

ANALYSIS OF A NOVEL V_HS107 HAPLOTYPE
IN CLA-2 AND WSA MICE

Evidence for Gene Conversion Among IgV_H Genes
in Outbred Populations

BY STACY E. FERGUSON,*† MICHAEL P. CANCRO,†
AND BARBARA A. OSBORNE*§

*From the *Program in Molecular and Cellular Biology, University of Massachusetts, Amherst, Massachusetts 01003; the †Immunobiology Division, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104; and the §Department of Veterinary and Animal Sciences, University of Massachusetts, Amherst, Massachusetts 01003*

Comparative analyses of Ig variable region genes provide a means for studying evolutionary mechanisms active in large multigene families. The Ig heavy chain V-region (V_H) gene families of inbred laboratory mice have been examined in detail, and several potential evolutionary mechanisms have been inferred from these studies. Among these, gene conversion has been suggested as the basis for many V_H allelic differences, particularly in the V_HS107 family (1-8). Gene conversion within this family was initially suggested based on the cDNA sequence of a CBA/J hybridoma (6), and subsequent examination of CBA/J germline genes indicated the occurrence of gene conversion in all four members of the V_HS107 family (7). Additionally, gene conversion events have been postulated to explain the differences between the V_HS107 haplotypes of laboratory inbred strains (8).

Whether gene conversion occurs with significant frequency in outbred populations has not been addressed directly. Extrapolation from the existing data is constrained by both the small group of progenitors common to most inbred laboratory strains, and the possibility that long-term homozygosity fosters events that otherwise rarely occur. Thus, the examination of homologous gene families in feral populations may provide insight regarding this issue.

A wild population of *Mus musculus domesticus* exhibiting low immunologic responsiveness to phosphorylcholine (PC)¹ was described by Lieberman et al. (9), and several inbred lines have been independently derived from this population. These include the WSA (10; this article) and CLA-2/Cn strains (11). In recent studies, the nucleotide sequence of heavy chain cDNA from a PC-specific hybridoma suggested gene conversion events within the CLA-2/Cn V_HS107 family (11). Since the CLA-2/Cn and WSA strains were recently and independently inbred, germline sequences

This work was supported in part by U.S. Public Health Service research grants GM-36344 and AI-20450, and Research Career Development Award AI-00550.

Address correspondence to Dr. Stacy E. Ferguson, University of Pennsylvania School of Medicine, Department of Pathology and Laboratory Medicine, Philadelphia, PA 19104-6082.

¹Abbreviations used in this paper: PC, phosphorylcholine.

consistent with conversion would strengthen the possibility that such events are widespread and unlikely artifacts of inbreeding. Accordingly, we have examined the V_HS107 families of the CLA-2/Cn and WSA strains. The results indicate that: (a) The CLA-2 and WSA strains bear identical but novel V_HS107 family haplotypes; (b) low PC responsiveness in these populations is unlikely due to an inability to express the V1 member of the V_HS107 gene family; and (c) gene conversion events have occurred with considerable frequency in the evolution of the murine V_HS107 family, especially among the V3, V13, and V11 members.

Materials and Methods

Mice. The CLA-2/Cn strain was derived in the laboratory of M. P. Cancro, from founders originally obtained from Dr. M. Potter (Laboratory of Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD; under the auspices of NIH contract NZ01-CB-71085). These mice were selectively inbred for PC low responsiveness, and this segregates as a single autosomal Mendelian trait (11). Adult breeding pairs of the WSA line were also obtained from Dr. M. Potter. The progenitors of the CLA and WSA lineages were members of an outbred population, and were initially separated based on coat color and marking traits before inbreeding regimes (9, 10).

DNA Preparation. High molecular weight DNA was prepared using adult female animals, freshly sacrificed by cervical dislocation. Nuclei were prepared according to the method of Liou et al. (12), with some modifications. Briefly, livers were immediately removed, minced, and dounce-homogenized in ice-cold PBS, containing 1 mM PMSF. All subsequent steps were performed on ice. The homogenate was strained through two layers of cotton gauze to remove connective tissue. Cells were spun down and washed several times in fresh PBS/PMSF. The pellet was resuspended in 25 ml NIB (60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 1 mM EGTA, 15 mM Tris-HCl, pH 7.6, 0.5 mM DTT, 1 mM PMSF), containing 0.3 M sucrose.

Cell membranes were solubilized by the addition of 0.3 ml 10% NP-40 detergent. Aliquots of 12.5 ml were gently layered over 12.5 ml sucrose cushions (1.7 M sucrose in NIB). Nuclei were spun down at 4°C for 15 min (13,000 rpm, HB-4 rotor, RC-5B centrifuge, Beckman Instruments, Fullerton, CA). The supernatant was discarded and the pellet was resuspended in a total volume of 5.0 ml NIB. Nuclei were lysed by the addition of 0.5 ml 5% SDS/100 mM EDTA and proteinase K to a final concentration of 0.5 mg/ml. Tubes were incubated overnight at 37°C.

The DNA preparation was extracted once in an equal volume of phenol, previously equilibrated with TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA), extracted twice with phenol/CHCl₃ (2/1, vol/vol) and once with CHCl₃. DNA was precipitated from the aqueous phase by the addition of 2 vol 100% ethanol, spooled, air-dried, and resuspended to a final concentration of ~1 mg/ml.

Individual liver DNA samples from 13 members of the original breeding stocks from which the CLA-2/Cn and WSA strains were derived were generously provided by Drs. L. D'Hoos-telaere and M. Potter.

Genomic Libraries. Phage genomic libraries were prepared through two different protocols. Partial Sau3a I digests were performed and fragments ranging from 9 to 23 kb in size were isolated by fractionation over sucrose gradients (13). Appropriate fractions were pooled, precipitated, and ligated into the Bam HI-digested λ phage vector, λEMBL3 (Stratagene, La Jolla, CA) according to supplier's instructions. Ligations were packaged with commercial extracts (Stratagene).

An alternative method for isolation of individual V_HS107 gene segments was also used. Genomic DNA was digested to completion with Eco RI and electrophoresed through 0.8% agarose slab gels. Regions containing V_HS107 gene segments, as determined by Southern blot analysis, were excised. Gel slices were placed in dialysis tubing for electroelution of fragments, which were precipitated and ligated into the vector λgt10 (Stratagene). Reactions were packaged as described for EMBL3 libraries.

Library Screening. λ EMBL3 and λ gt10 libraries were plated and plaque-lifts were performed in duplicate using nylon membranes (Hybond-N; Amersham Corp., Arlington Heights, IL). Lifts were denatured by floating on 0.2 M NaOH/1.5M NaCl. Filters were air-dried and baked for 2 h at 80°C. Lifts were washed for several hours with frequent changes of a wash solution containing 50 mM Tris-HCl, pH 8.0, 0.1% SDS, 1.0 mM EDTA, and 1.0 M NaCl at 42°C. Membranes were transferred to 50% formamide buffer for overnight hybridization.

Probes. A cDNA probe containing the BALB/c V1 gene segment was originally obtained from Dr. S. Crews (14); California Institute of Technology, Pasadena, CA). A Pst I fragment, containing the V1 coding region, was excised for use as a hybridization probe. A plasmid containing a rearranged V1 gene segment as well as associated 5'-sequences, was obtained from P. Brodeur (Department of Pathology, Tufts University, Boston, MA). A 1.7-kb Bam HI fragment containing V1-specific sequence was used to identify upstream sequences of that gene segment.

Probes were labeled with α -[³²P]dCTP using the random-priming method of Feinberg and Vogelstein (15). A commercial kit was used (Oligolabeling Kit; Pharmacia Fine Chemicals, Piscataway, NJ).

M13 Subclones. A minimum of two λ phage clones from original, unamplified libraries were sequenced per gene segment. All M13 manipulations were performed as described by Messing (16). Plaque-lifts and screenings were performed as described earlier for genomic library screenings, except that filters were neutralized after denaturation. Subcloning strategies were based on the restriction maps of BALB/c S107 V_H phage clones described by Siu et al. (17).

The V13 gene segments of the CLA-2/Cn and WSA strains were subcloned as 2.7-kb Eco RI fragments into Eco RI-digested M13mp18. Clones were also generated by digestion with Sau3a I and shotgun cloning of resulting fragments into Bam HI-digested vector.

The V1 gene segments of both strains were subcloned as 2.5-kb Bam HI fragments into Bam HI-digested vector. Larger, 5-kb Xba I fragments were also subcloned for flanking region analysis.

The V11 gene segments were force cloned into both orientations by double-digestion with the enzymes, Eco RI and Hind III. These fragments were cloned into the vectors M13mp18 and M13mp19. The novel V11 gene segments (V11-2) of the two strains were subcloned as 4.7-kb Eco RI fragments into Eco RI-digested M13mp18.

Orientations of inserts in all M13 subclones were determined either directly, through sequence analysis or through complementation testing (16).

Dideoxynucleotide Sequencing. Sequencing of M13 templates was performed with a commercial sequencing kit using supplier's instructions (Sequenase; United States Biochemical Corp., Cleveland, OH).

Oligonucleotide Primers. All genes were sequenced in two orientations using the M13 universal primer, as well as a set of primers designed to anneal to particular regions of the coding and flanking sequences.

Results

S107 VH Allele Assignments. To examine the contribution of gene conversion to allelism among V_HS107 haplotypes, it is first necessary to distinguish the V1, V11, V13, and V3 family members. Family member assignment was thus based on three criteria. First, genomic Southern blot analysis of Eco RI-digested CLA-2/Cn and WSA DNA reveals the presence of four V_HS107 gene segments (Fig. 1, last four lanes), three of which comigrate with V_HS107 elements in BALB/c (V1, V11, and V13), rendering preliminary family member assignment feasible. DNA from 13 members of the outbred CLA and WSA progenitors are also included in Fig. 1, and show that at least three V_HS107 haplotypes were segregating in this population. Second, coding region sequence analysis was used to distinguish the V1 element

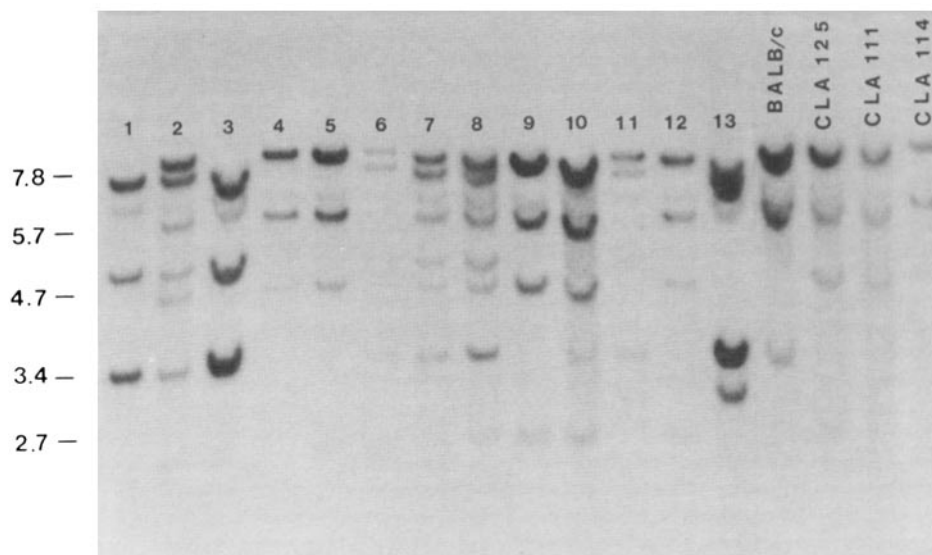


FIGURE 1. Southern Blot analysis of Eco RI-digested BALB/c, CLA-2/Cn, and outbred progenitor DNA. 7.5 μ g per lane of Eco RI-digested genomic DNA was electrophoresed, transferred to nylon filters, and hybridized to the p107V1 probe (14). Individual liver DNA samples from members of the outbred progenitor population of CLA-2/Cn and WSA are shown in lanes 1-13. BALB/c and three individual CLA-2/Cn DNA samples are in the last four lanes.

from the other three gene segments. Finally, noncoding, flanking region sequence analyses were performed. While the coding regions of the V11, V13, and V3 gene segments exhibit a high degree of homology, the flanking, noncoding sequences have diverged considerably, allowing for unambiguous family member assignment of novel $V_H S107$ sequences.

The CLA-2/Cn and WSA V13 Gene Segments are Pseudogenes. As in the BALB/c strain, the V13 gene segments of WSA and CLA-2/Cn lie on 2.8-kb Eco RI fragments. Alleles of V13 examined to date are believed to be functional based on nucleotide sequence analysis (14, 17, 18).

Subcloning of the CLA-2/Cn and WSA V13 gene segments revealed the presence of a novel *Sau* 3a restriction site within the V_H coding region. Subcloning and sequencing of the gene on the larger, 2.8-kb Eco RI fragment revealed that the novel *Sau* 3a site is attributable to a single nucleotide insertion within codon 37 (GTC to GATC), resulting in a shift in reading frame (Fig. 2). The new reading frame includes four translational stop codons, likely rendering the V13 gene segments of the CLA-2/Cn and WSA strains pseudogenes. A second mutation is present at codon 82c. This defect involves a single base deletion (CTG to TG) that restores the original reading frame. The defects present in these V13 gene segments are not cloning artifacts, as libraries from the two strains were generated independently. Furthermore, a minimum of three independent clones from unamplified libraries were examined per strain, all of which contained the same coding region defects. No obvious defects are present outside the V13 coding region.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
BALB/c	V13	GAG	GTG	AAG	CTG	ATG	GAG	TCT	GGA	GGA	GGC	TTG	GTA	CAG	CCT	GGG	GCT	TCT	CTG	AGA	CTC	TCC	TGT	GAA	GCT	TCT	GGA
CLA-2/Cn	V13
WSA	V13

		27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
BALB/c	V13	TTC	ACC	TTC	ACT	GAT	TAC	TAC	ATG	AGC	TGG	G TC	CGC	CAG	CCT	CCA	GGG	AAG	TCA	CCT	GAG	TGG	TTG	GCT	TTG	ATT	AGA
CLA-2/Cn	V13
WSA	V13

		52a	52b	52c	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
BALB/c	V13	AAC	AAA	GCT	AAT	GGC	TAT	ACA	ACA	GAG	TAT	AGT	GCA	TCT	GTT	AAG	GGT	CGG	TTC	ACC	ATC	TCC	AGA	GAT	AAT	TCT	CAA
CLA-2/Cn	V13
WSA	V13

		76	77	78	79	80	81	82	82a	82b	82c	83	84	85	86	87	88	89	90	91	92	93	94	95
BALB/c	V13	AAC	ATC	CTC	TAT	CTT	CAA	ATG	AAC	ACC	CTG	AGA	GCT	GAG	GCC	AGT	GCC	ACT	TAT	TAC	TGT	GCA	AAA	GAT
CLA-2/Cn	V13	-G-
WSA	V13	-G-

FIGURE 2. The CLA-2/Cn and WSA V13 gene segments. The coding regions of the CLA-2/Cn and WSA V13 genes are compared with their allele in BALB/c. Insertion and deletion mutations are boxed. The insertion generates four translational stop codons. These sequence data have been submitted to the EMBL/GenBank Data Libraries.

Two V11 Gene Segments are Present in the CLA-2/Cn and WSA Genomes. Two V11 gene segments were identified in the CLA-2/Cn and WSA strains. One of these comigrates with and is identical in sequence to the V11 gene segment of BALB/c (Fig. 3). A second copy of V11 is also present in the CLA-2/Cn and WSA genomes on a novel 4.7-kb Eco RI fragment not present in the BALB/c strain. This extra V11 gene segment will be referred to as V11-2, to distinguish it from the V11 gene segment that comigrates with that present in BALB/c.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
BALB/c	V11	GAG	GTG	AAG	CTG	GTG	GAG	TCT	GGA	GGA	GGC	TTG	GTA	CAG	CCT	GGG	GGT	TCT	CTG	AGA	CTC	TCC	TGT	GCA	ACT	TCT	GGG
CLA-2/Cn	V11
WSA	V11
1C4D8.8	cDNA
CLA-2/Cn	V11-2
WSA	V11-2

		27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
BALB/c	V11	TTC	ACC	TTC	ACT	GAT	TAC	TAC	ATG	AGC	TGG	GTC	CGC	CAG	CCT	CCA	GGG	AAG	GCA	CTT	GAG	TGG	TTG	GGT	TTT	ATT	AGA
CLA-2/Cn	V11
WSA	V11
1C4D8.8	cDNA
CLA-2/Cn	V11-2
WSA	V11-2

		52a	52b	52c	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
BALB/c	V11	AAC	AAA	GCT	AAT	GGT	TAC	ACA	ACA	GAG	TAC	AGT	GCA	TCT	GTG	AAG	GGT	CGG	TTC	ACC	ATC	TCC	AGA	GAT	AAT	TCC	CAA
CLA-2/Cn	V11
WSA	V11
1C4D8.8	cDNA
CLA-2/Cn	V11-2
WSA	V11-2

		76	77	78	79	80	81	82	82a	82b	82c	83	84	85	86	87	88	89	90	91	92	93	94	95		
BALB/c	V11	AGC	ATC	CTC	TAT	CTT	CAA	ATG	AAC	ACC	CTG	AGA	GCT	GAG	GAC	AGT	GCC	ACT	TAT	TAC	TGT	GCA	AGA	GAT		
CLA-2/Cn	V11	
WSA	V11	
1C4D8.8	cDNA
CLA-2/Cn	V11-2
WSA	V11-2

FIGURE 3. The CLA-2/Cn and WSA V11 and V11-2 Gene Segments. The V11 and V11-2 gene segments of the CLA-2/Cn and WSA strains are compared with their homologues in BALB/c. The cDNA V_H sequence expressed by the CLA-2/Cn-derived hybridoma, 1C4D8.8 (11), is shown to demonstrate that the novel V11-2 gene segments analyzed herein are functional. These sequence data have been submitted to the EMBL/GenBank Data Libraries.

Comparison of the coding and flanking region sequences from CLA-2/Cn V11, CLA-2/Cn V11-2, and BALB/c V11 (Fig. 4) shows that a high degree of homology exists among the three segments. It is clear from the flanking region sequences of the CLA-2/Cn V11-2 gene that it is a second copy of the V11 segment present in CLA-2/Cn and BALB/c.

The CLA-2/Cn and WSA V11-2 gene segments are functional and expressed. All promoter, leader peptide, and heptamer-nonamer rearrangement sequences are present and maintain features believed to be required for V_H expression. Most persuasively, this gene segment encodes the V_H sequence expressed in the heavy chain of the PC-binding, CLA-2/Cn-derived hybridoma, 1C4 (11). The cDNA sequence generated from 1C4 contained differences at six nucleotide positions, relative to the BALB/c V11 gene sequence. Five of these differences are present in the germline sequence of the CLA-2/Cn V11-2 gene segment (Fig. 3). The sixth difference (codon

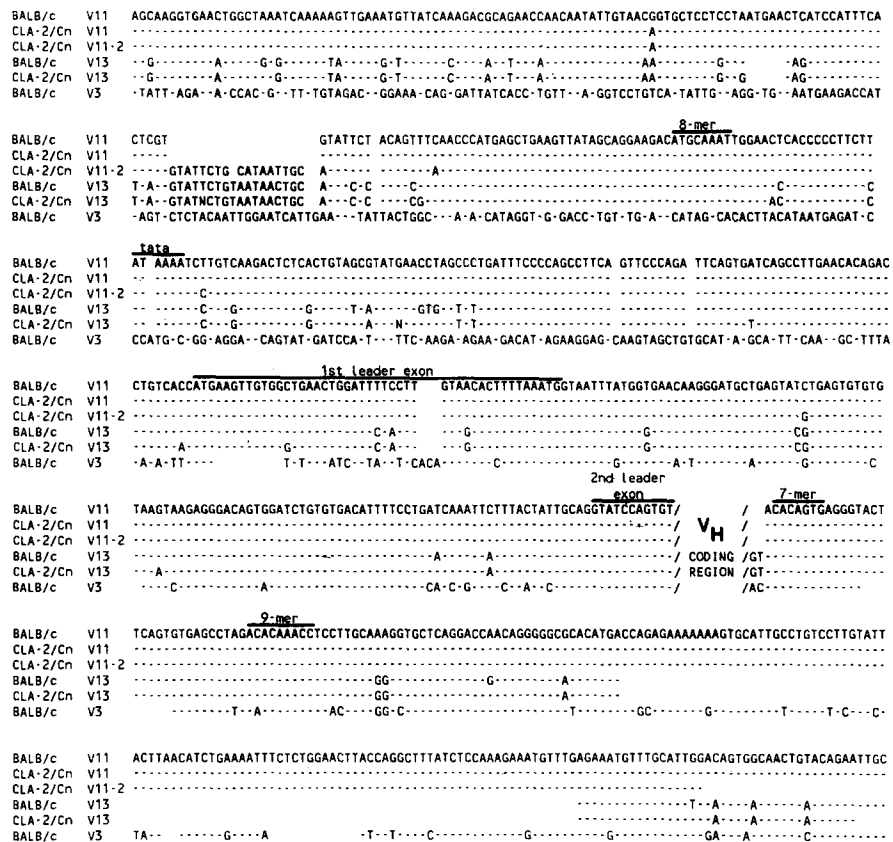


FIGURE 4. Flanking region analysis of V11, V13, and V3 gene segments. CLA-2/Cn V_H sequences were designated alleles of V11 or V13 through comparisons of flanking regions with those determined for the BALB/c strain (17). Two different V11 and a single V13 gene segments are present in the CLA-2/Cn germline. There are no V3-like pseudogene segments present in this strain. Promoter (conserved octamer and TATA sequences) and heptamer-nonamer rearrangement sequences are indicated.

95) is not found in any of the CLA-2/Cn S107 germline sequences and probably represents junctional diversity.

The CLA-2/Cn V1 Gene Segment. In Fig. 5 the CLA-2/Cn V1 gene segment and its associated flanking sequences is compared with its allele in the BALB/c strain. The coding region sequences of the two V1 alleles are identical. Proper promoter elements (the conserved V_H octamer and TATA elements) are present upstream of the CLA-2/Cn V1 gene segment. The heptamer-nonamer rearrangement signal is present as well. While several nucleotide differences exist in the 5'-flanking region sequences of the two strains, none of these would be expected to preclude proper expression of the V1 gene segment of CLA-2/Cn. None of the regions in which these differences occur are conserved in sequence or length among functional V_H gene segments.

CLA-2/Cn and WSA Do Not Possess an Allele of V3. Examination of the V_HS107 gene segments of the BALB/c, C57BL/10, and CBA/J strains revealed the presence of a pseudogene segment, designated V3 (7, 14, 17, 18). This gene segment has been designated a pseudogene, as it contains several coding region defects, a nonfunc-

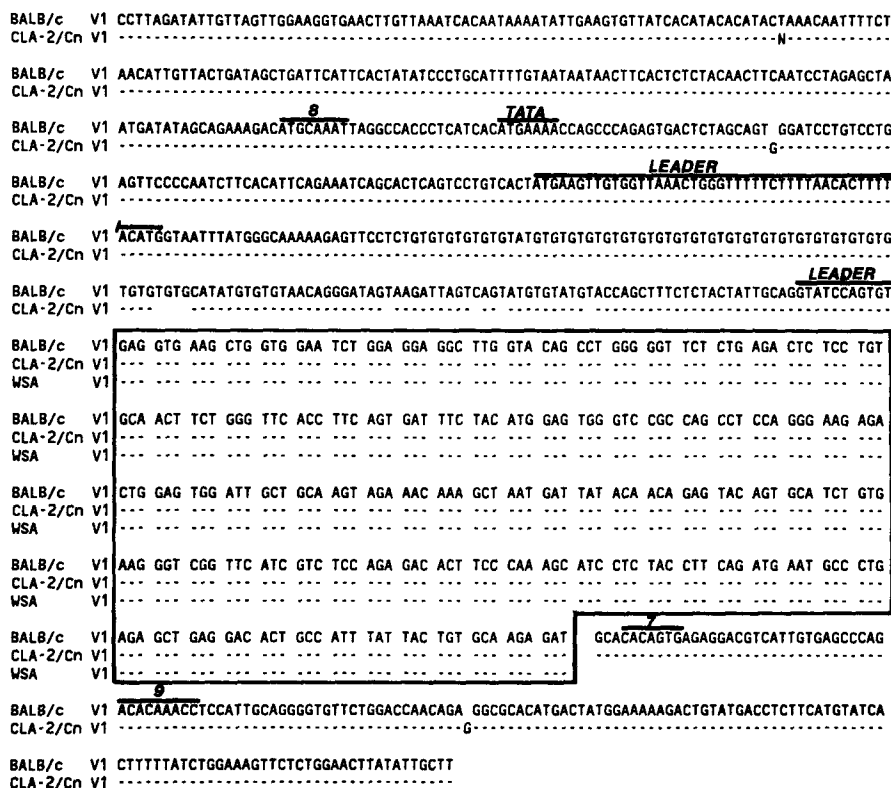


FIGURE 5. The CLA-2/Cn and WSA V1 gene segments. The V1 germline element of the low PC responder strain, CLA-2/Cn is compared with its allele in a responder strain, BALB/c. The coding region of the WSA strain has also been determined. Octamer, TATA, and heptamer-nonamer elements are marked. Exons encoding the leader peptide are indicated. These data have been submitted to the EMBL Gen Bank Data Libraries.

tional heptamer-nonamer rearrangement signal and lacks functional promoter and leader peptide sequences. None of the CLA-2/Cn or WSA gene segments contain such defects. While the V13 gene segment of CLA-2/Cn and WSA is a pseudogene, it is clearly not an allele of the V3 gene segment based on Southern blot as well as flanking region sequence analyses. We believe that we have isolated all V_H S107 gene segments of the CLA-2/Cn and WSA strains, none of which are alleles of V3.

Discussion

The V_H S107 genes of CLA-2/Cn and WSA mice have been examined in detail. The results suggest that: (a) Low PC responsiveness in these populations is unlikely due to an inability to express the V1 member of the V_H S107 gene family; and (b) gene conversion events have occurred with considerable frequency in the evolution of the murine V_H S107 family.

In accord with preliminary analyses that suggested that PC low responsiveness is unlinked to the IgH complex (11; Cancro, M. P., D. Hilbert, and H. Kersten, unpublished results), the CLA-2/Cn V1 segment contains no features that would preclude its use. All coding regions, including leader peptide sequences, are identical to BALB/c, and no defects are found in promoter or rearrangement sequences. Although a 4-bp deletion is found in a $(dG-dT)_n:(dC-dA)_n$ dinucleotide repeat in the intron separating the two leader peptide exons of CLA-2/Cn strain, the length of this repeat is variable among all strains examined (7), including those known to express the V1 gene segment. Similarly, the four single base deletions found in this intron relative to BALB/c also are seen in C57BL/10 and CBA/J sequences (7, 8). Neither D nor J_H elements are likely responsible for the low responder phenotype, since there is no correlation of D region use with PC specificity, and while a preference for J_{H1} is observed, it is not requisite for PC binding (19). Additionally, the $J_{\kappa 5}$ element, which is common to most PC-specific antibodies (19), is intact in CLA-2 and identical to BALB/c (Ferguson, S. E., and M. P. Cancro, unpublished data). Thus, either V_{κ} genes are responsible for low responsiveness, or Ig structural gene defects do not underlie low PC responsiveness.

A large proportion (49%) of differences between alleles of V_H S107 gene family members are consistent with gene conversion. In Fig. 6 all known *Mus musculus domesticus* V_H S107 gene segments are compared. Gene segments are grouped as alleles of individual V_H S107 family members. The V11-2 gene segment is grouped with other V11 genes. Differences attributed to gene conversion are boxed. Several observations suggest that these differences probably represent a minimal estimate of conversion events.

First, a difference between alleles at a particular nucleotide position was considered a conversion only if both alternative bases could be found at an identical position in at least one other member of the V_H S107 gene family. For example, alleles of the V13 gene segment vary at codon position 6 (GAA or GAG). Each of these codons are present in alleles of other V_H S107 family members. Because this conservative criterion of gene conversion requires identification of both donor and recipient sequences, some of the nucleotide differences shown in parentheses and assumed here to represent point mutations might also reflect conversion events. For example, at codon position 5 the V13 allele in the C57BL/10 strain contains GTG, whereas the BALB/c, CBA/J, and CLA-2/Cn alleles all contain the codon ATG. This

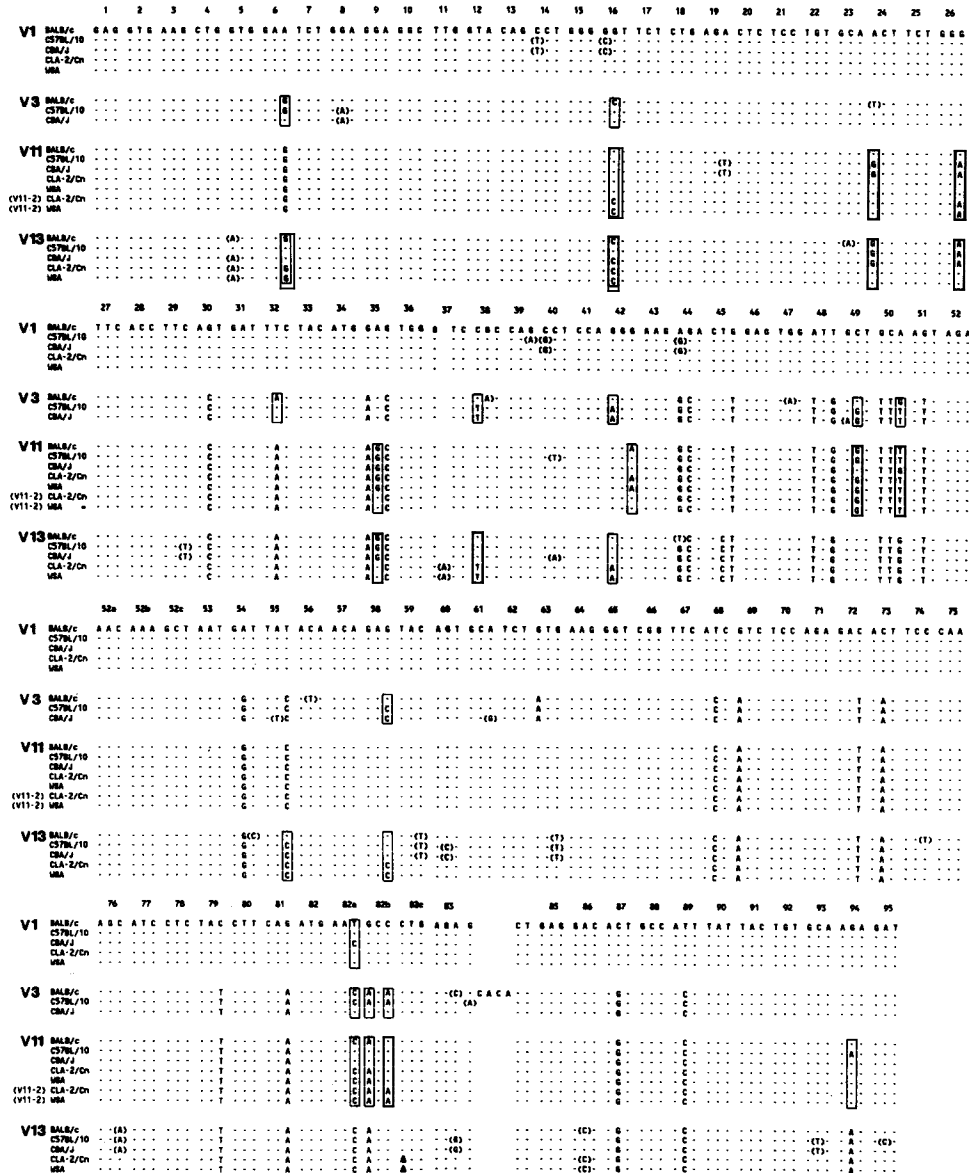


FIGURE 6. Gene conversion among V_H S107 gene segments. The coding region sequences of all V_HS107 gene segments examined in strains of *Mus musculus domesticus* are compared. The BALB/c, C57BL/10, and CBA/J sequences have been published previously (7, 8, 14, 17). Alleles of individual gene family members are grouped. Allelic nucleotide variations attributable to gene conversion (see text for criteria) are boxed. Parentheses indicate differences which do not fulfill the criteria for conversion used herein.

indeed may reflect a point mutation that occurred in the progenitor of the BALB/c, CBA/J, and CLA-2/Cn V13 alleles but not the progenitor of the C57BL/10 allele. Alternatively, all V13 gene segments may have once contained ATG, but the C57BL/10 allele may have served as a recipient for the codon GTG, using any of the other

members of the V_HS107 family as a donor sequence. In fact, the nucleotides that flank this difference (positions 6 and 16) are likely the result of gene conversion, strengthening this possibility.

Second, a striking feature of the V_HS107 family is the paucity of examples where more than two different bases are observed at a particular nucleotide position (this occurs at only 4 of the 64 positions where differences are observed: in codons 24, 40, 44, and 83). If nucleotide substitutions were random within the V_HS107 gene family, the proportion of positions where identical substitutions would exist can be calculated as $(0.33)(1-H)^N$, where H is the degree of homology among the sequences being compared, N is the total number of sequences less the putative donor and recipient, and 0.33 is the probability of choosing one of three possible nucleotides (8, 11, 18). If calculated for the 16 known alleles this value is 21%, and the 49% of observed variations (32 out of 65) consistent with conversion is significantly higher when compared by χ^2 analysis ($\chi^2 = 31.23_{df = 1}$; $p < 0.001$).

The distribution of putative conversion events is skewed to particular family members, as 31 of the 32 putative conversion events involve V11, V3, and V13 alleles as donor/recipient combinations. This is consistent with the observation that sequence homology is positively correlated with the rate of intrachromosomal gene conversion (20), since the degrees of homology between V1 and other segments range from only 85–90%, whereas this figure is 90–100% for all V11, V13, and V3 combinations. If only the V13, V11, and V3 members are considered, the expected random occurrence of matching differences is only 13%, but 53% (31 of 59) of the differences among these alleles are consistent with conversion. This deviation from the expected value based on random occurrence is highly significant ($\chi^2 = 81.08_{df = 1}$; $p < 0.001$). Thus, nonrandom events consistent with conversion have contributed to V_HS107 allelic differences, and are focused among the V11, V3, and V13 alleles.

The paucity of substitutions in V1 attributable to gene conversion might reflect a selective advantage afforded by the specificity of V1-derived antibodies. However, several considerations suggest that this is secondary to homology relationships among individual V_HS107 genes. Many of the observed conversions are silent (e.g., amino acid position 6), would not alter antibody specificity, and should thus serve as innocuous donor sequences for V1. Additionally, many conversions occur in framework regions, and it is clear from point mutations among V1 alleles that some framework changes can be tolerated. In fact, some of these appear less conservative than would be a potential conversion at the same position. For example, the gly → arg substitution between C57BL and BALB/c V1 is less conservative than the gly → ala substitution that would result if this codon were converted by the V3 or V13 members of the BALB/c haplotype.

The CLA-2/Cn and WSA sequences extend understanding of several features of the V_HS107 gene family. First, although recent comparisons of BALB/c V_HS107 members revealed that flanking regions are evolving faster than coding regions (17), this is not true for comparisons made between the alleles of individual gene segments among the CLA-2/Cn and BALB/c strains. In such comparisons (Table I), the V11-2 and V13 gene segments are evolving approximately twice as fast as are the noncoding sequences. Gene conversion may be the basis for the apparent conflict, since it can result in short-term allelic diversity, generating more than one base change through a single event. While we cannot determine the frequency of gene conversion

TABLE I
Drift between BALB/c and CLA-2 V_HS107 Coding and Flanking Regions

Comparison	Region	Number of nucleotides compared	Number of base mismatches	Drift %
CLA-2 V11	Noncoding flanks	571	1	0.2
vs.	Leader exons	57	0	0.0
BALB/c V11	VH segment	303	0	0.0
CLA-2 V11-2	Noncoding flanks	557	5	0.9
vs.	Leader exons	57	0	0.0
BALB/c V11	VH segment	303	5	1.6
CLA-2 V13	Noncoding flanks	569	15	2.6
vs.	Leader exons	57	1	1.6
BALB/c V13	VH segment	303	16	5.3
CLA-2 V1	Noncoding flanks	628	7	1.1
vs.	Leader exons	57	0	0.0
BALB/c V1	VH segment	303	0	0.0

The drift for alleles of V_HS107 family members in BALB/c and CLA was calculated for coding, flanking, and leader exon regions. Drift is calculated as the proportion of positions where base mismatches are observed in the designated region, and is expressed as a percentage.

with these data, it is clear that much gene conversion has occurred even during the relatively short time frame involved in divergence of alleles within a single species.

Second, these data argue against the notion that conversion events seen in the V_HS107 family are inbreeding artifacts. The CLA-2/Cn and WSA gene segments are identical in sequence, and these strains were independently derived from the same feral population in which several V_HS107 haplotypes were segregating (see Fig. 1). Since it is unlikely that identical mutations occurred independently in each strain subsequent to their separation and inbreeding, these gene segments probably represent those present in one haplotype of the original feral population of CNV mice.

Finally, the composition of the CLA/WSA haplotype strengthens the likelihood that a progenitor haplotype similar to CE may have given rise to many of the extant V_HS107 haplotypes. As originally proposed by Perlmutter and Schroeder (18), unequal crossovers within this haplotype could account for the BALB/c and A/J haplotypes through an unequal crossover event. Similarly, the CLA/WSA haplotype could be generated through an unequal crossover in the same progenitor, resulting in two V11 members (Fig. 7).

Based upon the data presented herein, as well as other recent studies suggesting that gene conversion contributes to sequence diversity among mouse and human V_H segments (3, 4, 6–8, 21, 22), it appears that gene conversion globally affects germline V_H genes, and has contributed substantially to *Igh* locus evolution.

Summary

Gene conversion has been suggested as the basis for many V_H allelic differences, particularly in the murine V_HS107 family. Whether conversion among IgV_H genes

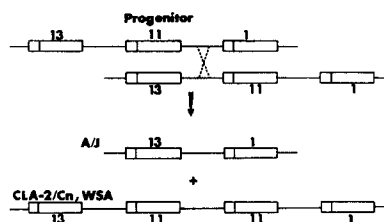


FIGURE 7. Proposed model for the evolution of the CLA-2/WSA $V_H S107$ haplotype. An unequal crossover in progenitor haplotype similar to CE is proposed to have resulted in the CLA/WSA haplotype, which is characterized by two V11-like segments. The reciprocal product would thus bear only V13 and V1 segments, similar to the extant A/J haplotype.

is likely to have occurred in outbred populations has not been directly addressed. The CLA-2/Cn and WSA strains, which were recently and independently derived from a feral population exhibiting low responsiveness to PC, provide the opportunity to approach this question. In previous studies, the heavy chain cDNA sequence of a PC-specific hybridoma derived from CLA-2/Cn suggested gene conversion events within the $V_H S107$ family. Accordingly, we have examined the germline $V_H S107$ genes of CLA-2/Cn and WSA. The results indicate that: (a) The CLA-2 and WSA strains bear an identical but novel $V_H S107$ family haplotype, which lacks a V3 element and contains a V1, a V13, and two V11 genes; (b) low PC responsiveness in these populations is unlikely due to an inability to express the V1 member of the $V_H S107$ gene family; and (c) when compared with the other known $V_H S107$ haplotypes, the proportion of differences consistent with gene conversion greatly exceeds that expected by random base substitution. Thus, gene conversion events appear to have occurred with considerable frequency in the evolution of the murine $V_H S107$ family, especially among the V3, V13, and V11 members.

We are grateful to Dr. Stuart Rudikoff for providing some of the oligonucleotide primers used in these studies, and Drs. Michael Potter and Lawrence D'Hoostelaere for providing the founder populations for the WSA and CLA-2/Cn strains, as well as the DNA samples from early breeding stocks. The helpful discussions and support of Dr. Roger Perlmutter are also gratefully acknowledged.

Received for publication 19 July 1989.

References

1. Bentley, D. L., and T. Rabbitts. 1981. Evolution of Immunoglobulin V genes. Evidence that recently duplicated human V_k sequences have diverged by gene conversion. *Cell*. 32:181.
2. Akimenko, M.-A., B. Mariame, and F. Rougeon. 1986. Evolution of the immunoglobulin kappa locus in the rabbit: Evidence for differential gene conversion events. *Proc. Natl. Acad. Sci. USA*. 83:5180.
3. Krawinkel, U., G. Zoebelen, M. Bruggemann, and A. Radbruch. 1983. Recombination between antibody heavy chain variable region genes: Evidence for gene conversion. *Proc. Natl. Acad. Sci. USA*. 80:4997.
4. Rathbun, G. A., F. Otani, E. C. B. Milner, J. D. Capra, and P. W. Tucker. 1988. Molecular characterization of the A/J J558 family of heavy chain variable region gene segments. *J. Mol. Biol.* 202:385.
5. Reynaud, C.-A., V. Anquez, H. Grimal, and J.-C. Weill. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell*. 48:379.

6. Clarke, S. H., J. L. Claffin, and S. Rudikoff. 1982. Polymorphisms in immunoglobulin heavy chains suggesting gene conversion. *Proc. Natl. Acad. Sci. USA.* 79:3280.
7. Ferguson, S. E., S. Rudikoff, and B. A. Osborne. 1988. Interaction and sequence diversity among T15 VH genes in CBA/J mice. *J. Exp. Med.* 168:1339.
8. Perlmutter, R. M., B. Berson, J. A. Griffith, and L. Hood. 1985. Diversity in the germ-line antibody repertoire: Molecular evolution of the T15 gene family. *J. Exp. Med.* 162:1998.
9. Leiberman, R., L. A. D'Hoostelaere, W. Humphrey, S. Nishinarita, and M. Potter. 1983. An allotypelinked gene is associated with a negative or very low anti-phosphorylcholine response (PC) in wild mice (CNV). *J. Immunol.* 131:1570.
10. D'Hoostelaere, L., and M. Potter. 1986. Ig kappa polymorphism in *M. musculus domesticus* populations from Maryland and Delaware. In *The Wild Mouse in Immunology*. M. Potter, M. Cancro, and J. Nadeau, editors. *Curr. Top. Microbiol. Immunol.* 127:175.
11. Hilbert, D. M., and M. P. Cancro. 1988. Characterization of an inbred mouse exhibiting low responsiveness to phosphorylcholine: antibodies from low-responder mice suggest gene conversion events in the S107 VH family. *J. Immunol.* 140:4364.
12. Liou, H.-C., B. S. Polla, D. Aragnol, L. D. Leserman, I. J. Griffith, and L. H. Glimcher. 1988. A tissue-specific DNaseI-hypersensitive site in a class II Aa gene is under trans-regulatory control. *Proc. Natl. Acad. Sci. USA.* 85:2738.
13. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
14. Crews, S., J. Griffin, H. Huang, K. Calame, and L. Hood. 1981. A single VH segment encodes the immune response to phosphorylcholine: Somatic mutation is correlated with class of antibody. *Cell.* 25:59.
15. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction fragments to high specific activity. *Anal. Biochem.* 132:6.
16. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20.
17. Siu, G., E. A. Springer, H. V. Huang, L. E. Hood, and S. T. Crews. 1987. Structure of the T15 VH gene subfamily: Identification of gene promoter homologies. *J. Immunol.* 138:4466.
18. Perlmutter, R., and H. W. Schroeder, Jr. 1987. Molecular evolution of the T15 V_H family. In *Evolution and Vertebrate Immunity: The Antigen-Receptor and MHC Gene Families*. G. Kelsoe, and D. H. Schulze, editors. University of Texas Press, Austin. 149-162.
19. Rudikoff, S. 1983. Immunoglobulin structure function correlates: Antigen-binding and idiotypes. *Contemp. Top. Mol. Immunol.* 9:169.
20. Liskay, R. M., A. Letsou, and J. L. Statchelek. 1987. Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics.* 115:161.
21. VanDijk, K. W., H. J. Schroeder, R. M. Perlmutter, and E. C. B. Milner. 1989. Heterogeneity in the human Ig VH locus. *J. Immunol.* 142:2547.
22. Schulze, D. H. 1987. A copy-substitution mechanism characterized in MHC mutants appears to be involved in the evolution of a VH multigene family. In *Evolution and Vertebrate Immunity: The Antigen-Receptor and MHC Gene Families*. G. Kelsoe and D. H. Schulze, editors. University of Texas Press, Austin. 413-427.