Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 7401–7404, August 1991
Evolution

Gene sequences suggest inactivation of α-1,3-galactosyltransferase in catarrhines after the divergence of apes from monkeys

(primate evolution/molecular evolution/gene inactivation/PCR sequencing)

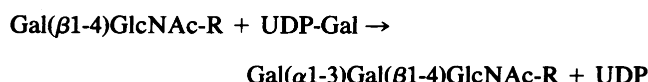
URI GALILI*†‡ AND KAREN SWANSON*

*Department of Laboratory Medicine, University of California, San Francisco, CA 94143; and †Department of Microbiology and Immunology, The Medical College of Pennsylvania, Philadelphia, PA 19129

Communicated by Charles G. Sibley, June 3, 1991 (received for review April 10, 1991)

ABSTRACT **The glycosylation enzyme α-1,3-galactosyltransferase (α1,3GT; UDPgalactose:β-D-galactosyl-1,4-N-acetyl-D-glucosaminide α-1,3-galactosyltransferase, EC 2.4.1.151) displays a unique pattern of distribution in mammals. It synthesizes an abundance of Gal(α1-3)Gal(β1-4)GlcNAc-R (α-galactosyl) epitopes within the Golgi apparatus of cells of nonprimate mammals, prosimians, and New World monkeys (platyrrhines). The catarrhines, which include Old World monkeys, apes, and humans, lack this enzyme activity because of the inactivation of the α1,3GT gene. In contrast, the catarrhines produce large amounts of antibodies, designated anti-Gal, against the α-galactosyl epitope. The inactivation of the α1,3GT gene in ancestral catarrhines was probably the result of an intensive evolutionary pressure for alteration in the makeup of cell surface carbohydrates (i.e., suppression of α-galactosyl epitope expression) and for the production of the anti-Gal antibody. To determine the period in which the α1,3GT gene was inactivated in ancestral catarrhines, comparative sequencing of a 370-base-pair region of this gene was performed by polymerase chain reactions with DNA of various primates. The data suggest that α1,3GT inactivation occurred rather late in the course of catarrhine evolution (less than 28 million years ago), as separate events in apes and in Old World monkeys, after the two groups diverged from each other.**

The catarrhines [Old World monkeys (OWM), apes, and humans] differ from other mammals in the expression of the carbohydrate epitope Gal(α1-3)Gal(β1-4)GlcNAc-R (the α-galactosyl epitope). Whereas the α-galactosyl epitope is abundantly expressed on cells ($>10^6$ epitopes per cell) and on secreted glycoproteins of nonprimate mammals, prosimians (lemurs), and New World monkeys (NWM, or platyrrhines), it is not found in catarrhines (1–4). Humans and other catarrhines, however, produce large amounts of a natural antibody (1% of circulating IgG) that interacts specifically with the α-galactosyl epitope and has been designated anti-Gal (5–10). The biosynthetic basis for this distribution pattern of the α-galactosyl epitope in mammals is the differential activity of the glycosylation enzyme α-1-3-galactosyltransferase (α1,3GT, EC 2.4.1.151). α1,3GT synthesizes the α-galactosyl epitope within the Golgi apparatus of nonprimate mammal, prosimian, and NWM cells by the following reaction (2, 11–14):



Gal(β1-4)GlcNAc-R + UDP-Gal →

Gal(α1-3)Gal(β1-4)GlcNAc-R + UDP

In contrast, α1,3GT activity is diminished in catarrhine cells (2). Inactivation of α1,3GT in ancestral catarrhines is likely to have resulted from an evolutionary event that exerted a selective pressure for suppression of α-galactosyl epitope expression and the subsequent appearance of a natural antibody (anti-Gal) in amounts manyfold higher than those of any other known natural antibody (15, 16). One of us previously speculated (17) that this major molecular change in cell surface carbohydrate makeup, which seems to represent a landmark in catarrhine evolution, also may be of pathologic significance if the α-galactosyl epitope is aberrantly expressed on human cells, since the interaction of the natural anti-Gal antibody with this carbohydrate epitope could result in the initiation of autoimmune processes. The purpose of the present study was to determine the evolutionary period in which the α1,3GT gene underwent inactivation in ancestral catarrhines.

Studies in molecular aspects of evolution have indicated that the evolutionary history of organisms is inscribed in the DNA of contemporary species (18). This principle has been demonstrated in the phylogenetic relationships among various primate species, as indicated by similarities in DNA sequences of homologous genes (19–21). We assumed that studying the α1,3GT gene sequences in primates might provide information on the evolutionary period in which this gene underwent inactivation in catarrhines. The working assumption has been that the smaller the difference between human, ape, or OWM sequences (no α1,3GT activity) and the sequences in NWM (active α1,3GT), the later the inactivation of this gene must have occurred in the evolution of the catarrhines. This assumption was based on one of the main principles in molecular evolution, which states that sequences in genes evolve in a "clockwise" manner and that the rate of mutation is faster in silent DNA than in coding sequences because it is not impeded by natural selection (18). The similarity in α1,3GT gene sequences and the occurrence of specific nucleotide sequences in various primate groups imply that this enzyme underwent suppression in catarrhines after OWM and apes diverged from each other.§

MATERIALS AND METHODS

DNA Source. The DNAs studied were extracted from fibroblasts of three hominoids (apes) [pygmy chimpanzee (*Pan paniscus*), gorilla (*Gorilla gorilla*), and orangutan (*Pongo pygmaeus*], two OWM [rhesus monkey (*Macaca mulatta*) and patas monkey (*Erythrocebus patas*)], and five NWM [spider monkey (*Ateles geoffroyi*), squirrel monkey (*Saimiri sciureus*), titi (*Callicebus molloch*), wooly monkey (*Lagothrix lagotricha*), and howler monkey (*Alouatta caraya*)]. The fibroblast cell lines were obtained from the Na-

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.

Abbreviations: α1,3GT, α-1,3-galactosyltransferase; NWM, New World monkey(s); OWM, Old World monkey(s).

‡To whom reprint requests should be addressed at: Department of Microbiology and Immunology, The Medical College of Pennsylvania, 3300 Henry Avenue, Philadelphia, PA 19129.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M72426).

tional Institute on Aging Cell Repository (Camden, NJ). DNA of an African green monkey (*Cercopithecus aethiops*, an OWM) was extracted from the COS-7 cell line and baboon (*Papio papio*, OWM) DNA was extracted from blood leukocytes. Some of the sequences were confirmed with DNA samples of chimpanzee, gorilla, orangutan, langur (*Pygathrix nemaeus*, OWM) and spider monkey received from O. A. Ryder (Center for Reproduction of Endangered Species, San Diego Zoo), and DNA of rhesus monkey purchased from Clontech.

DNA Sequencing by Polymerase Chain Reaction (PCR). Generation of single-stranded DNA by PCR and its application to direct sequencing were carried out as a slight modification of the method of Gyllenstein and Erlich (22). Genomic DNA (1 μg) was subjected to PCR amplification (40 cycles of 1 min at 94°C, 1 min at 55°C, and 1.5 min at 72°C) with 50 pmol of the primers 5′-GTCATATTTTACATCATG-GTGGAT-3′ and 5′-TATCTGAAGGCAGGCCTATAT-GAT-3′ (bases 559–582 and complementary sequence of bases 1121–1098, respectively, in the human α1,3GT sequence; ref. 23). These primers were chosen based on their high degree of conservation in the human (23), murine (24), and bovine (25) α1,3GT gene. The resulting 564-base-pair (bp) amplified double-stranded DNA molecules were purified by electrophoresis in 1% agarose and electroelution of the amplified DNA. Subsequently, 10–50 ng of the purified product were subjected to asymmetric PCR (50 pmol of the upstream primer and 1 pmol of the downstream primer and the reverse combination) for the generation of single-stranded DNA of both strands. The products were isolated on Qiagen P-25 columns (Qiagen, Studio City, CA) and subjected to direct sequencing with the corresponding upstream or downstream primer using the Sequenase kit (United States Biochemical) and [α-[^{35}S]thio]dATP. The mixture was loaded in three successive loads on a 6% polyacrylamide/7 M urea gel, electrophoresed, and exposed overnight to Kodak X-Omat film, as described (22). Both strands of DNA were analyzed. In DNA samples where fewer than 150 overlapping bases were found between the two sequenced strands, two additional internal primers were used for sequencing. The additional oligonucleotides were 5′-CACGAGGTCGACT-TCCTCTTCT-3′ and 5′-CTTTCATCATGCCACTTG-GCTT-3′ (bases 723–744 and complementary sequence of bases 1029–1008, respectively, in the human α1,3GT sequence; ref. 23).

RESULTS AND DISCUSSION

The cDNA of α1,3GT has been cloned from a cow cDNA library by Joziasse *et al.* (25) and from a mouse cDNA library by Larsen *et al.* (24). Using these probes, both groups found homologous regions in human DNA and thus suggested that this gene has been conserved within the human genome in a nonexpressed form (23, 25). Furthermore, Larsen *et al.* (23) used the mouse α1,3GT cDNA probe to clone a homologous region from a human genomic library. The cloned human DNA represents an 801-bp exon at the 3′ end of the α1,3GT gene, and it displays 81% and 88% sequence identity with the mouse and cow α1,3GT cDNAs, respectively (23–25). We used these mouse (24), cow (25), and human (23) α1,3GT sequences to design oligonucleotide primers that have been evolutionarily conserved in these three species. A 514- to 516-bp region of the α1,3GT DNA of various primates was amplified with these oligonucleotides by PCR. Single-stranded DNA was produced by asymmetric PCR and sequenced using the two terminal oligonucleotides, as well as internal oligonucleotides, as primers (Fig. 1). With this approach, between 210 bp and 430 bp of the α1,3GT gene were sequenced from 13 different platyrrhine and catarrhine species, as listed in Table 2. The sequence of 370 bp from nine of these primate species, along with the human (23) and cow (25) sequences, are aligned in Fig. 1. Pairwise comparison, expressed as percentage of nucleotide changes (i.e., substitutions and deletions) are shown in Table 1. No PCR amplification was obtained with DNA samples from lemur (*Lemur catta*), marmoset (*Callithrix jaccus*), tamarin (*Saguinus fuscicollis*), and siamang (*Hylobates syndactylus*), suggesting that the primers used in this study were not suitable for amplifying the α1,3GT gene in these species.

Comparison of the cow α1,3GT sequence (used as an external reference) with that of NWM or OWM illustrates the principle of the evolutionary clock (Fig. 1, Table 1). Primate and ungulate lineages diverged from each other 75–110 million years ago (26). Comparison of overall differences in sequences of pseudogenes that have been silent in both OWM and NWM lineages, such as the η-globin pseudogene, demonstrates a 38.2% and 37.5% difference in goat vs. rhesus monkey (OWM) and goat vs. owl monkey (*Aotus trivirgatus*, NWM), respectively (19). The α1,3GT gene that has been active in the lineage of NWM and that of cow exhibit, as expected, a much smaller difference, ranging between 9.7% and 11.9% (average, 10.9%). However, since this gene has been inactivated during the evolution of OWM, the differences between OWM and cow sequences are greater than those of NWM and cow, and range between 13.2% and 15.1% (average, 14.1%). The small differences in α1,3GT sequences among OWM (1.8–3.6%) or apes (0.3–4.9%) also indicate that these differences indeed reflect phylogenetic relationships. The 3.2% human/orangutan difference is similar to that observed in the η-globin pseudogene (19), suggesting that α1,3GT has likewise been inactive since the divergence of the orangutan lineage from the human lineage (estimated to be 12–18 million years ago; refs. 19, 20, 26–29). Furthermore, the 7.6–8.1% human/OWM differences are also equal to that of the η-globin pseudogene, suggesting that, for much of the period since the divergence of apes from OWM (estimated to be 20–28 million years ago; refs. 19, 20, 26–29), the α1,3GT gene, like the η-globin pseudogene, has been conserved in catarrhines in an inactive form. The human and chimpanzee sequences are almost identical, with only one base difference. This finding supports the suggestion that these two species diverged from each other after their common ancestor diverged from the gorilla lineage (21, 27–29).

The α1,3GT gene was probably not inactivated early in the evolution of catarrhines. This is suggested by the comparison of the human sequence with the NWM sequences. There is only a 4.6% difference between human and spider monkey sequences and 6.2% between human and squirrel monkey. These differences are less than those expected, had the gene been suppressed immediately after divergence of catarrhines from platyrrhines, ≈40 million years ago (30). The expected difference would have been at least 8.0% [based on 0.14% mutations per million years in the ape η-globin pseudogene (19)] and 0.07% mutations per million years in NWM α1,3GT gene (as calculated from the difference between NWM and cow sequences). Furthermore, the human/spider monkey difference is only 1.4-fold greater than the human/orangutan difference, suggesting that the gene was inactivated no more than 17–25 million years ago. The overall higher OWM/NWM difference as compared with ape/NWM may reflect the reported tendency of certain DNA regions to mutate faster in OWM than in apes (19, 26, 27).

Additional information on the period of α1,3GT inactivation could be obtained by comparison of specific nucleotide changes in the various primates. Larsen *et al.* (23) reported two deletion mutations at positions 822 and 904 in the human α1,3GT DNA sequence and suggested that the premature stop codons generated by the frameshifts may be the cause of α1,3GT inactivation in humans (Fig. 1). Sequencing data on the region of these two deletion mutations in primates re-

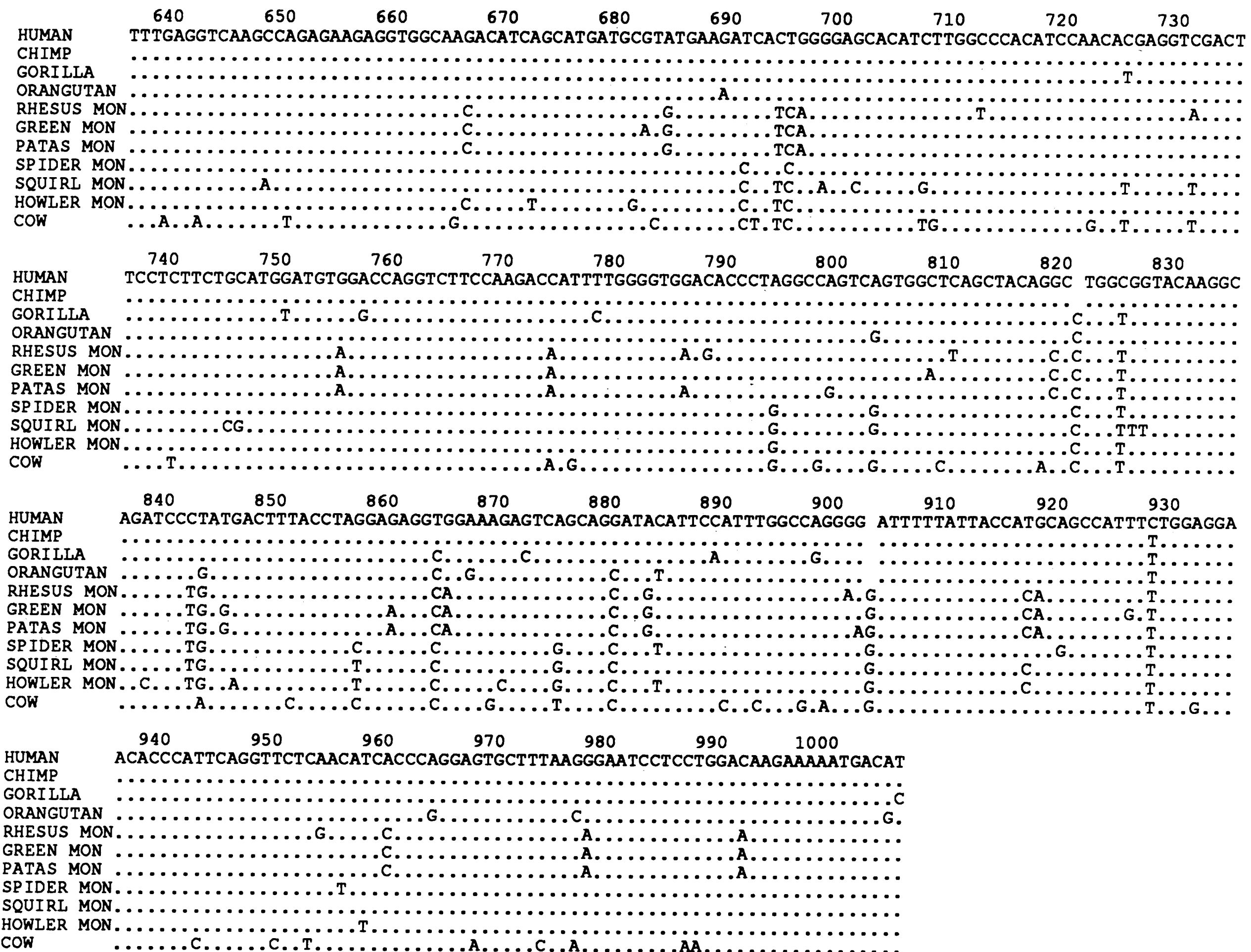


FIG. 1. Aligned DNA sequences of a 370-bp region in the 3′ portion of the α1,3GT gene from various primates, human (23), and domestic cow (25). The base numbers in this figure are according to the open reading frame of the mouse α1,3GT cDNA described by Larsen *et al.* (23, 24). The numbered base is under the second digit. Dots represent sequences identical to that of the human α1,3GT gene (23). Sequences of additional NWM and OWM species described in Table 2 had less than 370 bases and were not included in this figure.

vealed a distinct evolutionary pattern. OWM, like NWM, have no such deletions (Fig. 1, Table 2). Among apes, orangutan and gorilla have only the second deletion (position 904), whereas chimpanzees, like humans, have both deletions. These findings suggest that the deletion at position 904, and subsequently that at position 822, may have evolved as a means of inactivating α1,3GT only in apes, and that OWM have undergone another set of mutations to inactivate this gene. The implication of this observation, therefore, is that the selective pressure for inactivating α1,3GT affected the catarrhines after the divergence of apes from OWM, and that this inactivation took place in ancestral OWM and in apes as separate events independent of each other. The actual mechanism(s) inactivating the α1,3GT gene in these two catarrhine groups is not clear. Joziasse *et al.* (25) detected no α1,3GT mRNA in African green monkey and human cells, suggesting diminished transcription of this gene in catarrhines. Nevertheless, structural mutations, such as the two deletions within the human α1,3GT gene (23), might also have played an important role in inactivating this gene.

A detailed comparison of sequences in Fig. 1 reveals that OWM share specific sequences that are not found in apes or

Table 1. Pairwise comparison of nucleotide changes in a 370-bp 3′ region of the α1,3GT gene

	% changes									
Species	Human	Chimpanzee	Gorilla	Orangutan	Rhesus monkey	Green monkey	Patas monkey	Squirrel monkey	Spider monkey	Howler monkey
Chimpanzee (ape)	0.3	—								
Gorilla (ape)	3.0	3.0	—							
Orangutan (ape)	3.2	3.0	4.9	—						
Rhesus monkey (OWM)	8.1	7.8	8.9	8.6	—					
Green monkey (OWM)	7.6	7.6	7.0	7.6	3.2	—				
Patas monkey (OWM)	7.6	7.3	8.1	8.1	2.7	1.6	—			
Squirrel monkey (NWM)	7.0	7.0	7.3	7.0	8.9	8.6	8.6	—		
Spider monkey (NWM)	4.6	4.3	5.4	4.1	7.8	7.0	7.3	4.3	—	
Howler monkey (NWM)	6.2	5.9	6.5	6.2	7.6	7.3	7.0	5.1	5.3	—
Cow	12.4	12.1	12.1	13.5	15.1	13.2	14.1	11.1	9.7	11.9

Nucleotide changes include base substitutions and deletions (expressed as percent). Data are based on Fig. 1.

Table 2. Comparison of sequences at the sites of the two frameshift mutations within the human α1,3GT gene

Species	Bases 820–824 (mutation 1)	Bases 902–906 (mutation 2)
Human	GC TG	GG AT
Chimpanzee (ape)	GC TG	GG AT
Gorilla (ape)	GCCTG	GG AT
Orangutan (ape)	GCCTG	GG AT
Rhesus monkey (OWM)	CCCTG	AGGAT
Green monkey (OWM)	CCCTG	GGGAT
Patas monkey (OWM)	CCCTG	GAGAT
Baboon (OWM)	CCCTG	AGGAT
Langur (OWM)	CCCTG	GGGAT
Squirrel monkey (NWM)	GCCTG	GGGAT
Spider monkey (NWM)	GCCTG	GGGAT
Titi (NWM)	GCCTG	GGGAT
Howler monkey (NWM)	GCCTG	GAGAT
Wooly monkey (NWM)	GCCTG	GGGAT
Domestic cow	GCCTG	GGGAT

Base numbers are according to mouse α1,3GT cDNA described by Larsen *et al.* (23, 24).

NWM. Such sequences include C in position 667 (C-667), G-685, A-697, A-775, C-820, A-866, G-884, A-919, C-961, A-979, and A-993. These sequences, which are unique to OWM (found also in baboon and langur; data not shown), support the assumption that ancestral OWM had carried the active α1,3GT gene for a period after divergence from apes, and that the functional constraints on this gene resulted in the conservation of the many unique sequences shared by different OWM species at present. It may also be possible that some of the amino acid substitutions resulting from these shared mutations impair the catalytic activity of α1,3GT in OWM and have been conserved to prevent reactivation of α1,3GT [and the ensuing autoimmune phenomena (17)] in this group of primates. Other nucleotides—such as C-696, shared by NWM and OWM but not found in apes; G-844 and C-881, shared by NWM, OWM, and orangutan; C-865, shared by NWM, OWM, orangutan, and gorilla; and T-929, shared by all primates except for humans—indicate that many of the substitution mutations reflect a sequential rather than a random pattern, further supporting the notion that the α1,3GT gene has been subjected to functional constraints for long periods in the course of catarrhine evolution.

We have speculated that the inactivation of α1,3GT in the catarrhines resulted from a selective pressure exerted by a pathogen that was detrimental to primates and was endemic to the Old World. Such a pathogen could affect primates via cell surface receptors containing the α-galactosyl epitope. Alternatively, this hypothetical pathogen may have expressed α-galactosyl-like epitopes (2, 17). The exerted selective pressure is likely to have resulted in the evolution of catarrhine species that suppressed autologous α-galactosyl epitope synthesis. This suppression led to loss of immune tolerance toward the α-galactosyl epitope and thus enabled the production of the anti-Gal antibody as a protective measure against pathogens expressing similar carbohydrate structures (2, 17). This evolutionary event, which seems to have occurred in the early Miocene after the divergence of apes from monkeys, probably less than 28 million years ago (30), would have had a considerable effect on the size and diversity of both OWM and ape populations during the Miocene. It is tempting to link these molecular evolutionary observations with the fossil record, which has indicated the decline of ape (Hominoidea) species diversity during the middle and late Miocene, and the suppression of OWM (Cercopithecoidea) diversity to only three species until the late Miocene (31). One may speculate that the environmental factors that seem to have caused inactivation of the α1,3GT gene in catarrhines could have played a major role in determining the size and diversity of primate populations in the Old World during the Miocene, as reflected in the fossil record.

We thank Dr. Oliver A. Ryder of Center for Reproduction of Endangered Species, San Diego Zoo, for the primate DNA samples; Dr. Arthur E. Greene of the Cell Repository, National Institute on Aging, for the primate cell lines; Dr. Naomi Galili for the helpful discussions; and Mr. James Harris for typing the manuscript. This research was supported by National Institutes of Health Grants AG06299 and GM40205.

1. Galili, U., Clark, M. R., Shohet, S. B., Buehler, J. & Macher, B. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1369–1373.
2. Galili, U., Shohet, S. B., Kobrin, E., Stults, C. L. M. & Macher, B. A. (1988) *J. Biol. Chem.* **263**, 17755–17762.
3. Spiro, R. G. & Bhoyroo, V. D. (1984) *J. Biol. Chem.* **259**, 9858–9866.
4. Thall, A. & Galili, U. (1990) *Biochemistry* **29**, 3959–3965.
5. Galili, U., Rachmilewitz, E. A., Peleg, A. & Flechner, I. (1984) *J. Exp. Med.* **160**, 1519–1531.
6. Galili, U., Macher, B. A., Buehler, J. & Shohet, S. B. (1985) *J. Exp. Med.* **162**, 573–582.
7. Galili, U., Buehler, J., Shohet, S. B. & Macher, B. A. (1987) *J. Exp. Med.* **165**, 693–704.
8. Towbin, H., Rosenfelder, G., Wieslander, J., Avila, J. L., Rojas, M., Szarfman, A., Esser, K., Nowack, H. & Timple, R. (1987) *J. Exp. Med.* **166**, 419–432.
9. Wieslander, J., Månsson, O., Kallin, E., Gabrielli, A., Nowack, H. & Timple, R. (1990) *Glycoconjugate J.* **7**, 85–100.
10. Davin, J. C., Malaise, M., Foidart, J. M. & Mahieu, P. (1982) *Kidney Int.* **31**, 1132–1139.
11. Basu, M. & Basu, S. (1973) *J. Biol. Chem.* **248**, 1700–1706.
12. Betteridge, A. & Watkins, W. M. (1983) *Eur. J. Biochem.* **132**, 29–35.
13. Blake, D. A. & Goldstein, I. J. (1981) *J. Biol. Chem.* **256**, 5387–5393.
14. Blanken, W. M. & Van den Eijnden, D. H. (1985) *J. Biol. Chem.* **260**, 12927–12934.
15. Guilbert, B., Dighiero, G. & Avrameas, S. (1982) *J. Immunol.* **128**, 2779–2787.
16. Lutz, H. U., Flepp, R. & Stringaro-Wipf, G. (1984) *J. Immunol.* **133**, 2610–2617.
17. Galili, U. (1989) *Lancet* **ii**, 358–361.
18. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639.
19. Koop, B. F., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. (1986) *Nature (London)* **319**, 234–238.
20. Andrews, P. (1986) in *Major Topics in Primate and Human Evolution*, eds. Wood, B., Martin, L. & Andrews, P. (Cambridge Univ. Press, Cambridge, U.K.), pp. 107–129.
21. Ruovolo, M., Disotell, T. R., Allard, M. W., Brown, W. M. & Honeycutt, R. L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1570–1574.
22. Gyllenstein, U. B. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7652–7656.
23. Larsen, R. D., Rivera-Marrero, C. A., Ernest, L. K., Cummings, R. D. & Lowe, J. B. (1990) *J. Biol. Chem.* **263**, 7055–7061.
24. Larsen, R. D., Rajan, V. P., Ruff, M. M., Kukowska-Latalo, J., Cummings, R. D. & Lowe, J. B. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8227–8231.
25. Joziasse, D. H., Shaper, J. H., Van den Eijnden, D. H., Van Tunen, A. H. & Shaper, N. L. (1989) *J. Biol. Chem.* **264**, 14290–14297.
26. Britten, R. J. (1986) *Science* **231**, 1393–1399.
27. Li, W. & Tamimura, M. (1987) *Nature (London)* **326**, 93–96.
28. Miyamoto, M. M., Slightom, J. L. & Goodman, M. (1987) *Science* **238**, 369–373.
29. Sibley, C. G. & Ahlquist, J. E. (1987) *J. Mol. Evol.* **26**, 99–121.
30. Pilbeam, D. (1984) *Sci. Am.* **250** (3), 84–97.
31. Andrews, P. (1981) in *Aspects of Human Evolution*, ed. Stringer, C. B. (Taylor & Francis, London), pp. 25–61.