

Molecular basis for the deficiency in humans of gulonolactone oxidase, a key enzyme for ascorbic acid biosynthesis¹⁻³

Morimitsu Nishikimi and Kunio Yagi

ABSTRACT The inability of humans to synthesize L-ascorbic acid is known to be due to a lack of L-gulono- γ -lactone oxidase, an enzyme that is required for the biosynthesis of this vitamin. Isolation of a cDNA for rat L-gulono- γ -lactone oxidase allowed us to study the basic defect underlying this deficiency at the gene level and led to isolation of a human genomic clone related to L-gulono- γ -lactone oxidase as well as three overlapping clones covering the entire coding region of the rat L-gulono- γ -lactone oxidase cDNA. Sequence analysis study indicated that the human L-gulono- γ -lactone oxidase gene has accumulated a large number of mutations since it stopped being active and that it now exists as a pseudogene in the human genome. *Am J Clin Nutr* 1991;54:1203S-8S.

KEY WORDS Gulonolactone oxidase, ascorbic acid, enzyme deficiency, nucleotide sequence, pseudogene

Introduction

Vitamin C is unique among the vitamins discovered inasmuch as its essentiality as a nutrient is restricted to only a limited number of exceptional species among phylogenetically higher animals. In other words, nearly all higher animals can synthesize vitamin C, or L-ascorbic acid, in their body. In the 1950s, it was demonstrated that L-ascorbic acid is synthesized from D-glucose as shown in **Figure 1** (1). The first portion of the metabolic pathway (from D-glucose to L-gulonic acid) is part of the glucuronic acid cycle, one of the metabolic pathways starting from D-glucose; and the pathway of L-ascorbic acid synthesis branches from L-gulonic acid. The part after the branch point consists of two steps of enzymatic reactions, viz, lactonization and oxidation. The latter oxidation step is catalyzed by L-gulono- γ -lactone oxidase (GLO). This enzyme is missing in the above-mentioned exceptional species, eg, humans, other primates, and guinea pigs (2); as a consequence, these animals cannot synthesize L-ascorbic acid and thus need a dietary intake of vitamin C to prevent scurvy. This enzyme deficiency arose during evolution, and this trait has been carried through generations. In a sense, this is the most prevalent genetic disorder, because all individuals of scurvy-prone species carry this trait.

We have been studying the genetic defect underlying GLO deficiencies in these scurvy-prone animals. Starting with purification and characterization of rat and goat liver GLO (3), we carried out immunochemical studies on the enzyme deficiencies

in the guinea pig and monkey (4). We accomplished isolation of a cDNA for rat-liver GLO (5); as a result, it became possible to investigate the nature of the genetic defect at the gene level (6). In this article, we summarize our molecular biological studies on the genetic basis of the GLO deficiency in humans.

Enzymatic aspects of GLO

Before dealing with the genetic basis of the GLO deficiency, we will first briefly discuss some enzymological aspects of GLO. The distribution of GLO in vertebrates is interesting from a phylogenetic standpoint (7). The enzyme resides in the liver and/or the kidney of L-ascorbic-acid-synthesizing animals: it is present in the kidney of amphibians and reptiles and in the liver of mammals. The enzyme is located in the kidney of primitive species, in the liver of highly evolved species, and in both the kidney and liver of certain species at the intermediate level of evolution (8). The enzyme is missing in certain species of birds (8) as well as in the scurvy-prone mammals mentioned above.

We purified GLO to homogeneity from rat and goat liver (3) and from chicken kidney (9) by improving the method of Nakagawa and Asano (10), by which the enzyme was successfully solubilized from rat liver microsomes by use of appropriate detergents. Characterization of the purified enzyme preparations revealed the following properties. The molecular weight as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis is $\sim 50\,000$ (3, 9). One molecule of flavin adenine dinucleotide is present per molecule of the enzyme, being linked to the N(1) position of a histidyl residue of the apoprotein through the methyl group at the C(8) position of the isoalloxazine ring (9, 11). During the catalysis by GLO, the flavin accepts two electrons from the substrate L-gulono- γ -lactone, and the resulting reduced flavin is oxidized by molecular oxygen. In this process, 2-oxo-L-gulono- γ -lactone and hydrogen peroxide are formed as the products; the former product isomerizes nonenzymatically to form L-ascorbic acid.

¹ From the Institute of Applied Biochemistry, Yagi Memorial Park, Mitake, Gifu, Japan.

² Supported in part by a Grant-in-Aid (01580156) for Scientific Research from the Ministry of Education, Science, and Culture of Japan.

³ Address reprint requests to M Nishikimi, Institute of Applied Biochemistry, Yagi Memorial Park, Mitake, Gifu 505-01, Japan.

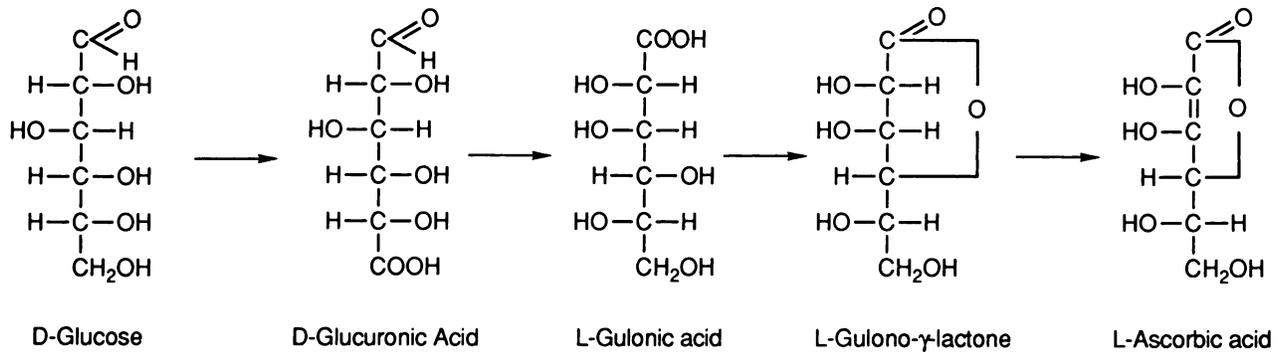


FIG 1. The metabolic pathway of L-ascorbic acid biosynthesis in animals.

Cloning of a cDNA for rat GLO

To elucidate the molecular mechanism underlying the GLO deficiencies in scurvy-prone animals, it is pivotal to obtain a cDNA encoding GLO of an L-ascorbic-acid-synthesizing animal. Because antibody directed against rat GLO was available to us, it was used as a probe to screen a rat-liver cDNA library in an expression vector, λ gt 11. As a result, a cDNA encoding the entire amino acid sequence of rat GLO was isolated (5). The nucleotide sequence of this cloned cDNA and its deduced amino acid sequence are shown in Figure 2. The amino-terminal 33-amino acid sequence of the enzyme was identical with that of the deduced amino acid sequence, and the amino acid composition of rat GLO was in reasonable agreement with that calculated from the deduced amino acid sequence (5). The authenticity of the cloned cDNA was further confirmed by the expression of functional GLO protein by transfection of COS-1 cells with the cDNA that was placed under the control of the SV40 late promoter in the eukaryotic expression vector pSVL (M Nishikimi and K Yagi, unpublished observations, 1991). The size of the enzyme protein synthesized was the same as that of rat GLO, as demonstrated by Western blot analysis. Moreover, the molecular weight presumed from the deduced amino acid sequence agreed with that of rat GLO determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The nucleotide sequence of the cDNA was found to contain an open reading frame of 1320 nucleotides. The sequence surrounding the first ATG codon (ATCATGG) agreed well with the consensus initiation sequence (PCCATGG) described by Kozak (12). Thus we concluded that the clone encodes 440 amino acids of the protein. The amino-terminal methionine is removed to form the mature GLO polypeptide with a presumed molecular weight of 50 483. The mature enzyme with a covalently bound flavin adenine dinucleotide (FAD) is thus presumed to have a molecular weight of 51 267. In fact, this value is comparable to that of rat GLO determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (3).

In addition to the coding region, the cDNA contains a 5'-noncoding region of 22 nucleotides and a 3'-noncoding region of 776 nucleotides. Because the latter region contains neither a poly-A tail nor a polyadenylation signal, the cDNA clone is not of full length. Yet, it was utilized as a versatile tool in our study on the genetic defect in GLO deficiency, as will be discussed below.

Molecular genetic approach to GLO deficiency in humans

Humans are lacking in GLO like other primates and guinea pigs, as mentioned in the introduction. In their investigation into the molecular basis of this deficiency, Nishikimi and Udenfriend (4) raised the question of whether scurvy-prone animals, such as guinea pigs and monkeys, have an aberrant form of GLO. They found by using the Ouchterlony technique that the liver of the guinea pig and of the African green monkey contains no detectable protein immunologically cross-reactive with antirat GLO rabbit antiserum. Their microcomplement fixation test also indicated that guinea pig liver microsomes contain no immunologically reactive protein or < 1/600 of that present in rat liver microsomes. Later, Sato and Udenfriend (13) established a radioimmunoassay for GLO by using the same antibody and showed that the amount of cross-reacting material in human liver microsomes was below the limit of detection. These studies indicated that GLO protein is not present in the liver of the scurvy-prone animals or that it is very aberrant if it is present at all.

The isolation of the rat cDNA enabled us to carry out Northern and Southern blot analyses to examine whether or not scurvy-prone animals possess GLO-specific mRNA in their liver poly(A)⁺ RNA preparation and any DNA sequence related to GLO in their genome. Northern blot analysis showed that no positive signal was detectable with poly(A)⁺ RNA from guinea pig liver under the conditions where rat liver poly(A)⁺ RNA gave a strong signal (6). Thus, it is clear that guinea pigs do not possess any detectable amount of GLO-specific mRNA in their liver. We have not yet determined whether human liver contains GLO-specific mRNA, because we have had no opportunity to test an undegraded preparation of human liver poly(A)⁺ RNA. Yet, humans should have no GLO-specific mRNA, in view of the observation that the GLO gene that was discovered in the human genome has deteriorated to a great extent (*see below*).

The measurement of gene products of the GLO gene at both the protein and mRNA levels indicated that the gene is not expressed in the scurvy-prone animals, as discussed above. The next question to be addressed is whether the gene related to GLO occurs in the genomes of the scurvy-prone animals. With the use of the rat GLO cDNA as a probe, Southern blot analysis of genomic DNAs from various animals gave an unequivocal answer to this question (6). In the hybridization analysis, not only DNAs from L-ascorbic-acid-synthesizing animals (mouse,

| | | | |
|------|--|-----------------------|-----|
| -22 | | GATCCTCTGATCACTGGAATC | |
| 1 | ATGGTCCATGGGTACAAGGGTCCAGTTCAAAATGGGCAAAGACCTATGGTTGCAGTCCAGAGGTACTACCAGCCCACCTCCGTG | | 30 |
| | <u>M V H G Y K G V Q F Q N W A K T Y G C S P E V Y Y Q P T S V</u> | | |
| 91 | GAGGAGGTCAGAGAGGTGCTGGCCCTGGCCCGGAGCAGAAGAAGAAAGTGAAGGTGGTGGTGGCCACTCGCCTTCAGACATTGCC | | 60 |
| | <u>E E V R E V L A L A R E Q K K K V K V V G G G H S P S D I A</u> | | |
| 181 | TGCACTGACGGTTTCATGATCCACATGGGCAAGATGAACCGGTTCTCCAGGTGGACAAGGAGAAGAAGCAGATAACAGTGAAGCCGGT | | 90 |
| | C T D G F M I H M G K M N R V L Q V D K E K K Q I T V E A G | | |
| 271 | ATCCTCCTGGGTGACCTGCACCCACAGCTGGATGAGCATGGCCTGGCCATGTCCAATCTGGGAGCAGTGTCTGATGTGACAGTTGCTGGT | | 120 |
| | I L L A D L H P Q L D E H G L A M S N L G A V S D V T V A G | | |
| 361 | GTCATTGGATCCGGAACACATAACACAGGGATCAAGCACGGCATCTGGCCACTCAGGTGGTGGCCCTGACCCTGATGACAGCTGATGGA | | 150 |
| | V I G S G T H N T G I K H G I L A T Q V V A L T L M T A D G | | |
| 451 | GAAGTTCGGAATGTTCTGAGTCAAGAAATGCAGATGTGTCCAGGCTGCACGGTGCACCTGGGTTGCCTGGGCATCATCTCCACCGTC | | 180 |
| | E V L E C S E S R N A D V F Q A A R V H L G C L G I I L T V | | |
| 541 | ACCTGCAGTGTGGCCTCAGTTTACCTTCAGGAGACATCTCCCTTCGACCCCTCAAAGAGGTCCTTGACAACCTAGACAGCCACCTG | | 210 |
| | T L Q C V P Q F H L Q E T S F P S T L K E V L D N L D S H L | | |
| 631 | AAGAGGTCGAGTACTTCCGCTTCTCTGGTTTCTCACACTGAGAACGTCAGCATCATCTACCAAGACCACCAACAAGGCCCTCC | | 240 |
| | K R S E Y F R F L W F P H T E N V S I I Y Q D H T N K A P S | | |
| 721 | TCTGCATTAAGTGGTTTGGGACTATGCCATCGGGTTTACCTACTGGAGTCTTGCTCTGGACCAGCACCTACCTGCCATGCCTCGTG | | 270 |
| | S A S N W F W D Y A I G F Y L L E F L L W T S T Y L P C L V | | |
| 811 | GGCTGGATCAACCGCTTCTTCTTCTGGATGCTGTTCAACTGCAAGAAGGAGAGCAGCAACCTCAGTCACAAGATCTTACCTACGAGTGT | | 300 |
| | G W I N R F F F W M L F N C K K E S S N L S H K I F T Y E C | | |
| 901 | CGTTCAAGCAGCATGTACAAGACTGGCCATCCTAGGAGAAGACCAAGGAGGCCCTATGGAGCTAAAGGCCATGTGGAGGCCACC | | 330 |
| | R F K Q H V Q D W A I P R E K T K E A L L E L K A M L E A H | | |
| 991 | CCCAAAGTGGTAGCCACTACCCGTAGAGGTGGCCTTACCCGAGGGATGACATTCTGCTGAGCCCTGCTCCAGAGGGACAGCTGC | | 360 |
| | P K V V A H Y P V E V R F T R G D D I L L S P C F Q R D S C | | |
| 1081 | TACATGAACATCATTATGTACAGGCCCTATGAAAGGACGTGCCTCGGCTAGACTACTGGCTGCCTATGAGACCATCATGAAGAAGTTT | | 390 |
| | Y M N I I M Y R P Y G K D V P R L D Y W L A Y E T I M K K F | | |
| 1171 | GGAGGAAGACCCACTGGGCAAAGGCCACAATTGCCACCCAGAAGGACTTTGAGGAAATGTACCCACCTTTCACAAGTTCTGTGACATC | | 420 |
| | G G R P H W A K A H N C T Q K D F E E M Y P T F H K F C D I | | |
| 1261 | CGTGAGAAGCTGGACCCACTGGAATGTTCTTGAATTCGTACCTGGAGAAAGTCTTCTACTAAAGCAGGAGTGGAAACAACCACCTGA | | |
| | R E K L D P T G M F L N S Y L E K V F Y | | |
| 1351 | CCCCTCACACTTCTGCTGCCCGGGGTCTGGGAGCAGAGAAGTGCCTCACAAGCACAATGGGAAGTGCCTCTCCTCTGACCACAA | | |
| 1441 | AGAAAGGCTGGCTCTGGCGGGTCTCTCTGCCTTCGGCATATTCCCTTACATCCAGGCGAAGAAGTGGCCTCTCAATTTCT | | |
| 1531 | GTTAGCATTCCATGGGTCACACATAAACTGCAATCCTCTCAGGAGAAGGGGATCCCTGATACATCATATCCAGACTAAGGATGT | | |
| 1621 | GGTTCTTCTAGATTCTATGGCTCCACCAGGTATAGAGAGATTCTGGGGCTGCAGTTCTCCATCCCTTTCAGAAGGGAGGGATCCCT | | |
| 1711 | TGGCGAGAGTTTGGCTCAGAGGTGGCATGAAGCATGCTCTGCTCTCTTACCCTTGAAGGTCCTTCGGATGCCAGAGATGCTGCTGG | | |
| 1801 | TCCTGGGCAAGCCATATTCAAAGGGTCCAACCTGGCCTTCTGCTGCCATGGCCTGACCCCTCGCAGTGTCTCTCCAGAGGTGTTAG | | |
| 1891 | AGTGGAACTCGCTTCAACCTTAAACAGTTGCTGATCCCTGTGTTCTCTCCCTTCTCCTTGGAGACTACTTGGAGGGGATCCAC | | |
| 1981 | CATGCTCTGGCTTCCCTGGGTATTGTTCTCTCTCTTCCCTTTCACAAATATGATTTCAAGTTGATTTGTTGGCCTTCTGGAGTGTCTCT | | |
| 2071 | TGGAGAACCAAGATGTTCCAGCTAC | | |

FIG 2. The nucleotide sequence of a cDNA encoding rat liver GLO and its deduced amino acid sequence. Revisions were made from our previous paper (5) as follows: a change from G to C at nucleotide position 567 accompanying a change from glutamine to histidine at amino acid residue 189, and deletion of the G at nucleotide position 1 and the Cs at positions 1460 and 2097 of the previous sequence. The amino-terminal amino acid sequence determined by the sequence analysis of purified rat liver GLO is underlined. Adapted from Koshizaka et al (5).

dog, cow, and chicken) but also DNAs from humans and guinea pigs gave positive signals (Fig 3). Accordingly, it is clear that both humans and guinea pigs do have a nucleotide sequence related to rat GLO. However, it should be noted from inspection of Figure 3 that the intensities of the signals for human DNA were far weaker than those for DNAs from the L-ascorbic-acid-synthesizing mammals and even from chickens. This finding indicates that the human GLO gene sequence has become altered to a great extent since the gene stopped functioning.

The structures of the GLO-related genes of humans and guinea pigs were studied in some detail by Southern blot analysis by using a 5'-terminal fragment (nucleotides -22-367) of the rat

GLO cDNA encoding the amino-terminal amino acid sequence and a fragment (nucleotides 947-1293) of the same cDNA encoding the amino acid sequence close to the carboxyl terminus (6). The results demonstrated that the human and guinea pig genes contain coding sequences corresponding to both amino- and carboxyl-terminal parts of GLO, and hence probably the entire coding sequence.

To more quantitatively analyze the alteration of the human GLO gene, we attempted to isolate both the rat and human GLO genes. Screening of partial *EcoRI*- and *HaeIII*-cut rat genomic DNA libraries in Charon 4A led to isolation of three overlapping clones covering the entire coding region of the rat

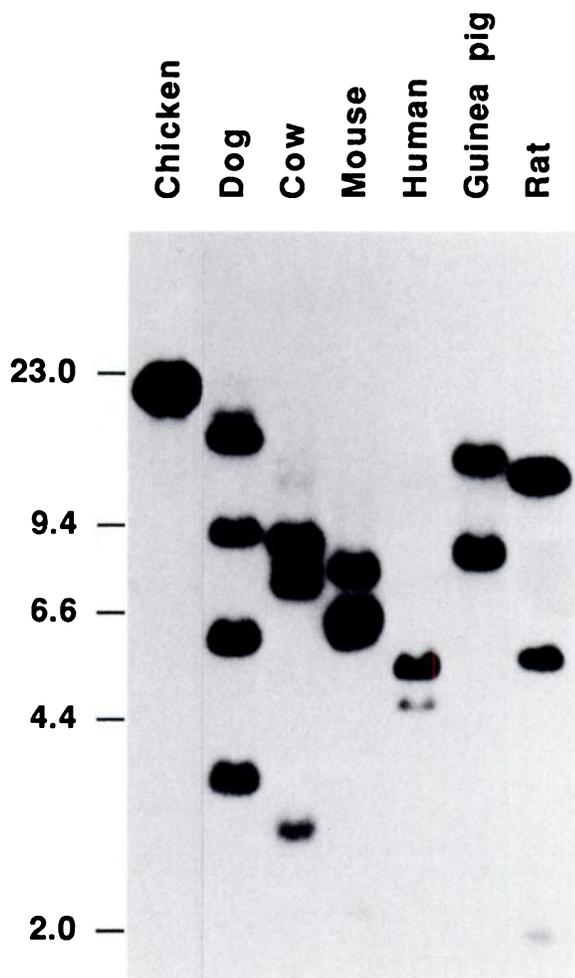


FIG 3. Southern blot analysis of genomic DNAs from various animals. Genomic DNAs from the indicated animals were digested with *Eco*RI and analyzed by Southern blot hybridization. The rat GLO cDNA fragment covering the 5'-noncoding region and most of the coding region (nucleotides -22-1293) was used as a probe. Size markers are shown on the left. Adapted from Nishikimi et al (6).

GLO cDNA (M Nishikimi and K Yagi, unpublished observations, 1991). Sequence analysis showed that all of the intron-exon junction sequences followed the GT/AG rule (14). The coding region was found to be divided into 12 exons by 11 introns. The determined nucleotide sequences of the exon regions were identical with the sequence of the rat GLO cDNA except for three differences. Two differences (G to A at position 252 and C to T at position 432) occur in the third position of the codons for glutamine and threonine, respectively, and most probably are due to allelic polymorphism. The other difference (A to G at position 1211), which alters the codon for glutamine to that for arginine, may also be a polymorphism or may be an artifact that arose during the cloning procedures.

Screening of a human genomic DNA library in EMBL3 resulted in isolation of a clone that contained nucleotide sequences homologous to several exons of the rat GLO gene (M Nishikimi and K Yagi, unpublished observations, 1991). Perhaps, this clone encompasses the regions corresponding to the six rat GLO gene exons that encode from phenylalanine at position 202 to the

carboxyl-terminal tyrosine. The nucleotide sequences of four exon-related regions of the six expected ones were determined and the result is shown in Figure 4. The other two regions corresponding to the two relatively short rat exons were not identified by Southern blot analysis using the rat cDNA as a hybridization probe, even at lower stringency. This may be due to too weak hybridization resulting from a large number of substitutions in the human sequences. In the determined human sequences of the four exon-related regions, there were three deletions (a single nucleotide for two deletions and three nucleotides for one deletion) and one single-nucleotide insertion. It was also found that two out of the eight intron-exon junctions sequenced in the human GLO gene did not conform to the GT/AG rule: the dinucleotides at the splice-donor site of the introns were GC and GG. When the sequences of the human four exon-related regions, a total of 548 nucleotides in number, were compared with those of the corresponding four rat exons, the overall homology was ~80% between the species. At the amino acid level, the homology was yet lower (~70%); and, what is more, many substitutions were not subtle and two stop codons were present in the human sequences. These results led us to conclude that the human GLO gene accumulated a large number of mutations under no selective pressure once it ceased to be active; in other words, the GLO gene now exists as a pseudogene in the human genome.

Molecular evolution of GLO genes

We analyzed the sequence data of Figure 4 by Nei and Gojobori's modified version (15) of the method of Miyata and Yasunaga (16) for calculating the number of amino acid substitutions per synonymous (silent) site (K_s) and per nonsynonymous (amino-acid-altering) site (K_n) between two genes. The values of K_s and K_n corrected for superimposed substitutions (K_s^c and K_n^c , respectively) (17) were 0.86 and 0.16, respectively. Because synonymous substitutions are not subjected to selective pressure during evolution, their number calculated for various genes is similar and an average of the K_s^c value estimated between human and rat genes is 0.823 (18). The value of K_s^c calculated for the GLO genes is comparable to this one, indicating that the synonymous substitutions in the GLO genes occurred at the same rate as in other genes.

In regard to the K_n^c value obtained, we can make the following discussion on the basis of the neutral theory of evolution. It may be reasonable to suppose that the substitution rate at nonsynonymous sites is far less in the rat GLO gene than in the human GLO gene, because nonsynonymous substitutions in the rat gene have been restricted by the selective pressure during evolution whereas the nonsynonymous substitutions in the human gene have occurred as frequently as the synonymous substitutions since this gene stopped being active. Therefore, the number of substitutions per nonsynonymous site that occurred between the human gene and the common ancestor gene could be close to the value of 0.16, which was obtained as the number of substitutions per nonsynonymous site between the rat and human GLO genes. On this assumption, we can date the primates' loss of GLO at ~70 million years ago, assuming the rate of substitutions at neutral sites as 2.3×10^{-9} substitution \cdot site $^{-1} \cdot$ y $^{-1}$ for the primate lineage (19). The estimated date corresponds to the time of mammalian radiation. By letting the substitution rate at the synonymous sites be equal to that at the nonsynonymous

sites in the human GLO gene, we can calculate the rate of synonymous substitution in the rat GLO gene to be $\sim 10 \times 10^{-9}$ substitution \cdot site $^{-1} \cdot$ y $^{-1}$. Interestingly, the synonymous substitution rates calculated for the rat and human GLO genes are consistent with the presumption by Li and Tanimura (19) that the average synonymous substitution rate in rodents is 4–10 times higher than that in higher primates. As to the date of the loss of the L-ascorbic-acid-synthesizing ability in primates, it is thought to have been before the divergence of New World monkeys and Old World monkeys (35–45 million years ago) because monkeys of both lineages are reported to be unable to synthesize L-ascorbic acid (20). Thus, the above calculations seem reasonable, though they are very rough and based on a number of assumptions. 

We thank the Japanese Cancer Resources Bank for donating an EMBL3 human genomic DNA library that was constructed by Yoshiyuki Sakaki.

References

- Burns JJ. Biosynthesis of L-ascorbic acid: basic defect in scurvy. *Am J Med* 1959;26:740–8.
- Burns JJ. Missing step in man, monkey and guinea pig required for the biosynthesis of L-ascorbic acid. *Nature* 1957;180:553.
- Nishikimi M, Tolbert B, Udenfriend S. Purification and characterization of L-gulonolactone oxidase from rat and goat liver. *Arch Biochem Biophys* 1976;175:427–35.
- Nishikimi M, Udenfriend S. Immunologic evidence that the gene for L-gulonolactone oxidase is not expressed in animals subject to scurvy. *Proc Natl Acad Sci USA* 1976;73:2066–8.
- Koshizaka T, Nishikimi M, Ozawa T, Yagi K. Isolation and sequence analysis of a complementary DNA encoding rat liver L-gulonolactone oxidase, a key enzyme for L-ascorbic acid biosynthesis. *J Biol Chem* 1988;263:1619–21.
- Nishikimi M, Koshizaka T, Ozawa T, Yagi K. Occurrence in humans and guinea pigs of the gene related to their missing enzyme L-gulonolactone oxidase. *Arch Biochem Biophys* 1988;267:842–6.
- Chatterjee IB. Evolution and the biosynthesis of ascorbic acid. *Science* 1973;182:1271–2.
- Chaudhuri CR, Chatterjee IB. L-Ascorbic acid synthesis in birds: phylogenetic trend. *Science* 1969;164:435–6.
- Kiuchi K, Nishikimi M, Yagi K. Purification and characterization of L-gulonolactone oxidase from chicken kidney microsomes. *Biochemistry* 1982;21:5076–82.
- Nakagawa H, Asano A. Ascorbate synthesizing system in rat liver microsomes I. Gulonolactone-reducible pigment as a prosthetic group of gulonolactone oxidase. *J Biochem (Tokyo)* 1970;68:737–46.
- Kenney WC, Edmondson DE, Singer TP, Nakagawa H, Asano A, Sato R. Identification of the covalently bound flavin of L-gulonolactone oxidase. *Biochem Biophys Res Commun* 1976;71:194–200.
- Kozak M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNA. *Nucleic Acids Res* 1984;12:857–72.
- Sato P, Udenfriend S. Scurvy-prone animals, including man, monkey, and guinea pig, do not express the gene for gulonolactone oxidase. *Arch Biochem Biophys* 1978;187:158–62.
- Mount SM. A catalogue of splice junction sequences. *Nucleic Acids Res* 1982;10:459–72.
- Nei G, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 1986;3:418–26.
- Miyata T, Yasunaga T. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *J Mol Evol* 1980;16:23–6.
- Jukes T, Cantor CR. Evolution of protein molecules. In: Munro HN, ed. *Mammalian protein metabolism III*. New York: Academic Press, 1969:21–123.
- Hayashida H, Miyata T. Unusual evolutionary conservation and frequent DNA segment exchange in class I gene of the major histocompatibility complex. *Proc Natl Acad Sci USA* 1983;80:2671–5.
- Li W-H, Tanimura M. The molecular clock runs more slowly in man than in apes and monkeys. *Nature* 1987;326:93–6.
- Stone I. Studies of a mammalian enzyme system for producing evolutionary evidence on man. *Am J Phys Anthropol* 1965;23:83–6.

