A Recurrent Breakpoint in the Most Common Deletion of the Ig Heavy Chain Locus (del A1-GP-G2-G4-E)

Alfredo Brusco, Silvia Saviozzi, Fernanda Cinque, Andrea Bottaro and Mario DeMarchi

*J Immunol* 1999; 163:4392-4398; http://www.jimmunol.org/content/163/8/4392

**References** This article *cites 49 articles*, 9 of which you can access for free at: http://www.jimmunol.org/content/163/8/4392.full#ref-list-1

**Subscription** Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

**Permissions** Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts
A Recurrent Breakpoint in the Most Common Deletion of the Ig Heavy Chain Locus (del A1-GP-G2-G4-E)1,2

Alfredo Brusco,3* Silvia Saviozzi,*, Fernanda Cinque,*, Andrea Bottaro,† and Mario DeMarchi‡

Human Ig heavy chain constant regions are encoded by a cluster of genes, the \( \text{IGHC} \) locus, on 14q32.3. Several forms of \( \text{IGHC} \) deletions and duplications spanning one to five genes have been described in different populations, with frequencies of 1.5–3.5% and 4.5–44%, respectively. Despite the common occurrence of these gene rearrangements, little is known about the breakpoint sites; evidence obtained from deletions in the \( \text{IGHC} \) locus and in other regions of the human genome suggests that they preferentially occur in highly homologous regions and might be favored by a variety of recombinogenic signals. We present here a detailed study of three homozygotes for the most common type of \( \text{IGHC} \) multiple gene deletion, spanning the A1-GP-G2-G4-E genes. Using a combination of Southern blotting, long-range PCR, and automated sequencing, the unequal crossover events of all of the six studied haplotypes have been mapped to a region of \( \sim 2 \) kb with almost complete homology between E1-A1 and E-A2, flanked by two minisatellites. These results are consistent with the hypothesis that segments of complete homology may be required for efficient homologous recombination in humans. The possible role of minisatellites as recombination signals is inferred, in agreement with current knowledge. The Journal of Immunology, 1999, 163: 4392–4398.

Received for publication June 7, 1999. Accepted for publication August 3, 1999.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 This work was supported by Ministero dell’Università e della Ricerca Scientifica e Tecnologica (60%).
3 Address correspondence and reprint requests to Dr. Alfredo Brusco, Dipartimento di Genetica, Biologia e Biochimica, via Santena 19, 10126 Torino, Italy. E-mail address: a.brusco@cios.to.cnr.it
4 Abbreviations used in this paper: \( \text{IGHC} \), Ig heavy chain constant region locus; RFLP, restriction fragment length polymorphism; VNTR, variable number of tandem repeats.

Haplotypes with deletions or duplications are present with a variable frequency in all populations analyzed thus far, with a significantly higher frequency of duplications over deletions (Refs. 20 and 21 and the references therein). According to the unequal crossover model, duplications and deletions should originate at the same rate. Therefore, the higher frequency of duplicated vs deleted haplotypes suggests that selective pressures may be operating on the loci to eliminate deletions and/or favor duplications.

Our understanding of the origins of this remarkable genetic instability and our interpretation of the role that selection may play in determining the distribution of the variant haplotypes in the population depend upon establishing the molecular mechanisms that lead to the individual recombination events. So far, the recombination sites of only two deletions have been mapped. The first study, based on a Southern blot approach followed by library construction and sequencing, succeeded in locating the breakpoint of a G1-G4 gene deletion to a highly homologous region located 3’ of G3 and G4, outside of the switch sequences (23). More limited analyses of an A1-E deletion mapped the breakpoint to a large region between the E genes and the switch A regions (24, 25).

In this report, we use Southern blotting, long-range PCR, and sequencing to define the deletion breakpoint in three individuals homozygous for a deletion from A1 to E, which is the most common type of deletion in the Italian population (20).

Materials and Methods

Subjects

 Peripheral blood samples were obtained from three unrelated subjects (BON, DEM, and MON) and 10 controls from different peninsular Italian regions. These three subjects are homozygous for a deletion of \( \sim 150 \) kb encompassing the A1, GP, G2, G4, and E genes. Analysis of \( \text{IGHC} \) RFLP markers and Gm allotypes suggests a different origin of the deleted haplotypes in MON and DEM (10, 13); BON is homozygous for all tested markers (26). Informed consent was obtained from all subjects.

Southern blot analysis

 A total of 5 \( \mu \)g of genomic DNA, extracted using standard phenol/chloroform methods (27), was digested with SacI, PstI, PvuII, and SacI+PstI restriction endonucleases (Roche Diagnostics, Monza, Italy) and run in 0.8% TBE (0.045 M Tris-borate, 0.002 M EDTA)-agarose. DNA was transferred onto Hybond N+ nylon filters (Amersham Pharmacia, Little
Chalfont, U.K.) and hybridized to the following genomic probes using standard techniques:

- pHSα1-0.9 is a 0.9-kb PstI- PvuII fragment (see Fig. 1A) subcloned from pH-Igα11 (kindly provided by T. Honjo, Kyoto University, Kyoto, Japan).
- pHSα2-2.6 is a 2.6-kb SacI/ PstI fragment (see Fig. 1A) subcloned from the l-TOUe-phage (kindly provided by M. P. Lefranc, University of Montpellier II, Montpellier, France).

These probes cross-hybridize with two regions between EP1-A1 and E-A2 genes.

**Cloning and sequencing**

The subcloning of l-TOUe-α and pH-Igα11 was performed using standard procedures (27). The clones obtained were sequenced using a primer walking strategy and dye-terminator chemistry on a model 377 ABI Prism automatic sequence analyzer (Applied Biosystems, Perkin-Elmer, Monza, Italy). Data were elaborated using the Navigator program (Perkin-Elmer) and compared with GenBank with BLAST.

**PCR**

To amplify the regions downstream of EP1 and E, specific forward primers 886 (5'-acactgggtcagctgacactacactgggttc-3') and 889 (5'-ggcctcacacagccctccggtgtaccacag-3') were used with the same reverse primer (891, 5'-tgtgtccacagatgtggctggatgctgagg-3') located 9 kb downstream of each gene. This PCR product was used in a heminested reaction to obtain regions A, B, and C distinctive of the EP1-A1 or the E-A2 regions.

**Region C** was screened for the following sequences involved in hot spots and chromosomal rearrangements using Lalign 2.0 DNA analysis software (52): the E. coli consensus sequence GCTGGTGG, which stimulates generalized recombination (28); the human hypervariable minisatellite consensus sequences GGGCAGGARG (29) and AGAGGTGGGCAGGTGG (30); the consensus sequences for Translin (ATGCAG).

**Sequence comparison**

Region C was screened for the following sequences involved in hot spots and chromosomal rearrangements using Lalign 2.0 DNA analysis software (52); the consensus sequence GCTGGTGG of Escherichia coli, which stimulates generalized recombination (28); the human hypervariable minisatellite consensus sequences GGGCAGGARG (29) and AGAGGTGGGCAGGTGG (30); the consensus sequences for Translin (ATGCAG).
and GCCWCWCT), which are found at chromosomal translocation breakpoints in lymphoid malignancies (31); the ATGAGT sequence present in the recombinogenic M26 allele of Schizosaccharomyces pombe ad6 (32); the tetranucleotide repeats TCTG<sub>4</sub> and CAGG<sub>9</sub> and the 289-bp consensus sequence identified in two recombinational hot spots of mouse MHC, between the A<sub>b</sub>, β and A<sub>b</sub>, γ genes, and in the E<sub>β</sub> gene, respectively (33–36); and the retrotransposon long terminal repeat sequence TCATA-CACCAGCAAGGTTAGAGGACT, located at the recombinational hot spot of the E<sub>β</sub> gene in mouse MHC (37).

**Results**

**Mapping the crossover breakpoint by Southern blotting**

The structure of the two homologous EP1-A1 and E-A2 regions as reported in the literature (2) and integrated by present data is shown in Fig. 1A.

Southern blot analyses of BON, DEM and MON had previously demonstrated a BamHI-RFLP containing EP1 at the 5′ end of the deleted region (i.e., absence of the BamHI site circled in Fig. 1) and a PstI-RFLP containing A2 at its 3′ end (Refs. 10, 13, and 26; and data not shown); therefore, we assumed that the breakpoint was located in the intergenic region between EP1 and A2 in all six deleted chromosomes.

In normal chromosomes, three restriction sites distinguish the EP1-A1 from the E-A2 region (Fig. 1A): a PvuII site upstream of the A1 switch absent in the corresponding position upstream of A2, a SacI site present only in the middle of the switch A2 sequence, and the above-cited BamHI site present within the E gene but not within EP1.

Thus, genomic DNA of the three probands was further analyzed with probes pHS<sub>α2</sub>-2.6 and pHS<sub>α1</sub>-0.9 (see Fig. 2) that, after digestion with the aforementioned enzymes, hybridize to fragments specific for the EP1-A1 or the E-A2 regions (Fig. 3) (38).

On SacI-digested DNA, pHS<sub>α2</sub>-2.6 and pHS<sub>α1</sub>-0.9 showed a single 4.8-kb band in BON and DEM and a heterozygous 4.8/5.4-kb pattern in MON, with the latter band arising from a polymorphism (Fig. 2A); the normal EP1-A1 fragment of 7.1 kb was lacking in all three individuals. Similarly, both probes showed on PvuII filters a 7.1-kb band in BON and DEM and 7.1/7.8-kb bands in MON; again, no EP1-A1-specific fragment (5.4 kb) was detected. Lack of A1 and retention of A2 was also supported by the hybridization of probe pHS<sub>α1</sub>-0.9 to double SacI/PstI digest (Fig. 2C): BON, DEM, and MON showed bands ~2.0 kb, confirming the presence of the switch A2 SacI site. Furthermore, the two probes showed the specific 2.9- to 3.1-kb fragments on PstI filters, indicating that the region upstream of A2 is not deleted (Fig. 2C). On the basis of these results, the breakpoints map upstream of the indicated SacI and PvuII sites in all probands (six haplotypes) and can therefore be assigned to an interval of ~10 kb between these sites and the aforementioned BamHI site (see Fig. 1A).

**Characterization of the EP1-A1 and E-A2 intergenic regions**

The genomic sequences available in GenBank only partially cover the intervals between EP1-A1 and E-A2. Unsequenced regions of ~2.6 kb and ~6.5 kb lay between E-A2 and between EP1-A1, respectively (see Fig. 1A). We tried to fill these gaps by sequencing the entire 2.6-kb SacI-PstI subclone from λ-TOU ε-α phage for E-A2 and the corresponding region from plasmid pH-lg11 for EP1-A1, as well as the S′ adjacent SacI-SacI fragment of 2.3 kb.

The two sequences were almost completely identical (identity equaling 98%). Their analysis in deleted haplotypes was exploited to refine breakpoint mapping by restriction analysis and sequencing (see below).

Variable numbers of tandem repeats (VNTRs) were also detected in the corresponding positions of the two regions. A 33-mer VNTR sequence (consensus AAGGATTTCG AGGTGARGGM VCCCT

GGATT GTC) spans approximately one-third of the SacI-PstI fragment (see Fig. 1A) in both the EP1-A1 and the E-A2 regions; because of its large extension (>0.9 kb), we could not overlap sequences from the two strands. PCR with primers flanking the minisatellite allowed us to establish a wide length variability on genomic DNA, with smaller alleles of ~25–30 repeats and larger alleles of ~40–45 repeats. The consensus sequence did not reveal any homology with known sequences or hypervariable minisatellites. The VNTR polymorphisms can show identical patterns in both the EP1-A1 and E-A2 regions; this makes them less suitable for defining deletion breakpoints (see also Fig. 2B).
In addition, within the 0.5-kb SacI-PstI fragment of the SacI-SacI subclone from both the EP1-A1 and E-A2 regions, sequencing revealed a VNTR with repeats of 40–41 bp (consensus TCCCT TCACC ACCCCCTCCT GCAYCTATC YDCAGCCCCC A), each possessing a canonical minisatellite core (CNTCCTGCCC) and a sequence (CCCACCCC) perfectly matching Nakamura’s consensus (CCCACNNC) (39). Two repeats also showed a \(x\)-like sequence (CCACAGC) with a single nucleotide missing from the consensus (CCACCAGC) (28).

Restriction analysis of long-range PCR products

Most of the critical interval in which the breakpoints are located was amplified by long-range PCR. Two specific forward primers (886 for EP1 or 889 for E) were designed where the two sequences diverge (i.e., EP1 becomes a truncated pseudogene); a common reverse primer (891) was tailored 9 kb downstream. For each region, two semi-nested reactions (primers indicated in Fig. 1B) produced fragment A (~3.5 kb) and fragment B (~4 kb).

The two amplification products from four control subjects (eight haplotypes) were then tested with several frequent-cutting enzymes to detect patterns specific for the EP1-A1 and the E-A2 regions. Four enzymes (DdeI, AluI, MaeIII, and SmalI) in fragment A and four (AluI, DdeI, NlaIII, and MaeIII) in fragment B were found to differentiate the EP1-A1 region from the E-A2 region (Fig. 4A).

In region A, both chromosomes of probands BON and MON showed only EP1-A1-specific sites (Fig. 4B); proband DEM was shown to be homozygous for EP1-A1 sites with SmalI and MaeIII, but EP1-A1/E-A2 was heterozygous with DdeI and HinII, which might indicate a gene conversion event on one chromosome (DEM2 in Fig. 4B). In fragment B, all sites showed an E-A2-typical pattern in the three probands. Thus, the breakpoints could be restricted to a region comprising the two VNTRs and their flanking regions (region C in Fig. 1), which was further analyzed by direct sequencing of overlapping PCR products.

**Fine mapping of the breakpoint**

The sequencing of region C in four control subjects in \(\lambda\)-TOU e-\(\alpha\) and in pH-Ig\(\alpha\)11 permitted us to draw the map shown in Fig. 5. Only positions differing between EP1-A1 and E-A2 in all of the chromosomes studied are reported.

The differences used to map the breakpoint sites in deleted subjects are summarized in Fig. 5. The first seven base differences locate the breakpoint 3’ of the 41-mer VNTR. Two informative base differences are present 5’ of the second VNTR, whereas there are 10 single base differences and one deletion at the 3’. Thus, the

---

**FIGURE 3.** Observed alleles for A1 and A2 RFLPs. Multiple fragment sizes within the same box indicate allelic variants.

**FIGURE 4.** Analysis of regions A and B amplified by hemi-nested, long-range PCR. A, The EP1-A1 or E-A2 discriminatory restriction sites are indicated. B, The results for BON, DEM, and MON are indicated below. Numbers 1 and 2 refer to individual chromosomes. Black dots indicate EP1-A1 region sites, whereas white dots indicate the E-A2-associated restriction sites. In region B, DdeI and MaeIII (data not shown) display the same differences as AluI.
crossover event could be mapped to the 2-kb region of high homology between the two VNTRs in all subjects. The two nucleotide differences in the middle of the 2-kb region of homology are polymorphic, and were not considered for breakpoint mapping.

Known signals favoring recombination were searched for in region C (see Materials and Methods). One sequence and three binding sites for Translin showed complete identity with their consensus; in addition, a number of sites showed a single nucleotide difference from other putative recombinogenic sequences (Fig. 5 and Table I).

**Discussion**

Unequal crossing-over between mispaired homologous regions is the cause of many genetic disorders and of evolutionary variability in multigene families (reviewed in Ref. 40). One notable example is the \( IGHc \) locus, where deletions/duplications of single or multiple genes reach a polymorphic frequency. Because homology is often as elevated in intergenic regions of \( IGHc \) as in the coding sequences, intergenic exchanges giving rise to hybrid genes are rare among recombinant haplotypes (41). The first recombinational breakpoint within the \( IGHc \), described by Keyeux et al. (23) in a G1-G4 homozygous deleted individual, occurred in a highly homologous region, outside of the switch sequences, which the authors named “hsg”. These regions were subsequently shown to contain repetitive MER-4 elements (42). Hsg homology regions were also located downstream of each G gene except G2, suggesting that even other deletions (e.g., of EP1-GP and EP1-G4) could derive from intrachromosomal recombinations between these regions. By analogy, the A1-E and GP-A2 deletions might arise through recombinations between homologous regions downstream of the E or A genes but different from hsg. In general, it could be argued that extensive stretches of high sequence homology are necessary for the generation of deletions and duplications in the \( IGHc \) locus.

The most common deletion/duplication, involving the genes from A1 to E, depends upon mispairing between two G-G-E-A blocks. Recently, a homozygous A1-E deletion (T17) was independently investigated by Wiebe et al. (24) and Chen et al. (25). Using a Southern blot approach, the deletion breakpoint was located in two different fragments due to the different interpretations

| Table I. Homologies to recombinogenic consensus sequences in region C | | |
|---|---|---|---|
| Sequence Designation | Consensus | Species | Sequence Found<sup>a</sup> |
| 1 Translin consensus | GCCCWCTT | Homo sapiens | GCGGACCT |
| 2 Ade6 sequence | ATGACGT | Seccharomyces pombe | ATGACCT |
| 3 \( \chi \) sequence | GCTGTTGG | Escherichia coli | GCTGGTGG in EP1-A1 |
| 4 Ade6 sequence | ATGACGT | S. pombe | GGACGT |
| 5 Hypervariable minisatellite | GGCCAGGARG | H. sapiens | GGGGAGAAGG |
| 6 Hypervariable minisatellite | GGCCAGGARG | H. sapiens | GGGGAGAAGG |
| 7 Translin consensus | ATGACG | H. sapiens | ATGACG |
| 8 \( \chi \) sequence | GCTGTTGG | E. coli | GCTGATGG |
| 9 Translin consensus | GCCCWCTT | H. sapiens | GCCCAGCA |
| 10 Translin consensus | GCCCWCTT | H. sapiens | GCCCAGCA |
| 11 Translin consensus | ATGACG | H. sapiens | ATGACG |
| 12 Translin consensus | GCCCWCTT | H. sapiens | GCCCAGCA |
| 13 Ade6 sequence | ATGACGT | S. pombe | ATGACGT |

<sup>a</sup> Position numbering refers to Fig. 5.
<sup>b</sup> Sequence positions differing from consensus are boldface and underlined.
of the RFLP pattern. One group had assigned the breakpoint to a 5.2-kb SacI/SacI fragment containing the switch A2. However, this region is highly polymorphic because of both the switch and the 33-mer VNTR; therefore, the fragment was likely a normal polymorphism unrelated to the deletion. Chen et al. (25) restricted the breakpoint region to a 2.0-kb BamHI-SacI fragment just downstream of the EP1 gene. Unfortunately, this conclusion is also incorrect because Southern blotting can only place the T17 breakpoint between the discriminatory sites reported in Fig. 1. It should be noted that the Psrl digestion used by these authors cannot distinguish the EP1-A1 from the E-A2 derived fragments. Using long-range PCR, we have evidence that fragments of the same length can originate from either region, likely due to the 33-mer VNTR (see Fig. 3 and the legend to Fig. 2).

Our study localizes the recombination for the A1-E deletion in three unrelated homozygotes within the intergenic region between EP1/E and A1-A2 genes, which show ~98% sequence identity. More precisely, the breakpoint was pinpointed in all individuals by sequencing to a ~2-kb region of absolute homozygosis (only two nucleotide differences) between two VNTRs. Furthermore one haplotype (DEM2) also shows an intermixed restriction site pattern for EP1-A1 and E-A2, compatible with a gene conversion event in proximity of the exchange site. Gene conversion events have been reported by several authors in association with meiotic crossing-over and other homologous recombination events. These findings are in agreement with current double-strand break models of recombination (43).

Our data suggest that in humans, long stretches of identity favor recombination between two sequences. Indeed, sequence homology has been demonstrated in recombinational hot spots such as those of the mouse MHC locus (44), pseudautosomal region of sex chromosomes (45), hereditary neuropathy with liability to pressure palsies, and Charcot-Marie-Tooth 1A disease (46).

In addition to sequence identity, other specific factors or signals may have targeted the recombination event between the two G-G-E-A blocks. In fact, the same mispairing could theoretically also generate other deleted/duplicated haplotypes involving the GP-A2, EP1-G4, and G1-G4 genes: however, these are significantly rarer. Variable selective pressures are one potential explanation for this difference. However, it seems unlikely for a deletion of the AI-GP-G2-G4-E genes to be significantly favored in terms of fitness over one spanning, for instance, GP-G2-G4-E-A2. Alternatively, specific recombination-prone sequences between EP1-A1 and/or between E-A2 may have favored recombination in the interval. The low linkage disequilibrium level observed in particular within the E-A2 region (47–49) supports this model.

Repetitive DNA sequences have been proposed as preferred sites of recombination from microorganisms to humans. In the IGHC locus, the repetitive “switch” elements upstream of each functional CH gene are good candidates in light of their programmed involvement in somatic switch recombination. In an insertional mutagenesis assay, the Sμ region was shown to represent a recombinogenic hot spot with a 100-fold higher insertion frequency than the average genomic region (50).

In this respect, the two VNTRs found in the surroundings of the recombinational region may be particularly important. The consensus core sequence of human hypervariable minisatellites shows similarity to the E. coli χ sequence (39), a hot spot that promotes generalized recombination catalyzed by RecA-RecBC (51). Direct experimental evidence has also been obtained showing that minisatellite sequences enhance the ability of a DNA molecule to undergo recombination at a distance, perhaps acting as recognition sites for a specific enzyme or macromolecular complex (30).

In our study, the breakpoint critical region in all six investigated haplotypes lay in the same ~2-kb sequence containing two VNTRs and several consensus sequences that are known to be involved in recombination in humans, yeast, and bacteria. This strongly suggests, although it does not directly prove, that these recombinogenic elements may represent the site of the unequal crossing-over, and may even coincide with an alleged recombination hot spot of the IGHC locus.

Acknowledgments

We thank Profs. M. P. Lefranc and T. Honjo for providing clones, A. Plebani and M. G. Guerra for patients, Prof. N. Migone for critical discussion, and C. Boccazzi for technical support. We are also very grateful to Dr. Enza Ferrero for critical revision of the manuscript.

References


The Journal of Immunology 4397

Downloaded from http://www.jimmunol.org/ by guest on September 25, 2017


