic 1.2-kb germline VH5 transcripts were observed on Northern analyses of fetal liver (15 and 17 wk) and cord blood, but were not seen in adult T cell-depleted PBMC or in adult spleen (Fig. 5, A and B; confirmed by S1 analysis, data not shown). The sequence of a germline VH5 clone (CB-4) isolated from a cDNA library of normal cord blood (Fig. 5 B, lane 2) was essentially identical to functional VH5 cDNAs isolated from transformed B lineage cells (Fig. 2). As with the transformed lines, we did not detect germline-length transcripts with VH family probes other than VH5 (data not shown).

![Figure 5](image-url) Germline VH5 expression in normal tissues. (A) Germline VH5 expression in T cell-depleted PBMC (lanes 1-3) and fetal liver (15 wk, lane 4; 17 wk, lane 5). RNA from the LAZZ pre-B ALL line is shown for comparison. (B) Germline VH5 expression in normal cord blood (lane 2). RNA from B CLL/12 is shown for comparison. Complete VH5-Cc H chain mRNA was detected in fetal liver and in cord blood, and both VH5-Cc and VH5-Cc mRNA were seen in adult PBMC. Amounts of RNA per lane were: A1-3, 10 μg total RNA; A4-5, 6 μg poly(A)*; A6, 4 μg poly(A)*; B1-2, 5 μg poly(A)*.

(>98% similarity; Fig. 2). This is consistent with previous findings of limited VH251 coding region polymorphism (14). Unexpectedly, a high degree of conservation is also found in the 3' flanking regions of the functional VH5 sequences, all of which are >98% similar. Even the pseudo-VH5 sequences are >98% similar throughout both the coding and flanking regions (Fig. 3).

Discussion

Limited Polymorphism among Germline VH5 cDNAs. We have demonstrated that unrearranged functional and pseudo-VH5 gene segments are transcribed in normal and transformed B lineage cells. Although germline transcripts of the poly-

![Figure 4](image-url) Germline transcripts of functional and pseudo-VH5 genes are highly related. The coding regions of functional (L3-4) and pseudo-VH5 (L2-9) cDNAs (isolated from LAZZ, see text) are >98% similar (only FR3 is shown), and the 3' flanking regions are >98% similar throughout both the coding and flanking regions (Fig. 3).

Open Reading Frames (ORF) Are Found in the Germline VH5 cDNAs. In the germline cDNAs derived from the functional VH5 gene, an ORF extends throughout the VH coding region and ends in a termination codon within the nonamer recombination sequence (see asterisk in Fig. 2). Beyond the nonamer there are multiple stop codons in all reading frames. A search of GenBank did not reveal significant homologies between this untranslated region and other human or murine genes. The germline cDNAs potentially encode a VH region the same as that of a rearranged gene, except for 11 COOH-terminal residues (His-Thr-Val-Arg-Pro-Val) identical encoded within the heptamer/23-bp spacer/nonamer region of all the functional VH5 cDNAs (Fig. 2). A similar ORF was also found in the pseudo-VH5 cDNAs, again terminating within the nonamer, but only three additional amino acids (Gly-Thr-Ile) are encoded beyond FR3 due to a deletion within the recombination sequences (Figs. 3 and 4; see Discussion).

Identical Germline VH5 Transcripts Are Expressed in Normal Tissues. Characteristic 1.2-kb germline VH5 transcripts were observed on Northern analyses of fetal liver (15 and 17 wk) and cord blood, but were not seen in adult T cell-depleted PBMC or in adult spleen (Fig. 5, A and B; confirmed by S1 analysis, data not shown). The sequence of a germline VH5 clone (CB-4) isolated from a cDNA library of normal cord blood (Fig. 5 B, lane 2) was essentially identical to functional VH5 cDNAs isolated from transformed B lineage cells (Fig. 2). As with the transformed lines, we did not detect germline-length transcripts with VH family probes other than VH5 (data not shown).

Discussion

Limited Polymorphism among Germline VH5 cDNAs. We have demonstrated that unrearranged functional and pseudo-VH5 gene segments are transcribed in normal and transformed B lineage cells. Although germline transcripts of the poly-

![Figure 4](image-url) Germline transcripts of functional and pseudo-VH5 genes are highly related. The coding regions of functional (L3-4) and pseudo-VH5 (L2-9) cDNAs (isolated from LAZZ, see text) are >98% similar (only FR3 is shown), and the 3' flanking regions are >98% similar throughout both the coding and flanking regions (Fig. 3).

Open Reading Frames (ORF) Are Found in the Germline VH5 cDNAs. In the germline cDNAs derived from the functional VH5 gene, an ORF extends throughout the VH coding region and ends in a termination codon within the nonamer recombination sequence (see asterisk in Fig. 2). Beyond the nonamer there are multiple stop codons in all reading frames. A search of GenBank did not reveal significant homologies between this untranslated region and other human or murine genes. The germline cDNAs potentially encode a VH region the same as that of a rearranged gene, except for 11 COOH-terminal residues (His-Thr-Val-Arg-Pro-Val) identical encoded within the heptamer/23-bp spacer/nonamer region of all the functional VH5 cDNAs (Fig. 2). A similar ORF was also found in the pseudo-VH5 cDNAs, again terminating within the nonamer, but only three additional amino acids (Gly-Thr-Ile) are encoded beyond FR3 due to a deletion within the recombination sequences (Figs. 3 and 4; see Discussion).

Identical Germline VH5 Transcripts Are Expressed in Normal Tissues. Characteristic 1.2-kb germline VH5 transcripts were observed on Northern analyses of fetal liver (15 and 17 wk) and cord blood, but were not seen in adult T cell-depleted PBMC or in adult spleen (Fig. 5, A and B; confirmed by S1 analysis, data not shown). The sequence of a germline VH5 clone (CB-4) isolated from a cDNA library of normal cord blood (Fig. 5 B, lane 2) was essentially identical to functional VH5 cDNAs isolated from transformed B lineage cells (Fig. 2). As with the transformed lines, we did not detect germline-length transcripts with VH family probes other than VH5 (data not shown).

Discussion

Limited Polymorphism among Germline VH5 cDNAs. We have demonstrated that unrearranged functional and pseudo-VH5 gene segments are transcribed in normal and transformed B lineage cells. Although germline transcripts of the poly-

![Figure 5](image-url) Germline VH5 expression in normal tissues. (A) Germline VH5 expression in T cell-depleted PBMC (lanes 1-3) and fetal liver (15 wk, lane 4; 17 wk, lane 5). RNA from the LAZZ pre-B ALL line is shown for comparison. (B) Germline VH5 expression in normal cord blood (lane 2). RNA from B CLL/12 is shown for comparison. Complete VH5-Cc H chain mRNA was detected in fetal liver and in cord blood, and both VH5-Cc and VH5-Cc mRNA were seen in adult PBMC. Amounts of RNA per lane were: A1-3, 10 μg total RNA; A4-5, 6 μg poly(A)*; A6, 4 μg poly(A)*; B1-2, 5 μg poly(A)*.
The sequences of the functional and pseudo-V<sub>n</sub>5 transcripts are highly conserved in the 3' flank as well as in the coding regions. It has been suggested that the limited polymorphism seen in coding regions of the smaller V<sub>n</sub> gene families (V<sub>n</sub>4, -5, -6) may be related to critical antigen specificities encoded by these genes (14, 15). This explanation does not account for the high degree of conservation among the pseudo-V<sub>n</sub>5 coding region sequences or among the 3' flanking regions of both the functional and pseudo-V<sub>n</sub>5 genes. Such conservation could be due to recent emergence and duplication of the V<sub>n</sub>5 family, but may also be explained by the fact that these regions are expressed in germ-line transcripts. Conservation of the 3' untranslated regions could be due to a functional role such as regulating mRNA stability. The V<sub>n</sub>5 pseudogene is likely to be nonfunctional with regard to VDJ rearrangement due to its abnormal recombination sequences; however, conservation of the ORF in the pseudo-V<sub>n</sub>5 gene might reflect a function of an encoded germline V<sub>n</sub>5 protein.

3' Flanking Regions of Functional and Pseudo-V<sub>n</sub>5 Germline cDNAs. The functional and pseudo-V<sub>n</sub>5 cDNA sequences are highly homologous throughout the coding (91% similar) and 3' flanking regions (86% similar). A comparison of the 3' flanking regions of functional (L3-4) and pseudo-V<sub>n</sub>5 (L2-9) cDNAs reveals two major regions of divergence: ~0.16 kb 3' of the nonamer and immediately 3' of the coding regions (Fig. 4). A 24-bp deletion in the immediate 3' flank of the pseudogene results in loss of the heptamer and part of the spacer region, thus placing the nonamer 12 bp 3' of the internal 'V replacement' heptamer (16, 17). Therefore, on the basis of the 12/23 joining rule (reviewed in reference 1) there could be direct rearrangement between the V<sub>n</sub>5 pseudogene (12-bp spacer) and J<sub>n</sub> segments (23-bp spacer). As yet there is no evidence for direct pseudo-V<sub>n</sub>5 to J<sub>n</sub> joining perhaps due to the orientation of the internal heptamer relative to the nonamer. Although it may be a coincidence, it is notable that the 24-bp deletion in the pseudogene does not shift the reading frame; thus, the translational site in the nonamer is conserved.

Comparison of Murine and Human Germline V<sub>n</sub> Transcription. Several aspects of human and murine germline V<sub>n</sub> transcription are similar. Although mouse and human germ-line V<sub>n</sub> transcripts vary in length, murine germline V<sub>n</sub>5 cDNAs have an ORF similar in size to that of the human germline V<sub>n</sub>5 transcripts, also ending in a termination codon within the recombination signals (but in the 23-bp spacer region rather than in the nonamer; reference 7). Thus, murine and human germline V<sub>n</sub> transcripts potentially encode truncated V<sub>n</sub> proteins of a similar size. Transcription of only one V<sub>n</sub> family has been readily found in humans (V<sub>n</sub>5) as well as in mice (J558). Although the V<sub>n</sub>5 family has no obvious murine counterpart (i.e., >70% nucleic acid similarity), V<sub>n</sub>5 J558 members are the most closely related. Of course germline transcripts from other murine and human V<sub>n</sub> gene families may exist at undetected, lower steady-state levels. In mice and humans, most germline V<sub>n</sub> transcription initiates at the normal V<sub>n</sub> promoter, although more 5' initiation has been seen with V<sub>n</sub> pseudogenes that still retain ORFs (see above and reference 7).

Specificity of Germline V<sub>n</sub>5 Expression. The consistent expression of germline V<sub>n</sub>5 in pre-B ALLs may provide an additional diagnostic marker for disease in tissues such as PBMC, which normally lack detectable germline V<sub>n</sub> mRNA. In contrast to the pre-B specificity of murine germline V<sub>n</sub> expression, we also detected human germline V<sub>n</sub>5 transcripts in more mature B lineage cells. Germline V<sub>n</sub>5 expression in some B-CLLs and in some EBV-transformed B cell lines might argue against the proposal that germline V<sub>n</sub>5 transcription reflects processes that regulate V<sub>n</sub> to D-J<sub>n</sub> rearrangement in pre-B cells. On the other hand, germline V<sub>n</sub>5 expression beyond the pre-B cell stage may be unique to transformed cells and may reflect the difficulty of staging human B lineage tumor cells. Germline V<sub>n</sub>5 expression may also be restricted to certain B cell subpopulations. Germline V<sub>n</sub>5 mRNA was observed in fetal liver, but was not detected in PBMC or in adult spleen cells, suggesting that normal mature B cells do not express germline V<sub>n</sub>5. Nevertheless, germline V<sub>n</sub>5 transcripts may be expressed in some normal CD5<sup>+</sup> B cells based on their presence in some CD5<sup>+</sup> B-CLLs and in cord blood and fetal liver, tissues that are rich in CD5<sup>+</sup> cells (18, 19). We saw no evidence for germline V<sub>n</sub>5 transcripts analogous to the unusual species found in human T cells, consisting of a novel 5' exon spliced to an unrearranged V<sub>n</sub>4 gene (20).

Function of Germline V<sub>n</sub>5 Transcription. Germline V<sub>n</sub>5 expression may regulate accessibility of the V<sub>n</sub> locus during H chain rearrangement; alternatively, truncated V<sub>n</sub> proteins encoded by germline transcripts may play a role in B cell development (7). Germline V<sub>n</sub>5 cDNAs presumably possess the same 5' signals required by mature H chain transcripts for translation since the transcription start sites seem similar. Furthermore, in vitro transcription and translation of the germline V<sub>n</sub>5 cDNAs demonstrated that the transcripts are translatable into a protein consistent with the size (13.9 kD) of the predicted ORF (J. E. Berman and F.W. Alt, unpublished results). These putative truncated V<sub>n</sub>5 H chain proteins are reminiscent of other members of the Ig gene superfamily, such as Vpre-B (21), and would consist of a single Ig domain possessing the two cysteine residues necessary for formation of the characteristic intra-chain disulfide-bonded loop.

Conservation of germline V<sub>n</sub> gene transcription in humans as well as in mice suggests an important functional significance. The complexity of the J558 family has made it difficult to ascertain the heterogeneity of murine germline transcripts and to identify the corresponding genomic germline genes that are transcribed. Detectable germline J558 expression in mice could be the result of low level transcription...
of multiple genes and/or high level expression of a limited subset of the large J558 family. The relative simplicity of the human \(V_{\alpha}5\) family allows us to conclude that human germ-line transcripts derived from just two unrearranged \(V_{\alpha}\) genes are readily detectable and their genomic counterparts have already been isolated and are available for further analysis. This should facilitate future study of factors such as associated regulatory regions, which may control germ-line \(V_{\alpha}\) gene transcription.

This work was supported by the Howard Hughes Medical Institute and by grants from the National Institutes of Health (CA-44016, CA-40427, and AI-20047). J. Berman was a fellow of the Howard Hughes Medical Institute.

Address correspondence to Frederick W. Alt, Howard Hughes Medical Institute Research Laboratories, Columbia University College of Physicians and Surgeons, 701 West 168th Street, New York, NY 10032. J.E. Berman's present address is the Medical Biotechnology Center and the Department of Microbiology and Immunology, University of Maryland, 655 West Baltimore Street, Baltimore, MD 21201.

Received for publication 5 November 1990 and in revised form 11 March 1991.

References

1. Alt, F.W., T.K. Blackwell, R.D. DePinho, M.G. Reth, and G.D. Yancopoulos. 1986. Regulation of genome rearrangement events during lymphocyte differentiation. Immunol. Rev. 89:5.

2. Berman, J.E., and F.W. Alt. 1990. Human H-chain variable region gene diversity, organization, and expression. Int. Rev. Immunol. 5:203.

3. Shen, A., C. Humphries, P.W. Tucker, and F.R. Blattner. 1987. Human H-chain V region gene family non-randomly rearranged in familial chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA. 84:8563.

4. Humphries, C.G., A. Shen, W.A. Kuziel, J.D. Capra, F.R. Blattner, and P.W. Tucker. 1988. A new human Ig \(V_{\alpha}\) family preferentially rearranged in immature B-cell tumours. Nature (Lond.). 331:446.

5. Berman, J.E., S.J. Mellis, R. Pollock, C.L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C.R. Cantor, and F.W. Alt. 1988. Content and organization of the human Ig \(V_{\alpha}\) locus: definition of three new \(V_{\alpha}\) families and linkage to the Ig \(C_{\alpha}\) locus. EMBO (Eur. Mol. Biol. Organ.) J. 7:727.

6. Sen, R., and D. Baltimore. 1989. Factors regulating Ig-gene transcription. In Immunoglobulin Genes. T. Honjo, F.W. Alt, and T. Rabbitts, editors. Academic Press, San Diego, CA. 327–342.

7. Yancopoulos, G.D., and F.W. Alt. 1985. Developmentally controlled and tissue-specific expression of unrearranged \(V_{H}\) gene segments. Cell. 40:271.

8. Livant, D., C. Blatt, and L. Hood. 1986. One H-chain V region gene segment subfamily in the BALB/c mouse contains 500-1000 or more members. Cell. 47:461.

9. Poons, K.A., and R.F. Todd. 1986. Immunologic classification of leukemia and lymphoma. Blood. 68:1.

10. Nickerson, K.G., J.E. Berman, E. Glickman, L. Chess, and F.W. Alt. 1989. Early human IgH gene assembly in Epstein-Barr virus-transformed fetal B cell lines. Preferential utilization of the most \(I_{c}\)-proximal D segments (DQ52) and two unusual \(V_{c}\)-related rearrangements. J. Exp. Med. 169:1391.

11. Nakamura, M., S.E. Burastero, A.L. Notkins, and P. Casali. 1988. Human monoclonal rheumatoid factor-like antibodies from CD5 (Leu-1)\(^{-}\) B cells are polyreactive. J. Immunol. 140:4180.

12. Berman, J.E., K.G. Nickerson, R.R. Pollock, J.E. Barth, R.K.B. Schuurman, D.M. Knowles, L. Chess, and F.W. Alt. \(V_{\alpha}\) Gene usage in humans: biased usage of the \(V_{\alpha}6\) gene in immature B lymphoid cells. Eur. J. Immunol. In press.

13. Jackson, R.J., and N. Standart. 1990. Do the poly(A) tail and 3' untranslated region control mRNA translation? Cell. 26:15.

14. Sanz, I., P. Kelly, C. Williams, S. Scholl, P. Tucker, and J.D. Capra. 1989. The smaller human \(V_{\alpha}\) gene families display remarkably little polymorphism. EMBO (Eur. Mol. Biol. Organ.) J. 8:3741.

15. Van Dijk, K.W., H.W. Schroeder, Jr., R.M. Perlmutter, and E.C.B. Milner. 1989. Heterogeneity in the human Ig \(V_{\alpha}\) locus. J. Immunol. 142:2547.

16. Reth, M., P. Gehrmann, E. Petrac, and P. Wiese. 1986. A novel \(V_{\alpha}\) to \(V_{\delta}D_{\gamma}\) joining mechanism in H-chain-negative (null) pre-B cells results in H-chain production. Nature (Lond.). 322:840.

17. Kleinfield, R., R. Hardy, D. Tarlinton, J. Dangl, L. Herzenberg, and M. Weigert. 1986. Recombination between an expressed Ig H-chain gene and a germine Ig V gene in a Ly-1\(^{-}\) B-cell lymphoma. Nature (Lond.). 322:843.

18. Antin, J.H., S.G. Emerson, P. Martin, N. Gadol, and K.H. Ault. 1986. Leu-1\(^{-}\) (CD5\(^{+}\)) B cells. A major lymphoid sub-population in human fetal spleen: phenotypic and functional studies. J. Immunol. 136:505.

19. Hardy, R.R., K. Hayakawa, M. Shimizu, K. Yamashiki, and T. Kishimoto. 1987. Rheumatoid factor secretion from human Leu-1\(^{-}\) B cells. Science (Wash. DC). 236:81.

20. Baer, R., A. Forster, I. Lavenir, and T.H. Rabbitts. 1988. Ig \(V_{\alpha}\) genes are transcribed by T cells in association with a new 5' exon. J. Exp. Med. 167:2011.

21. Kudo, A., and F. Melchers. 1987. A second gene, VpreB in the \(\lambda 5\) locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. EMBO (Eur. Mol. Biol. Organ.) J. 6:2267.