

# Duplication of the CD8 $\beta$ -chain gene as a marker of the man–gorilla–chimpanzee clade

(primate/evolution/lymphocyte/antigen)

CHRISTIANE DELARBRE\*, HIROMITSU NAKAUCHI†, RONALD BONTROP‡, PHILIPPE KOURILSKY\*, AND GABRIEL GACHELIN\*

\*Unité de Biologie Moléculaire du Gène, Unité 277, Institut National de la Santé et de la Recherche Médicale, Institut Pasteur, 25 rue du Dr. Roux, 75015 Paris, France; †Laboratory of Cell Growth and Differentiation, The Institute of Physical and Chemical Research, Tsukuba Life Science Center, Tsukuba, Japan; and ‡Institute of Applied Radiobiology and Immunology, Netherlands Central Organization for Applied Scientific Research, Rijswijk, The Netherlands

Communicated by François Jacob, May 4, 1993 (received for review March 4, 1993)

**ABSTRACT** In earlier studies we have found that the gene encoding the CD8  $\beta$  chain is duplicated in man. We demonstrate here that the duplicated genes are both located on chromosome 2. We have also studied the moment of the duplication event relative to the evolution of higher primates by using genomic DNA of a panel of primates. Our data strongly suggest that duplication occurred after the orangutan lineage had split and before the chimpanzee, gorilla, and man clade diverged, some 8–9.5 million years ago. This result makes the CD8  $\beta$ -chain gene duplication a convenient marker for the study of the evolution of higher primates.

The T-cell differentiation antigen CD8 is a glycoprotein present at the surface of thymocytes and T lymphocytes that recognize antigenic peptides in the context of the class I major histocompatibility complex (MHC) molecules (1). In the human, the CD8 antigen is a heterodimer product of the association of an  $\alpha$  and a  $\beta$  chain, both encoded by genes closely linked on chromosome 2 (2, 3) in the vicinity of the immunoglobulin  $\kappa$ -chain locus *IGK* (4). The  $\alpha$  chain binds to the  $\alpha 3$  domain of the MHC class I molecules (5), and, in the mouse, the  $\beta$  chain broadens the recognition specificity of CD8 for MHC class I molecules (6). In the course of studies on the expression of the human CD8  $\beta$ -chain gene, designated *CD8B*, and on the origin of the distinct forms of CD8  $\beta$ -chain cDNA, which presumably result from alternative splicing, two of us (H.N. and G.G.) discovered a second but inactive CD8  $\beta$ -chain gene, which we have named *CD8B2*, the functional gene being renamed *CD8B1* (7). The nucleotide sequences of the coding and noncoding regions of the *CD8B1* and *CD8B2* genes were found to be 98.9% homologous. The duplication event has thus occurred rather recently in the evolution of the human lineage. Thus, the CD8  $\beta$ -chain gene duplication, segregating a gene and a pseudogene, could become a useful marker for studies on the phylogeny of higher primates including man and for studies on the mechanisms of gene duplication.

We had previously shown that the *CD8B* duplicated genes existed in all of the 60 human (of Japanese and Caucasian origin) genomic DNA samples tested (7) and therefore that duplication had most likely occurred before the split of the human lineage. We have studied the moment of the duplication by searching for the presence of a single or of duplicated CD8  $\beta$ -chain gene(s) in the genomic DNA of a panel of primates. We show in the present paper that the duplicated genes are both located on chromosome 2 and that duplication had taken place in an ancestor common to the man, gorilla, and chimpanzee clade, most likely excluding the orangutan and other primates.

## MATERIALS AND METHODS

**Sources of Genomic DNA.** Human–hamster or human–mouse somatic cell hybrids. Genomic DNAs prepared from the cells of the National Institute of General Medical Sciences mapping panel no. 1 (Coriell Institute for Medical Research, Camden, NJ) were gifts of J. Weissenbach (Institut Pasteur).

**Primate DNAs.** Human genomic DNA isolated from cell line GPP15 was a gift of M. Fellous (Institut Pasteur). All other primate genomic DNAs were prepared at the Institute of Applied Radiobiology and Immunology, Netherlands Central Organization for Applied Scientific Research.

**Polymerase Chain Reaction (PCR).** The primers used are described in Fig. 1. For the screening of human–mouse and human–hamster hybrids, genomic DNA was hybridized with the common 5' primer and the *CD8B1*-specific or the *CD8B2*-specific 3' primers. PCR was carried out as follows: 0.2  $\mu$ g of genomic DNA, 1.5 mM  $MgCl_2$ , 0.5  $\mu$ M primers, 0.2 mM dNTP, and 1 unit of *Taq* polymerase (Promega) were mixed in a total volume of 50  $\mu$ l; after 2 min of denaturation at 94°C, 40 cycles (70 sec at 94°C, 1 min at 60°C, and 4 min at 72°C) were carried out, followed by a last cycle at 72°C for 10 min. The same conditions were used for the detection of *CD8B1* exon 9 by PCR with primers specific for exon 9 that extended over two nucleotides on each side within the noncoding flanking sequences.

**Sequencing of Amplified DNA.** Amplified DNA was separated by electrophoresis on a 2% agarose and electroeluted. Eluted fragments were made blunt-ended with T4 DNA polymerase (Pharmacia), phosphorylated with polynucleotide kinase (Pharmacia), ligated by using T4 DNA ligase (Pharmacia) to pBluescript KS open at the *EcoRV* site, and treated with calf intestine alkaline phosphatase (Boehringer Mannheim) according to conventional procedures (8). JM 101 was transformed with the ligation products. Sequencing was carried out on single-stranded DNA obtained with helper phage R408. Universal M13 (–40) primer and SK primer were used with the Sequenase sequencing kit (United States Biochemical).

## RESULTS

**Chromosome Assignment of the *CD8B2* Gene.** Since the chromosome assignment of the human *CD8B2* gene had not yet been defined, we first determined its chromosomal location. The two genes are 98.9% homologous over the entire sequence of the first seven exons. The homology between the two genes stops abruptly after exon 7. Exons 8 and 9 are missing in the *CD8B2* gene. From the sequences surrounding the breakpoint, we have designed a common 5' primer and two 3' primers specific for *CD8B1* and *CD8B2* sequences, respectively (Fig. 1). Their use on human genomic DNA

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



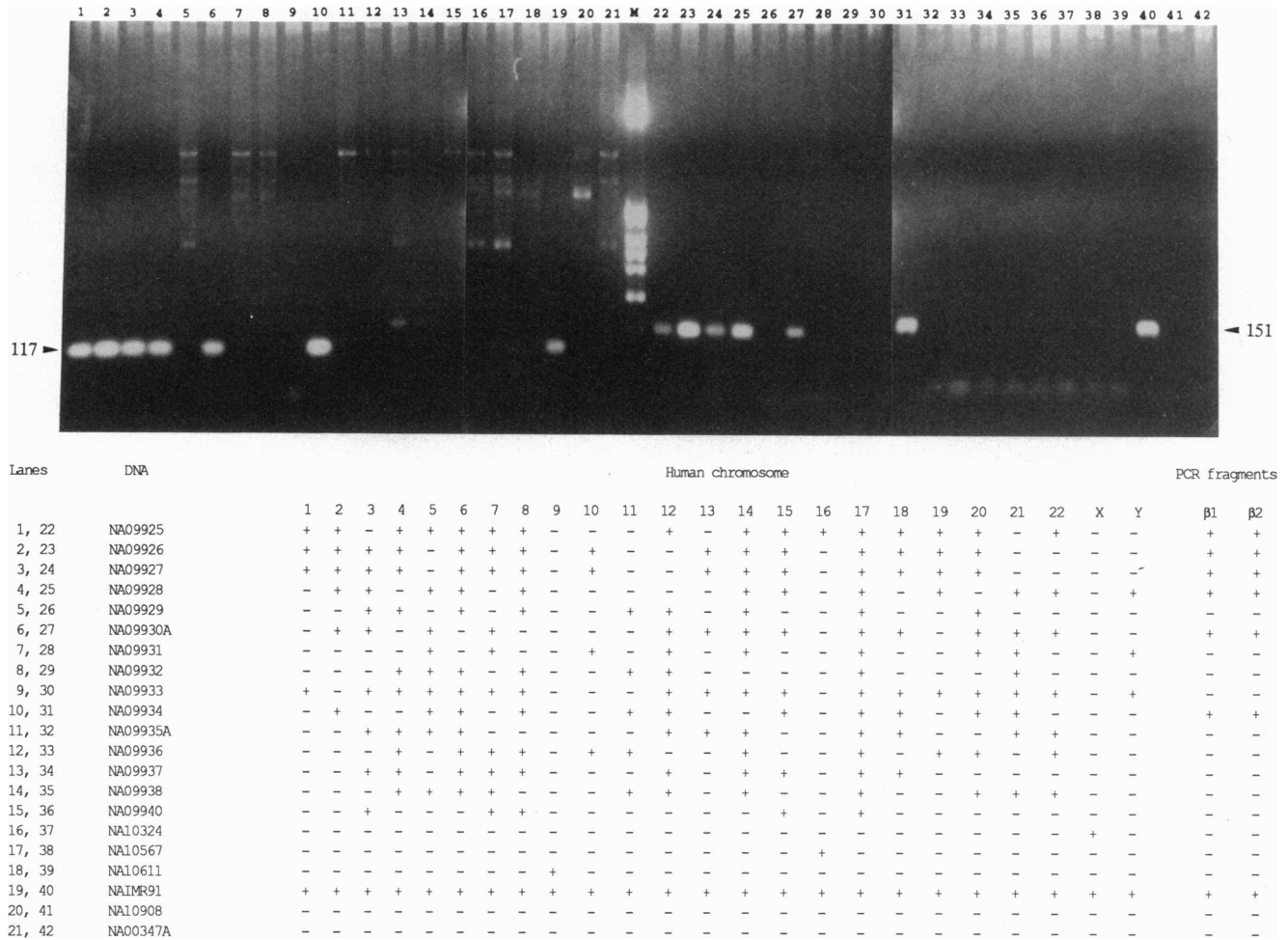


FIG. 2. Chromosomal location of *CD8B2* gene. (Upper) We used the National Institute of General Medical Sciences human-rodent somatic cell hybrids to define the chromosomal location of the *CD8B2* gene. Three hundred nanograms of each genomic DNA were amplified by PCR as described in text. Of each sample, 1/10th was analyzed by electrophoresis in a 2% agarose gel. The DNAs were used in the following order in lanes 1–21 and again in lanes 22–42: NA09925, NA09926, NA09927, NA09928, NA09929, NA09930A, NA09931, NA09932, NA09933, NA09934, NA09935A, NA09936, NA09937, NA09938, NA09940, NA10324, NA10567, NA10611 and parental DNAs NAIMR91 (human line IMR91), NA10908 (Chinese hamster ovary line UV-135), and NA00347A (mouse line B-82). In the series of lanes 1–21, the DNAs were amplified by using the *CD8B1*-specific 3' primer. In the series of lanes 22–42, the DNAs were amplified by using the *CD8B2*-specific 3' primer. pBR322 digested by *HinfI* supplied size markers (lane M). (Lower) The table indicates the presence (+) or the absence (–) of the indicated human chromosome in each of the human-rodent somatic cell hybrids (named DNA). Hybrids possessing more than 10% (of the indicated chromosome) are considered positive for the presence of the considered human chromosome. The two right lanes summarize the presence (+) and absence (–) of PCR amplification products based on the use of *CD8B1* ( $\beta 1$ )- and *CD8B2* ( $\beta 2$ )-specific primers.

amplification products with a *CD8B2* sequence, we carried out a primer extension on amplification products using the <sup>32</sup>P-labeled internal 5' extension primer (Fig. 1). The resulting radiolabeled extension products were analyzed on a sequence gel. A radiolabeled band migrating with the expected size of 128 bp was observed in the human control, in the two gorilla DNAs, and the three chimpanzee DNAs. No *CD8B2*-specific fragment could be identified among amplification products from genomic DNA of other primates, including the orangutan (Fig. 4). Finally, we subcloned PCR products of gorilla Rambo, chimpanzee Ingrid, and of the human control and determined the nucleotide sequence of several clones. We found that chimpanzee, gorilla, and human DNA sequences were identical to the already determined sequence of the human *CD8B2* gene (7).

DISCUSSION

A region of the human chromosome 2, including a large part of the gene encoding the CD8  $\beta$  chain, has been duplicated and transposed elsewhere on the same chromosome. In an

earlier publication (7), we had determined the relative locations of the CD8  $\alpha$ -chain gene (designated *CD8A*) and the *CD8B1* gene: the two genes are in the same orientation and *CD8A* is located 25 kb downstream from the *CD8B1* gene. Earlier the *CD8A* gene had been mapped within the 2p12 band, which also includes the *IGK* locus (11). *CD8A* and *IGK* are thus the only genes or loci so far known to exist in the vicinity of the *CD8B* chain gene. No evidence for a duplication of the *CD8A* gene and of the *IGK* locus has ever been found in the human (11). Thus, the transposed segment excludes the *CD8A* and *IGK* loci. The precise extent of the transposed area has not been determined yet: whereas its 3' end was found to be located 25 kb upstream of the *CD8A* gene (7), the 5' end of the transposed stretch of DNA could not be determined from available data.

Within the limits of the small number of independent primate DNA samples studied, the duplication of the *CD8B* gene appears to have taken place in a lineage common to the gorilla, chimpanzee, and man after the orangutan clade had split. We have not mapped yet the duplicated *CD8B* genes in the gorilla and in the chimpanzee. Such a study should be

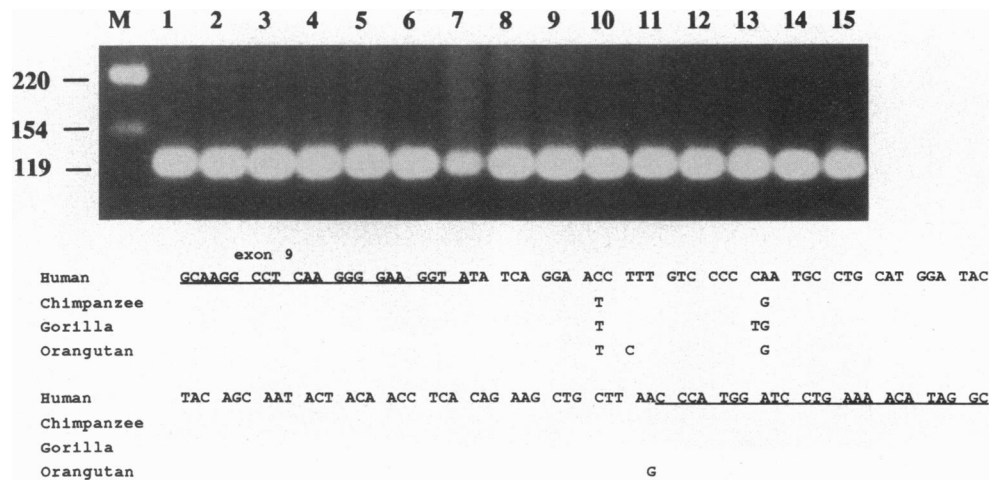


FIG. 3. (Upper) Presence of *CD8B1* exon 9 in all primates. Three hundred nanograms of genomic DNA was amplified by using *CD8B1* exon 9-specific primers depicted in Fig. 1. One tenth of each sample was analyzed by electrophoresis in a 2% agarose gel. The order of the samples was as follows: chimpanzees Ingrid (lane 1), Agnette (lane 2), and Louise (lane 3); gorillas Rambo (lane 4) and Dihi (lane 5); orangutans Elmar (lane 6) and Jago (lane 7); *Hamadryas* baboon cell line S594 (lane 8); rhesus monkeys 1WM (lane 9), 1RK (lane 10), and 2Y (lane 11); Stumptail macaques 125 (lane 12), 258 (lane 13), and 288 (lane 14); and human line GPP15 (a gift of M. Fellous) (lane 15). pBR322 digested with *HinfI* supplied size markers (lane M). (Lower) The amplified fragments of chimpanzee Ingrid, gorilla Rambo, and orangutan Elmar were sequenced after subcloning in pBluescript KS vector. The sequences are aligned on the human *CD8B1* exon 9 sequence (7).

replaced in the context of the origin of human chromosome 2. Indeed, the present human chromosome 2 is derived from the fusion of two chromosomes in other nonhuman primates—chromosomes 12 and 13 in the chimpanzee and chromosomes 11 and 12 in the gorilla (11). Further cytogenetical analyses are required to determine if the duplicated genes are on the same chromosome in the chimpanzee and in the gorilla.

The duplication of *CD8*  $\beta$ -chain genes appears to be a molecular event that unambiguously favors an earlier split of the orangutan clade, assuming parsimony, the duplication being one event present in the ancestor common to African apes and man and inherited. Our data dismiss cladograms that isolate man from apes (12) or relate man to *Pongo* (13), but our data are in good agreement with most molecular (14, 15), morphological (16), and paleontological (17) data indicating that the man-gorilla-chimpanzee lineage excludes the more ancient orangutan clade.

Thirty-three nucleotide differences have been scored over 2940 nucleotides (including 647 coding nucleotides) of the sequence shared by *CD8B1* and *CD8B2* genes. It has been noted for a long time that the rate of nucleotide substitutions is lower in primates, and especially in human, than in other vertebrates (18). By using the values of  $1.56 \times 10^{-9}$  substitutions per site per year, deduced from sequences of the IgE chain in chimpanzee and orangutan (19), and of  $1.3 \times 10^{-9}$  substitutions per site per year, deduced from sequences of the  $\psi\eta$ -globin gene in most primates (20), we calculated that duplication of *CD8*  $\beta$ -chain genes had occurred 8–9.5 million years ago. These values are close to the estimated dating of the branching point of the homininae (*Pan-Gorilla-Homo*)—namely, 5.3–8 million years (21) or 6.3–8.1 million years ago (22). Thus, duplication of the *CD8*  $\beta$ -chain gene occurred slightly before the human, gorilla, and chimpanzee families separated from each other and is a distinctive molecular feature of the homininae clade; our molecular data provide

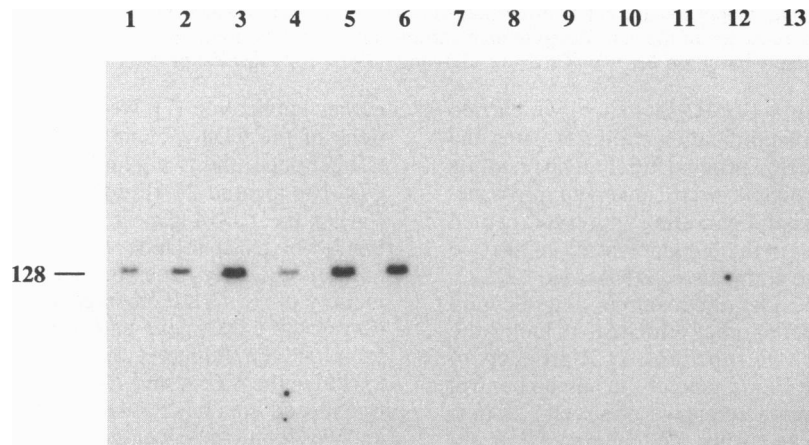


FIG. 4. Presence of *CD8B2* chain genes in upper primates. The region of genomic DNA around the breakpoint was PCR-amplified by using the *CD8* common 5' primer and the *CD8B2*-specific 3' primer. The composition and the final volume of the reaction mixture were as indicated in text, whereas PCR conditions were altered as follows: 94°C at 2 min followed by 25 cycles, each consisting of 70 sec at 94°C, 1 min at 68°C; and 1 min at 72°C followed by a last cycle of 10 min at 72°C. Amplification products (2  $\mu$ l) were annealed with the  $^{32}$ P-labeled internal 5' extension primer. After 2 min at 94°C, annealing at 60°C was followed by a single cycle of extension lasting 15 min at 72°C, and extension products were analyzed on a 5% acrylamide sequence gel (10). The order of the DNA samples was human DNA GPP15 used as a control (lane 1); chimpanzees Ingrid (lane 2), Agnette (lane 3), and Louise (lane 4); gorillas Rambo (lane 5) and Dihi (lane 6); orangutans Elmar (lane 7) and Jago (lane 8); *Hamadryas* baboon S594 (lane 9); rhesus monkeys 1RK (lane 10) and 2Y (lane 11); and macaques 258 (lane 12) and 288 (lane 13).

unequivocal support for the conclusion that human and African great apes form a monophyletic group.

We thank Dr. P. Couillin for the gift of samples of genomic DNA and Drs. P. Darlu and P. Tassy for helpful discussions.

1. Swain, S. L. (1983) *Immunol. Rev.* **74**, 129–137.
2. Spurr, N. K., Goodfellow, P. N. & Johnson, P. (1988) *Immunogenetics* **27**, 70–72.
3. Bruns, G., Kavathas, P., Shiloh, J., Sakai, K., Schwaber, J., Latt, S. A. & Herzenberg, L. A. (1986) *Hum. Genet.* **70**, 311–314.
4. Sukhatme, V. P., Volmer, A. C., Erickson, J., Isobe, M., Croce, C. & Parnes, J. R. (1985) *J. Exp. Med.* **161**, 429–434.
5. Salter, R. D., Benjamin, R. J., Wesley, P. K., Buxton, S. E., Garrett, T. P., Clayberger, C., Krensky, A. M., Norment, A. M., Littman, D. R. & Parham, P. (1990) *Nature (London)* **345**, 41–46.
6. Karaki, S., Tanabe, M., Nakauchi, H. & Taniguchi, M. (1992) *J. Immunol.* **149**, 1613–1618.
7. Nakayama, K. I., Kawachi, Y., Tokito, S., Minami, N., Yamamoto, R., Imai, T., Gachelin, G. & Nakauchi, H. (1992) *J. Immunol.* **148**, 1919–1927.
8. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
9. Lawlor, D. A. & Parham, P. (1992) *Immunogenetics* **36**, 121–125.
10. Gachelin, G., Delarbre, C. & Kourilsky, P. (1993) in *HLA 1991: Proceedings of the 11th Workshop on Histocompatibility*, eds. Tsuji, K., Aisawa, M. & Sasazuki, T. (Oxford Univ. Press, New York), Vol. 2, pp. 56–63.
11. O'Brien, S. J., ed. (1990) *Genetic Maps* (Cold Spring Harbor Lab. Press, Plainview, NY).
12. Kluge, A. G. (1983) in *New Interpretation of Apes and Human Ancestry*, eds. Ciochon, R. L. & Corruccini, R. S. (Plenum, New York), pp. 151–177.
13. Schwartz, J. H. (1984) *Nature (London)* **308**, 501–505.
14. Barriel, V. & Darlu, P. (1990) *Bull. Mem. Soc. Anthropol. Paris* **2**, 3–24.
15. Bailey, W. J., Fitch, D. H. A., Tagle, D. A., Czeluswiak, J., Slightom, J. L. & Goodman, N. (1991) *Mol. Biol. Evol.* **8**, 155–184.
16. Barriel, V. (1991) *Bull. Mem. Soc. Anthropol. Paris* **3**, 45–72.
17. Andrews, P. (1992) *Nature (London)* **360**, 641–646.
18. Li, W. H. & Tanimura, M. (1987) *Nature (London)* **326**, 93–96.
19. Sakoyama, Y., Hong, K.-J., Byun, S. M., Hisajima, H., Ueda, S., Yaoita, Y., Hayashida, H., Miyata, T. & Honjo, T. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1080–1084.
20. Miyamoto, M. M., Koop, B. F., Slightom, J. L., Goodman, M. & Tennant, M. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7627–7631.
21. Koop, B. F., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. (1986) *Nature (London)* **319**, 234–238.
22. Goodman, M. (1986) in *Evolutionary Perspectives and the New Genetics*, eds. Gershowitz, H., Rucknagel, D. L. & Tashian, R. E. (Liss, New York), pp. 121–132.