

Cloning and Chromosomal Mapping of the Human Nonfunctional Gene for L-Gulono- γ -lactone Oxidase, the Enzyme for L-Ascorbic Acid Biosynthesis Missing in Man*

(Received for publication, October 21, 1993)

Morimitsu Nishikimi†§, Ryuichi Fukuyama¶, Sinsei Minoshima¶, Nobuyoshi Shimizu¶, and Kunio Yagi‡

From the †Institute of Applied Biochemistry, Yagi Memorial Park, Mitake, Gifu 505-01, Japan and the ¶Department of Molecular Biology, Keio University School of Medicine, Tokyo 160, Japan

Man is among the exceptional higher animals that are unable to synthesize L-ascorbic acid because of their deficiency in L-gulono- γ -lactone oxidase, the enzyme catalyzing the terminal step in L-ascorbic acid biosynthesis. In the present study, we isolated a segment of the nonfunctional L-gulono- γ -lactone oxidase gene from a human genomic library, and mapped it on chromosome 8p21.1 by spot blot hybridization using flow-sorted human chromosomes and fluorescence *in situ* hybridization. Sequencing analysis indicated that the isolated segment represented a 3'-part of the gene, where the regions corresponding to exons VII, IX, X, and XII of the rat L-gulono- γ -lactone oxidase gene remain with probable deletion of the regions corresponding to exons VIII and XI. In the identified exon regions were found various anomalous nucleotide changes, such as deletion and insertion of nucleotide(s) and nonconformance to the GT/AG rule at intron/exon boundaries. When the conceptual amino acid sequences deduced from the four exon sequences were compared with the corresponding rat sequences, there were a large number of nonconservative substitutions and also two stop codons. These findings indicate that the human nonfunctional L-gulono- γ -lactone oxidase gene has accumulated a large number of mutations without selective pressure since it ceased to function during evolution.

Most phylogenetically higher animals can synthesize L-ascorbic acid from D-glucose. Humans, other primates, and guinea pigs are among the exceptional species that are unable to do so and consequently require a dietary intake of this vitamin to prevent scurvy. This metabolic defect arose during evolution of these animals and is currently carried in all of their individuals. In this sense scurvy is an unusual type of inborn error of metabolism (for a review, see Ref. 1).

Enzymological studies in the late 1950s revealed that the genetic defect was caused by the deficiency of L-gulono- γ -lactone oxidase (GLO)¹ (2), the enzyme that catalyzes the terminal step of L-ascorbic acid biosynthesis in most animals. Our recent

* This work was supported in part by Grant-in-aid 01580156 for Scientific Research from the Ministry of Education, Science, and Culture of Japan. 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) D17460 and D17461.

§ To whom correspondence should be addressed. Tel.: 81-574-67-5500; Fax: 81-574-67-5310.

¶ The abbreviations used are: GLO, L-gulono- γ -lactone oxidase; bp,

molecular biological study showed that the human genome contains a nucleotide sequence that cross-hybridizes to a cDNA for rat GLO (3). This finding indicates that the complete loss of the GLO locus has not occurred in the human genome and that a nonfunctional gene for GLO is present in the genome. Therefore, it is interesting to elucidate the defect in the human nonfunctional GLO gene² at the nucleotide level. For this purpose we attempted to isolate it and determine its nucleotide sequence for comparison with the sequence of the functional GLO gene of the rat. In addition, we performed the mapping of the GLO gene in the human chromosomes.

EXPERIMENTAL PROCEDURES

Materials—Human genomic DNA (peripheral blood of a Japanese) library in EMBL3 that had been constructed by Dr. Yoshiyuki Sakaki was donated from the Japanese Cancer Resources Bank. The commercial sources of materials used in cloning and sequencing of genomic DNA (4), spot blot hybridization using flow-sorted chromosomes (5), and fluorescence *in situ* hybridization (FISH) (6) were described elsewhere. The isolation of a cDNA clone for rat GLO was described previously (7). The positions of the nucleotide sequence of the cDNA were numbered starting from the initial nucleotide of its translated region.

Isolation and Sequencing of Human Genomic Clones—Approximately 1×10^6 recombinants of a human genomic library in EMBL3 were screened by the plaque hybridization technique using a ³²P-labeled fragment of the rat GLO cDNA (designated 15L in Ref. 3) as a hybridization probe. The methods for library screening, subcloning of isolated DNA, and DNA sequencing were the same as used previously for the study of the guinea pig nonfunctional GLO gene (4), except that subcloning of the insert DNAs was done with fragments produced by digestion with EcoRI plus SalI.

Spot Blot Hybridization Using Flow-sorted Human Chromosomes—Human chromosomes of a normal male karyotype were prepared from diploid B-lymphoblastoid GM00130B cells (NIGMS Human Genetic Mutant Cell Repository, Camden, NJ), and 50,000 chromosomes of each type were sorted directly onto nitrocellulose filters with a FACS440 cell sorter (Becton-Dickinson, Mountain View, CA) (8, 9), and hybridized with ³²P-labeled probe (5) as described previously. The probe used was an insert DNA (nucleotides 3006–3364 shown in Fig. 2B) from one of the deletion mutants that had been prepared for the nucleotide sequencing experiment.

FISH—FISH was carried out by the same method as described previously (6), except that the probe used was a GLO genomic clone DNA (pHGO8B, see Fig. 1).

RESULTS AND DISCUSSION

Cloning of a Segment of the Human GLO Gene—We previously cloned a cDNA encoding GLO from a rat liver cDNA library (7). An EcoRI fragment of the cDNA (nucleotides -23 to 1293, designated 15L), covering almost all of the translated

base pair(s); kbp, kilobase pair(s); FISH, fluorescence *in situ* hybridization.

² This gene has been designated GULO by the nomenclature committee of the Human Genome Mapping Workshop.

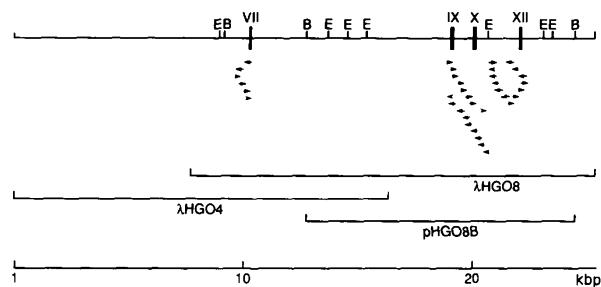


FIG. 1. Structure of a segment of the human nonfunctional GLO gene. The solid vertical lines indicate the regions corresponding to the indicated exons of the rat GLO gene. Restriction endonuclease mapping was carried out only for the region covered by clone λ HGO8. E, EcoRI; B, BamHI. pHGO8B is a plasmid clone used for FISH. Arrows indicate the direction and length of sequencing runs.

region, was used as a probe to screen a human genomic library in EMBL3 by the plaque hybridization technique. As a result, eight overlapping clones were isolated; two of them, designated λ HGO4 and λ HGO8, were analyzed by restriction endonuclease mapping and Southern blotting (Fig. 1). In the analysis of the λ HGO8 DNA digested with EcoRI plus *Sall*, the probe hybridized strongly to two fragments of 5.3 and 4.7 kbp and weakly to a fragment of 2.4 kbp (data not shown). In our previous study (3), the first two fragments were detected by genomic Southern blot analysis using the same probe, although the last one was not detected. This incapability of detection was probably caused by inadequacy of the probe used, because the probe was homologous, at its 3'-end portion, to only a short stretch (103 bp) of a single exon XII-related region contained in the 2.4-kbp fragment, as eventually demonstrated by sequencing analysis. The BamHI fragments of the insert DNA of clone λ HGO8 that hybridized to the probe were 11.6 and 3.6 kbp in length (data not shown), and their sizes were comparable with those of the BamHI fragments detected by genomic Southern blot analysis in our previous study (3). These findings indicate the authenticity of the isolated clone.

The isolated stretch of sequence appears not to cover the entire region of the human nonfunctional GLO gene. In our previous Southern blotting experiment, a positive band could be detected at a position of ~3.6 kbp when a BamHI digest of human genomic DNA was hybridized with a 5'-probe, i.e. the 5'-fragment of rat GLO cDNA (nucleotide -22 to 367, covering exons II-IV and parts of exons I and V) (3). However, the 5'-portion of the human nonfunctional GLO gene was not cloned in the present study, although as many as eight independent clones representing the 3'-half of the gene were obtained. Screening of the same library using the above 5'-probe did not yield any positive clone. The difficulty in cloning of the 5'-portion of the human gene is probably due to the methylation on its DNA, for the efficiency of recovery of recombinants containing methylated genomic insert in conventional cloning systems is known to be low (10).

Sequencing Analysis of the Human Nonfunctional GLO Gene—The regions corresponding to exons VII, IX, X, and XII of the rat GLO gene were definitively identified by sequencing analysis of clone λ HGO8, with reference to the sequences of the respective exons of the rat GLO gene (4) (Fig. 2). Exon VII appears to be deleted in the human nonfunctional gene, since no exon VIII-related sequence was found in the subclones of the cloned human DNA that had hybridized to a rat GLO cDNA (fragment 15L). Although a stretch of sequence in the region between human exon X and exon XII was found to be similar to the middle portion of rat GLO gene exon XI, it was short (34 bp) and the degree of homology was low (68%) compared with the

overall homology (80%) of the above four exon-related sequences. It may be that the similarity of the sequences is fortuitous and that exon XI is also deleted in the human sequence. In connection with this view, the presence of two *Alu* sequences in the region between exon X and exon XII (Fig. 2) indicates that the structure of this region was altered to a great extent by their insertion. Since *Alu* sequences are thought to have been derived from a common progenitor 7SL gene and evolved during primate evolution (11), the events of insertion should have occurred sometime during this period.

When the sequences of the four exon-related regions of the human nonfunctional GLO were compared with the corresponding regions of the rat GLO gene, they were found to have many anomalous nucleotide changes. There were two 1-base deletions, one in the exon VII-related region and the other in the exon X-related region, as well as a 3-base deletion and a 1-base insertion in the exon IX-related region. Two of the intron/exon boundaries identified did not follow the GT/AG rule (12), the dinucleotides flanking the 3'-ends of the exon VII- and exon IX-related regions being GC and GG, respectively, instead of GT. When the human sequences corresponding to rat exons VII, IX, X, and XII (a total of 548 bases) were compared with the respective exons of the rat gene, the overall homology was about 80%. The amino acid sequence deduced from the human exon-related sequences showed still lower homology (~70%) when compared with the corresponding amino acid sequence of rat GLO. Many substitutions were nonconservative according to Dayhoff's conservative category (13), and there were two stop codons in the human sequence. Since both New World and Old World monkeys are deficient in GLO, whereas prosimians possess this enzyme (14), the loss of GLO in primates is thought to have occurred before the divergence of New World monkeys and Old World monkeys (35–45 million years ago) and after the divergence time of the prosimian and simian lineages (50–65 million years ago). Thus the very many substitutions found in the human nonfunctional gene have occurred at random for over some 40 million years since the gene stopped functioning and became free of selective pressure.

Mapping of the Human Nonfunctional GLO Gene—Human metaphase chromosomes of GM00130B cells were separated by flow cytometric sorting after having been stained with propidium iodide (Fig. 3A) or Hoechst 33258 (Fig. 3B). Various deletion mutants prepared in the sequencing analysis experiment were analyzed by Southern blotting with a 32 P-labeled EcoRI-digest of human placental DNA as a probe, and one of the negative clones was chosen and its insert DNA was used as a GLO gene-specific probe for spot blot hybridization of DNA of each chromosome. As a result, the human nonfunctional GLO gene was assigned to chromosome 8 (Fig. 3, C and D). When FISH was performed with a subclone containing a ~12-kbp fragment of λ HGO8 (pHGO8B, see Fig. 1) as a probe, fluorescence signals were detected on the short arm of chromosome 8 (Fig. 4A). Comparison of the FISH signals with Q-banding of the chromosomes enabled us to localize the gene to sub-band 8p21.1 (Fig. 4B). This location was further confirmed by measurement of the physical distances of the gene from the centromere and from the short arm telomere: the map position of the human nonfunctional GLO gene was +38 relative to the centromeric region when expressed by the arbitrary reference point system (6), in which the centromere, short arm telomere, and long arm telomere were given numeric values of 0, +100, and -100, respectively.

With respect to the localization of the human nonfunctional GLO gene, it should be noted that the gene for human glutathione reductase is located at the same sub-band (15). The gene products of these two genes are enzymes participating in the metabolism of two redox-active substances, L-ascorbic acid and

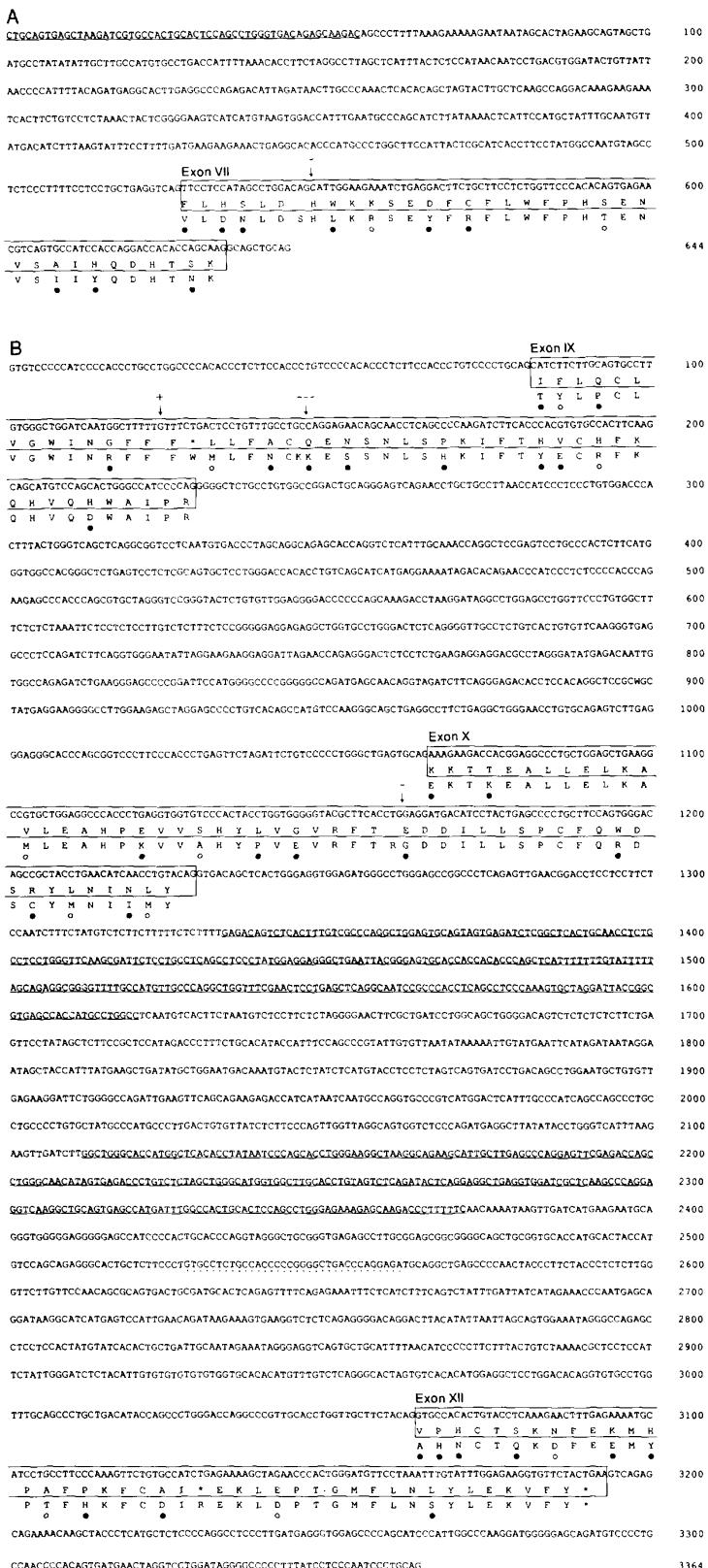


FIG. 3. Assignment of the nonfunctional GLO gene to human chromosome 8. A and B, flow karyotypes of B-lymphoblastoid GM00130B cells. Human chromosomes were prepared from the cells and stained with propidium iodide (A) or Hoechst 33258 (B) and then analyzed with a FACS440 cell sorter. Numbers or X and Y represent the positions of the corresponding chromosome types. The position of chromosome 8 is marked by an arrow. C and D, spot blot hybridization analysis of flow-sorted chromosomes. Chromosomes of each type that had been stained with propidium iodide (C) or Hoechst 33258 (D) were sorted directly onto nitrocellulose filters and hybridized with a 32 P-labeled human GLO gene-specific probe (nucleotides 3006–3364 in Fig. 2B). Positive signals are marked by arrowheads.

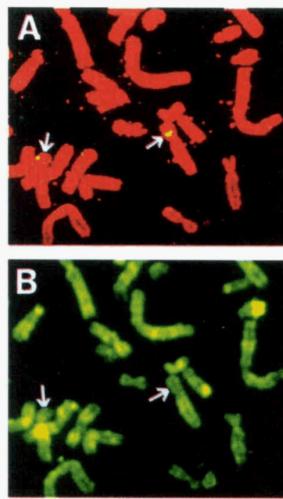
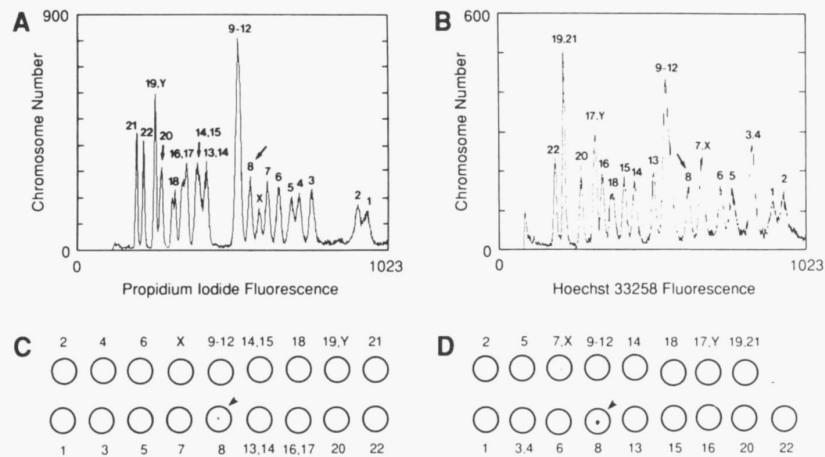


FIG. 4. Regional localization of the nonfunctional GLO gene. A, FISH. Metaphase chromosome spreads on a glass slide were hybridized with a fluorescein isothiocyanate-conjugated GLO gene DNA (pHGO8B DNA), and then counter-stained with propidium iodide. Arrows indicate chromosome 8. B, Q-banding of chromosomes. The preparation used for FISH was stained with quinacrine mustard for Q-banding. Arrows indicate chromosome 8. C, localization of the human nonfunctional GLO gene on chromosome 8. The position of the GLO gene (indicated by double dots) and boundaries of bands, p22/p21.3 and p21.1/p12, are indicated as relative numeric values by use of the centromere (0) and telomeres (+100/-100) as arbitrary reference points.

glutathione, which couple with each other in redox reaction as demonstrated in both *in vitro* and *in vivo* studies (16, 17). In this sense, the close proximity of the genes for GLO and glutathione reductase is interesting in terms of their evolution, although such proximity should be confirmed in other species possessing the functional GLO gene.

Acknowledgment—Experiments using radioisotope were done in the Radioisotope Center Medical Division, University of Nagoya.

REFERENCES

- Nishikimi, M., and Udenfriend, S. (1977) *Trends Biochem. Sci.* **2**, 111–113
- Burns, J. J. (1976) in *Metabolic Pathways* (Greenberg, D. M., ed) Vol. 1, pp. 394–411, Academic Press, New York
- Nishikimi, M., Koshizaka, T., Ozawa, T., and Yagi, K. (1988) *Arch. Biochem. Biophys.* **267**, 842–846
- Nishikimi, M., Kawai, T., and Yagi, K. (1992) *J. Biol. Chem.* **267**, 21967–21972
- Minoshima, S., Kawasaki, K., Fukuyama, R., Maekawa, M., Kudoh, J., and Shimizu, N. (1990) *Cytometry* **11**, 539–546
- Fukuyama, R., Ichijoh, Y., Minoshima, S., Kitamura, N., and Shimizu, N. (1991) *Genomics* **11**, 410–415
- Koshizaka, T., Nishikimi, M., Ozawa, T., and Yagi, K. (1988) *J. Biol. Chem.* **263**, 1619–1621
- Minoshima, S., Amagai, M., Kudoh, J., Fukuyama, R., Hashimoto, T., Nishikawa, T., and Shimizu, N. (1991) *Cytogenet. Cell Genet.* **57**, 30–32
- Sakai, K., Hirai, M., Kudoh, J., Minoshima, S., and Shimizu, N. (1992) *Genomics* **14**, 175–178
- Raleigh, E. A., Murray, N. E., Revel, H., Blumenthal, R. M., Westaway, D., Reith, A. D., Rigby, R. W. J., Elhai, J., and Hanahan, D. (1988) *Nucleic Acids Res.* **16**, 1563–1575
- Deininger, P. L., and Daniels, G. R. (1986) *Trends Genet.* **2**, 76–80
- Mounts, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472
- Dayhoff, M. O., Schwarz, R. M., and Orcutt, B. C. (1978) in *Atlas of Protein Sequences and Structure* (Dayhoff, M. O., ed) Vol. 5, Suppl. 3, pp. 345–352, National Biomedical Research Foundation, Silver Spring, MD
- Nakajima, Y., Shanta, T. R., and Bourne, G. H. (1969) *Histochemistry* **18**, 293–301
- Nevin, N. C., Morrison, P. J., Jones, J., and Reid, M. M. (1990) *J. Med. Genet.* **27**, 135–136
- Winkler, B. S. (1987) *Biochim. Biophys. Acta* **925**, 258–264
- Märtensson, J., and Meister, A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 11566–11568