

2/16/02

## Molecular Basis of Evolutionary Loss of the $\alpha$ 1,3-Galactosyltransferase Gene in Higher Primates\*

Received for publication, November 1, 2001, and in revised form, December 11, 2001  
Published, JBC Papers in Press, December 28, 2001, DOI 10.1074/jbc.M110527200

Chihiro Koike<sup>‡§¶</sup>, John J. Fung<sup>‡</sup>, David A. Geller<sup>‡</sup>, Reiji Kannagi<sup>||</sup>, Therese Libert<sup>§</sup>,  
Patrizia Luppi<sup>§</sup>, Izumi Nakashima<sup>\*\*</sup>, Jennifer Profozich<sup>§</sup>, William Rudert<sup>§</sup>,  
Sugandha B. Sharma<sup>‡§</sup>, Thomas E. Starzl<sup>‡</sup>, and Massimo Trucco<sup>§</sup>

From the <sup>‡</sup>Thomas E. Starzl Transplantation Institute and Department of Surgery, University of Pittsburgh and the <sup>§</sup>Division of Immunogenetics, Department of Pediatrics, Children's Hospital of Pittsburgh, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, the <sup>||</sup>Department of Pathology, Aichi Cancer Research Center, 81-1159 Kanokoden, Chikusa, Nagoya 465, Japan, and the <sup>\*\*</sup>Department of Immunology, Nagoya University School of Medicine, 65 Tsurumai, Showa, Nagoya 466, Japan

Galactose- $\alpha$ 1,3-galactose ( $\alpha$ Gal) epitopes, the synthesis of which requires the enzyme product of  $\alpha$ 1,3-galactosyltransferase ( $\alpha$ 1,3GT), are sugar chains on the cell surface of most mammalian species. Notable exceptions are higher primates including Old World monkeys, apes, and humans. The  $\alpha$ Gal-negative species as well as mice with deletion of the  $\alpha$ 1,3GT gene produce abundant anti- $\alpha$ Gal antibodies. The evolutionary loss of  $\alpha$ Gal epitopes has been attributed to point mutations in the coding region of the gene. Because no transcripts could be found in the higher primate species with Northern blot analysis, a potential alternative explanation has been loss of upstream regulation of the gene. Here, we have demonstrated that the rhesus promoter is functional. More importantly, a variety of full-length transcripts were detected with sensitive PCR-based methods in the tissues of rhesus monkeys, orangutans, and humans. Five crucial mutations were delineated in the coding region of the human and rhesus and three in the orangutan, any one of which could be responsible for inactivation of the  $\alpha$ 1,3GT gene. Two of the mutations were shared by all three higher primates. These findings, which elucidate the molecular basis for the evolutionary loss of  $\alpha$ Gal expression, may have implications in medical research.

Most mammals express the cell surface carbohydrate epitope galactose- $\alpha$ 1,3-galactose ( $\alpha$ Gal),<sup>1</sup> with notable exceptions that include Old World monkeys, apes, and humans (1). With the loss of the  $\alpha$ Gal epitopes, the synthesis of which is dependent

on enzyme product of the  $\alpha$ 1,3-galactosyltransferase ( $\alpha$ 1,3GT) gene, higher primates produce anti- $\alpha$ Gal antibodies (2) that are responsible for the hyperacute rejection of organs transplanted from  $\alpha$ Gal-positive donors (3).

In 1989, Joziassse *et al.* (4) reported the sequence of full-length cDNA clone of the bovine  $\alpha$ 1,3GT gene and demonstrated the presence of this gene in the DNA of human cell lines. Using an 804-bp fragment derived from bovine full-length cDNA as a probe, they also detected mRNA transcripts in bovine and marmoset (New World monkey) but not in human or African green monkey cell lines (4). The absence of  $\alpha$ 1,3GT mRNA has since been widely viewed as a feature of all Old World monkeys, apes, and humans (5, 6). The molecular basis for the inactivation of the  $\alpha$ 1,3GT gene in the  $\alpha$ Gal-negative species has been attributed to a mutation(s) localized to a partial sequence of exon 9 (4, 5, 7).

Stimulated by the current interest in producing transgenic pigs for clinical use as tissue and organ xenograft donors, the cDNA for the  $\alpha$ 1,3GT gene of the  $\alpha$ Gal-positive pig was isolated (8), and the coding regions were characterized (9–12). Further, the full genomic organization of the porcine  $\alpha$ 1,3GT gene including its upstream regulatory region, was completed (12). A CpG island (*i.e.* a C connected by a 3'-5' phosphodiester bond to a G) characteristic of housekeeping genes was observed around exon 1 of the pig  $\alpha$ 1,3GT gene (12). It is of interest that a CpG island was not present in the mouse gene (13–15), suggesting that the upstream regulatory regions might be species-specific even though the enzyme function of the  $\alpha$ 1,3GT gene product was identical (15). If so, it would be conceivable that the loss of  $\alpha$ Gal expression in higher primates could be due to mutation in the upstream regulatory region (16).

As a further step in a phylogenetic comparison of the  $\alpha$ Gal-positive pig with the higher primates, we employed a PCR-based cloning strategy to avoid the potential artifacts of extrapolating from partial sequences and cross-species hybridization. After obtaining the full coding region of the  $\alpha$ Gal-positive marmoset and cebus  $\alpha$ 1,3GT gene, we identified the full coding region of the gene in the  $\alpha$ Gal-negative rhesus monkey, orangutan, and human, as well as the genomic organization and upstream regulatory region in the rhesus and human. In all three  $\alpha$ Gal-negative primate species, we found that the  $\alpha$ 1,3GT gene produces a variety of transcripts that were confirmed in the rhesus samples with Northern blot analysis. The transcripts contained crucial mutations in addition to the one(s) described previously in exon 9 of the rhesus (6), orangutan (6), and human (5, 6). Taken together, our findings suggest substantial revisions of the characterization of the  $\alpha$ 1,3GT gene

\* This work was supported by National Institutes of Health Grants DK 29961, R01 AI/DK 38899, AI 40329, DK 49615, and R01 DK 54232 and by Juvenile Diabetes Foundation Grant 4-1999-807. 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™/EBI Data Bank with accession number(s) AF384428 (marmoset  $\alpha$ 1,3GT), AY034181 (cebus  $\alpha$ 1,3GT), AY026225–AY026237 (rhesus  $\alpha$ 1,3GT), AF456457 (orangutan  $\alpha$ 1,3GT), and AF378121–AF378123 (human  $\alpha$ 1,3GT).

¶ To whom correspondence should be addressed: Thomas E. Starzl Transplantation Institute, University of Pittsburgh, 3601 Fifth Ave., Falk Clinic, 4th Fl., Pittsburgh, PA 15213. Tel.: 412-624-0112; Fax: 412-624-0192; E-mail: koikec@msx.upmc.edu.

<sup>1</sup> The abbreviations used are:  $\alpha$ Gal, galactose- $\alpha$ 1,3-galactose;  $\alpha$ 1,3GT,  $\alpha$ 1,3-galactosyltransferase; RACE, rapid amplification of cDNA ends; PBMC, peripheral blood mononuclear cells; RT-PCR, reverse transcriptase PCR.

TABLE I

Sequences of oligonucleotides used as PCR primers

GenBank™ accession numbers are: human, J05421; orangutan, M73305; rhesus, M73306.

Human	
H9a:	5'-CTGAATGGGTGTTCTCCAGAAATGGCT-3'
H9b:	5'-TGGAAAGACCTGGTCCACATCCATGCAG-3'
H9m:	5'-GAAGAGGTGGCAAGACATCAGCATGATGC-3'
H9n:	5'-CTGCATGGATGTGGACCAGGTCTTCCA-3'
Orangutan	
U9a:	5'-TAGGCCAGTCGGTGGCTCAGCTACAGG-3'
U9b:	5'-TGGCGGTACAAGGCAGATCCCGATGAC-3'
U9m:	5'-GTGATGTTGAGAACCCTGAATGGGTGTTG-3'
U9n:	5'-CTGGCCAAATGGAATATATGCTGCTGACTC-3'
Rhesus	
R9a:	5'-GAGGTCAAGCCAGAGAAGAGGTGGCAACA-3'
R9b:	5'-GACTTCCTCTTCTGCATGGATGTAGACCAG-3'
R9m:	5'-ATGTTGAGAACCCTGAATGGGTGTTCTCC-3'
R9n:	5'-CTGGCCAAATGGAATGCATGCTGCTGACTC-3'

that have implications for the molecular study of primate evolution and for the planning of xenotransplantation initiatives.

## EXPERIMENTAL PROCEDURES

**Genomic Organization of Human, Orangutan, and Rhesus Monkey  $\alpha 1,3$ GT Genes**—After high molecular genomic DNA was isolated from the PBMC of multiple healthy adult male humans, two rhesus monkeys, and an orangutan, libraries for the three species were constructed using the Universal GenomeWalker™ Library Kit (CLONTECH, Palo Alto, CA). The initial gene-specific and nested primer pairs were designed from the partial exon 9 sequences of the  $\alpha 1,3$ GT gene available in the GenBank™ (Table I). Gene-specific and nested primer pairs for other regions were designed from the exonic sequences obtained from RACE or RT-PCR experiments (not shown).

**5'- or 3'-RACE Analyses**—Using Trizol™ reagent (Invitrogen) according to the manufacturer's instructions, total RNA was extracted from rhesus PBMC and spleen samples, from human PBMC and placenta (provided in the CLONTECH Marathon™ amplification kit), and from orangutan PBMC. After treatment with DNase I (Ambion Inc., Austin, TX), poly(A)+ RNA was separated from total RNA using the Dynabeads™ mRNA purification Kit (Dyna, Oslo, Norway) in accordance with the protocol provided. Using the primers listed in Table I, the 5'- and 3'-RACE procedures were performed with the Marathon™ cDNA amplification kit. The thermal cycling conditions on a PerkinElmer GeneAmp System 9600 or 9700 thermocycler and TaKaRa LA Taq (Takara Shuzo, Shiga, Japan) enzyme were used for PCR procedures.

**RT-PCR**—Using total RNA derived from two rhesus monkeys, first-strand cDNA templates for RT-PCR were generated with the SuperScript Preamplification™ system (Invitrogen) according to the manufacturer's protocol. Based on the sequences obtained from the RACE experiments, primers were designed and RT-PCR was performed. Comparable experiments were performed with the orangutan and human-derived specimens. The primers are listed in Table I.

**Subcloning and Sequencing of Amplified Products**—PCR products amplified in 5'- or 3'-RACE, GenomeWalker™-PCR, and RT-PCR were subcloned into the pCR II vector provided with the Original TA Cloning™ Kit (Invitrogen, Carlsbad, CA). Automated fluorescent sequencing of cloned inserts was performed using an ABI 377 Automated DNA Sequence Analyzer (Applied Biosystems, Inc., Foster City, CA).

**Northern Blot Analysis**—To confirm the results of RT-PCR in the rhesus species with NorthernMax™-GlyNorthern™ blots system (Ambion), total RNA was extracted from rhesus spleen tissue using Trizol™ as described above. Using the protocol provided with the kit, ~10  $\mu$ g of total RNA from each species was blotted to the BrightStar-Plus™ membrane. The BrightStar™ Psoralen-biotin nonisotopic labeling kit was employed for the biotinylated DNA probe (a 744-bp fragment of the rhesus  $\alpha 1,3$ GT exon 9).

**Luciferase Assay for Promoter Activity**—The Dual Luciferase™ reporter assay (Promega, Madison, WI) was used to evaluate the promoter activity. For this purpose, a fragment of 1431 bp of the 5'-flanking region of rhesus exon 1 was inserted into the multicloning site of the pGL3-Basic vector (termed pGL3-RA), which was co-transfected with pRL-TK vector for normalization of transfection efficiency into pig aortic endothelial cells using LipofectAMINE (Invitrogen). Relative luciferase activity of the pGL3-RA was measured by luminometer and compared with that of the pig promoter (12).

## RESULTS

**Genomic Organization of the Rhesus and Human  $\alpha 1,3$ GT Gene**—The full-length genomic organization determined by the combined results of GenomeWalker-PCR, RACE, and RT-PCR consisted of nine exons and eight introns (Fig. 1), corresponding to the previously delineated porcine and mouse genomic structure (8–14). However, some cryptic exons or aberrant splicing sites were observed as described below. The total rhesus and human gene size is calculated to be more than 50 kb.

**5'-RACE, 3'-RACE, and Northern Blot Analysis of the  $\alpha 1,3$ GT Gene**—5'- and 3'-RACE experiments revealed sequences of the  $\alpha$ Gal-negative species that were highly homologous with the  $\alpha 1,3$ GT gene of  $\alpha$ Gal-positive species. 3'-RACE demonstrated the presence of poly(A) (adenylation) signal and poly(A) tail. Based on these sequences, we performed RT-PCR using a pair of primers that spanned several exons. Multiple bands were seen in human (Fig. 2A, a) and rhesus (Fig. 2A, c), whereas only one major band was seen in the orangutan (Fig. 2A, b). The results of any exon of a given species were identical with all three PCR methods (*i.e.* RACE, RT-PCR, and GW-PCR) (Fig. 2A). Moreover, the presence of transcripts for the rhesus  $\alpha 1,3$ GT gene was confirmed with Northern blot hybridization (Fig. 2A, d). We concluded that the sequences identified with all of the detection methods were derived from authentic  $\alpha 1,3$ GT mRNA in the respective  $\alpha$ Gal-negative species.

**Cryptic Exons**—Cryptic exons were found in both the rhesus and human  $\alpha 1,3$ GT gene, *e.g.* between exons 2 and 4, designated 3H in human and 3R in rhesus (Fig. 2B, a and c). Exons 3H and 3R showed no homology with each other, and neither exon had any homology with the previously described mouse exon 3 (13). A cryptic human exon with two size variants, both of which had poly(A) tails, was found in 18 of 22 humans between exons 7 and 8 (Fig. 2B, a) and was designated 7H. At the end of the 7Ha variant, a donor splicing site was found that spliced to exon 9 (Fig. 2B, a). In rhesus, cryptic exons designated 6Ra and 6Rb were found between exons 6 and 7, and in addition, a third variant was found (6Rc) that included a retained intron (Fig. 2B, c).

**Aberrant Splicing Sites**—In the 5'-untranslated region, rhesus exon 2 is 112 bp, whereas human exon 2 is 269 bp (Fig. 3). Comparative analysis revealed that the first half (110 bp, termed 2A) of the human exon 2 is highly homologous to rhesus exon 2, and the last half (159 bp, termed 2B) is similar to the rhesus intron 2 that extends from exon 2. Although the first half (138 bp, termed 8Ra) of the rhesus exon 8 (327 bp) showed high homology with the porcine exon 8 (138 bp), the remaining half (189 bp, termed 8Rb) resembled the porcine intron 8 sequences extending from exon 8 (Figs. 2B, c, and 3). A single nucleotide deletion of G was found only in rhesus at the border between the region 8Ra and 8Rb (Fig. 3, position [d]), resulting in utilization of another downstream splicing site.

**Alternative mRNA Splicing of the Human and Rhesus  $\alpha 1,3$ GT Gene**—A variety of alternative splicing patterns were demonstrated in the RACE and RT-PCR experiments (Fig. 2, A and B). In the coding region, the predominant pattern in human consisted of exons 4 through 9. Because of the cryptic exons and aberrant splicing sites in the human and rhesus genes, the alternative splicing patterns were far more varied in these  $\alpha$ Gal-negative species than in  $\alpha$ Gal-positive murine or porcine species (8–13).

**Exon-Intron Splice Junctions of Human, Orangutan, and Rhesus  $\alpha 1,3$ GT Gene**—A comparative sequence analysis of splice junctions was performed on rhesus, orangutan, and human genomic DNA (Fig. 3). The results were further compared with the porcine species sequence (GenBank™ accession numbers AF221510–AF221517) (not shown), revealing high homol-

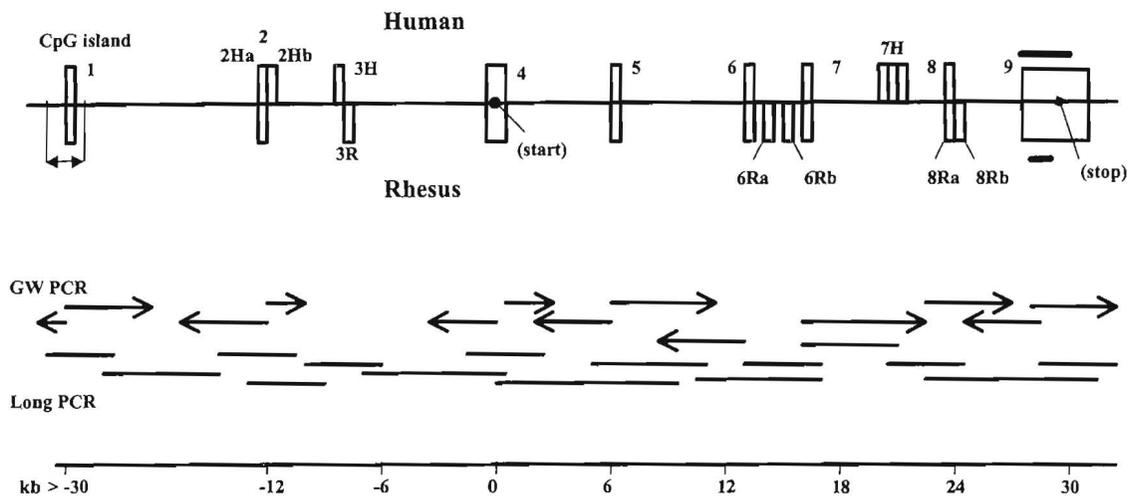


FIG. 1. Genomic organization on a kb scale of the  $\alpha 1,3$ GT gene of the human (above horizontal line) and rhesus (below horizontal line). The numbered boxes indicate the approximate location of exons within the locus of the two species but do not accurately reflect the length of each exon. The translation start and stop codons are located in exons 4 and 9. The horizontal thick bar above exon 9 denotes the GenBank™ data for human (J05421, 794 bp) contributed by Larsen *et al.* (5), and the short thick bar below exon 9 denotes the data for rhesus (M73306, 371 bp) contributed by Galili and Swanson (6). Shown below the gene locus representation is a contig of overlapping subgenomic clones derived from GenomeWalker™ libraries (horizontal arrows) or from long PCR experiments (horizontal lines).

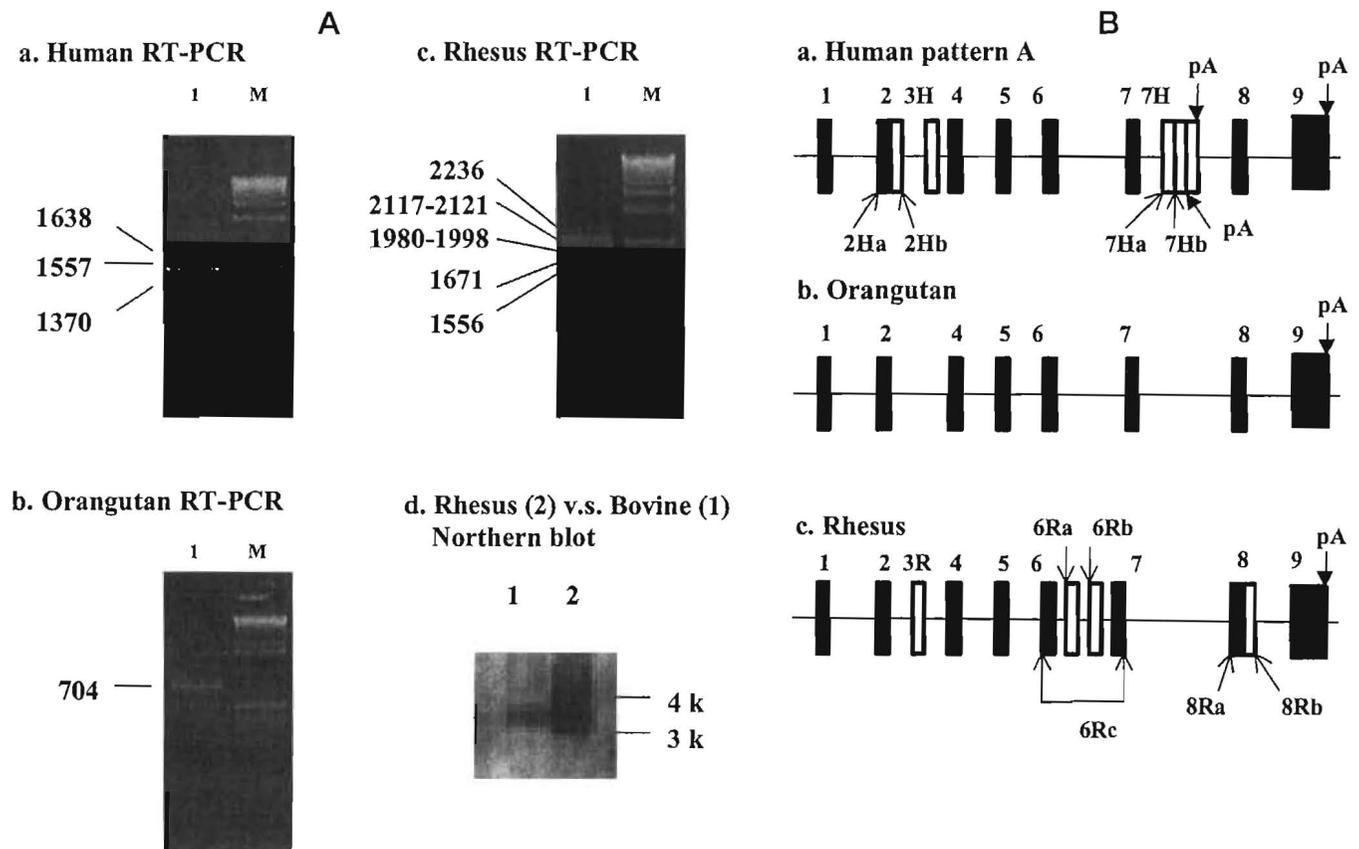


FIG. 2. A, electrophoresis of PCR (a-c) and Northern blot hybridization (d). Lane 1, electrophoresis results of RT-PCR in the human, orangutan, and rhesus cDNA libraries, respectively, using a pair of primers spanning the coding region. Lane M, DNA size marker ladder (Invitrogen, 1-kb DNA size ladder). d, Northern blot of total RNA from bovine (lane 1) or rhesus spleen tissue (lane 2) hybridized with a probe derived from rhesus exon 9 fragment. As expected, with in-species versus cross-species hybridization, gene expression was lower in the bovine than in the monkey spleen cells. B, schematic splicing variations and cryptic exons of human (a), orangutan (b), and rhesus (c)  $\alpha 1,3$ GT gene. The closed boxes denote exons in the respective species. The open boxes with arrows denote cryptic exons that were not observed in the  $\alpha$ Gal-positive species. Exon 6Rc in rhesus has exon 6, intron 6, and exon 7 (a retained intron).

ogy among the three species. As shown in Fig. 3, the pyrimidine richness in the intron 7–exon 8 boundary of both the rhesus and human was less than the pyrimidine content of the intron 3–exon 4 boundary, explaining the frequency of splicing of exon 8 in the 2 species.

*Regulatory Region and 5'-Untranslated Region of the Rhesus  $\alpha 1,3$ GT Gene*—GenomeWalker-PCR starting at exon 1 revealed that the rhesus and human upstream regulatory region contains several SP1/GC-box, AP2, and other consensus motifs in the absence of a TATA- or CAAT-box. The findings were

exon 1	Hu	---GGGCGCGGTCTGCTCAGCCAGCTCACCGCGCGCCGCGCCCTCGGCGCCCTGGTTCTGCGGATCAG	gtaggt---
	Rh	---CGCCGCCGGGGGCGGGAGCCGAGGTGTGGGCCATCCCCGAGCGCACCCAGCTTCTGCCGATCAG	gtaggt---
exon 2	Hu (269)	---gtctactccccatctcctgaaacagctgtttattctttcgacag	GAGTTGAAACC--GGCAGGAG GTATGT
	[2A:110]		---GGCTCAGGGACTTCACT gtaagt---
	Ou (115)	---gtctactccccatctcctgaaacagctgtttattctttcgacag	GATTTGAAACC--ACGATGAG gtagct---
	Rh (112)	---gtctactccctatctgctgaggcagctgtttattctttcgacag	GAGTTGAAACC--ACGATAAG gtagt---
exon 4	Hu (89)	---atttaactctgttctctttttttctccctcttctttttcttag	GAGAAAATAATG-- AAC AG gtaatt---
	Ou (89)	---gatttaactctgttctctttttttctccctcttctttttcttag	GAGAAAATAATG-- AAC AG gtaatt---
	Rh (89)	---atttaactctgttctctttttttctccctcttctttttcttag	GAGAAAATAATG-- AAC AG gtaatt---
exon 5	Hu (36)	---attaaaagtaacatgatatgtttttattctttttgcttttag	C ACA GAA--CAC TCA AA gtaagt---
	Ou (36)	---attaaaagtaacatgatatgtttttattctttttgcttttag	C CCA GAA--CAC TCA AA gtaagt---
	Rh (36)	---attaaaagtaacatgatatgtttttgctttctctttcacttttag	C CCA GAA--CGC TCA AA gtaagt---
exon 6	Hu (66)	---caccacattttctaaaaatatttgtaactttttcattcttag	A AAC CCA--AAC AAT GG gtaagg---
	Ou (66)	---acaccacattttctaaaaatatttgtaactttttcattcttag	A AAC CCA--AAC AAT GG gtaagg---
	Rh (66)	---caccacattttctaaaaacatttgtaactttttcattcttag	A AAC CCA--AAC AAT GG gtaagg---
exon 7	Hu (116)	---tcaccttccctgccttacctggtttactttccctttgacttag	G ATC CAC---ATC CAA A gtaaga---
	Ou (116)	---tcaccttccctgccttacctggtttactttccctttgacttag	G ATC CAC---ATC CAA A gtaaga---
	Rh (115)	---tcaccttccctgccttacctggtttactttccctttgacttag	G ATC CAC---ATC CAA A gtaaga---
exon 7H	Hu (803)	---ttaattcaacataaatttttaagggtgctgtttttcttccag	C TTG AAG--GGATGACCTG gtagtag---
	[7Ha:189]		
exon 8	Hu (138)	---ggcaaagctgccctgactcagggaaactctacgttttatcaccag	g AAA CGC--GTC GGA AG gtaggt---
	Ou (138)	---ggcaaagctgccctgactcagggaaactctacgttttatcaccag	g AAA CGC--ATT GGA AG gtaggt---
	Rh (327)	---tggcaaagctgccctgactcagggaaactctacgttttatcaccag	G AAA CAC--ATT GGA AG -TGGGT
	[8A:138]		---GAACAAAATAAGAGTAG gtaggt---
			[d] Deletion of G
exon 9	Hu (692)	---ggaatcccagaggataaaatgttttgcttttcttctttgctttag	A TAT AAT-AAT GTC TGA-- + poly A signal
	Ou (693)	---ggaatcccagaggataaaatgttttgcttttcttctttgctttag	A TAT AAT-AAC GTC TGA-- + poly A signal
	Rh (694)	---ggaatcccagaggataaaatgttttgcttttcttctttgctttag	A TAT ATT-ATC ATC TGA-- + poly A signal

FIG. 3. Comparison of exon-intron boundaries of the human (H), orangutan (Ou), and rhesus (R)  $\alpha 1,3$ GT gene. Capital or lowercase letters indicate the nucleotides of the exon or intron in each species, respectively. The numbers in parentheses indicate exon size in bp. The start codon in exon 4 and the authentic stop codon in exon 9 are underlined.

similar to those in the pig gene (12). In both the rhesus and human, as well as the pig, the region comprises a CpG island. In the upstream region between -340 and -470, the high degree of homology between the rhesus and pig, but not between the human and pig, gene is noteworthy (Fig. 4). The transcription initiation site has been identified in the pig (12). When the transcription initiation site is identified in the human and rhesus gene, the question may be clarified as to whether there are differences in gene regulation between the human and rhesus.

**Rhesus  $\alpha 1,3$ GT Gene Promoter Activity**—The intensity of expression promoted expressed by the pGL3-RA construct was 18-fold greater than the pGL3-Basic control, and that of pig was 81-fold higher (Fig. 5).

**3'-Untranslated Region of the Rhesus  $\alpha 1,3$ GT Gene**—The nucleotide divergence of the rhesus exon 9 sequence from that of the orangutan and human was found starting 1273 bp downstream from the 5'-end of exon 9 and spanning 420 bp (data not shown). A GenBank™ BLAST search revealed that this divergence had a striking homology with a retrovirus long terminal repeat sequence present in the baboon (GenBank™ accession number NC-000863; Ref. 17).

**Homology in the Coding Region**—We compared our results of the human, orangutan, and rhesus sequences to our  $\alpha$ Gal-positive marmoset and cebus New World monkey sequences (deposited in the GenBank™). Our marmoset sequence also was compared with an independently submitted sequence (18). Although the overall sequence from the start codon in exon 4 to the stop codon in exon 9 in marmoset is 1131 bp, the corresponding sequence of the human (from exon 4 to exon 9 through

exon 8) or rhesus (from exon 4 to exon 9 through exon 8Ra) was 1128 or 1129 bp, respectively. Homology between human, orangutan, or rhesus to marmoset is 94.3, 94.7, or 92.4%, respectively.

**Comparison of Predicted Amino Acid Sequences**—There were no frameshift mutations in the rhesus, orangutan, or human exons 4–6. However, five crucial mutations that might account for loss of the  $\alpha 1,3$ GT expression were identified in the human and rhesus, and three were detected in the orangutan. It is noteworthy that the mutations at the position [b] and [f] were shared by all three species.

In the rhesus, at position [a] and [b] (Fig. 7), corresponding to the deduced amino acid positions 75 and 81, respectively, in marmoset, a single nucleotide deletion resulted in a frameshift. At position 96 (position [c]), a single nucleotide substitution resulted in a premature stop codon. As described under "Aberant Splicing Sites," a single nucleotide deletion was detected at position [d] of the exon 8 splicing junction. At position 268 [f], a single nucleotide substitution also resulted in a premature stop codon.

In the orangutan, a single nucleotide deletion was observed at position [b] and [g]. A premature stop codon was observed at position [f].

In human, a single nucleotide deletion was observed at the position [b], [e] and [g]. A premature stop codon was observed at position [f] and [h].

Three critical mutations were observed within exon 7, leading to the predicted nonfunctional  $\alpha 1,3$ GT expression in all frames (Figs. 6 and 7). At position [a] or [b] in the rhesus, corresponding to the deduced amino acid at positions 75 or

	ZID	PADS	NMYC	VDR	AREB6			
R (-497):	TTCCAGCTCC	ATCACTGACA	CCACCACCAC	CACCCCGTGC	AGGCCAAGA	AGGCCAAAGA	GGCCAGTCA	CTCTCACC
	*		** ***** *	*****	*****	*****	*****	** * ***** *
P (-492):	T-----	-----C	CCTCCACCCC	CACCACGTGC	AGGCCAGAGA	AGGC-AAAGA	GGCCAGCCA	CCCTCACC
	NFKB	IRF1	AREB6	AREB6				
R (-417):	GGACTCTACT	TTCTTTTTTT	GCTGGTTTCA	GGCTTTTTTC	TGCCTGAGTG	AAAATGAAAC	AAACACACCC	-GCGCC-CCC
	*** * * *	*****	*****	*****	*****	*****	*****	*** * ***** *
P (-431):	GGAATTTCTT	TTCTTTTTTT	GCTGGTTTCA	GGCTTTTTTC	TGCCTGAGTG	AAAATGAAAC	AAACACCCCC	TGCGCCTCCC
	WT1	PAX8	USF	SP1				
R (-339):	GGC-----A	CACACA----	CACACTCATT	CATGCATTCA	CG-CGGCGGC	CTCCAGCCG	TCTGTGTCCA	GGCGGGGAC
	***	* *****	*** ** *	* ** ** *	** * ****	*** ** **	* *****	*** *****
P (-351):	GGCCACCAGA	CACACACGCG	CACCGGCACT	CGCGCACTCG	CGCCCTCGGC	CTCCTAG-CG	GCCGTGTCTG	GGGC-GGGAC
	NF1	SP1						
R (-270):	TGGCTTGCA	CCGACCCTG	CCGCCGGGC	ACAG-----	-TGGCCGCC	CGCTG-TCAG	CGCTCGCCG	-TCCCTCGC
	*** ** *	** ** *	* *****	**	*****	*** * **	** *****	* *****
P (-273):	CCGCTCTGCA	CAAACAGCCG	CGGCCGGGT	GGAGCGGGGA	GCTCGCCGC	CGCCGCCAG	TGCCCGCCG	CTTCTCGC
	NRSF	MTF-1	AP2					
R (-199):	CCCCTGCCCG	CTGCGCTATG	CCAGAGCGG	ACGCCAGCC	GGCCGGCCCA	CGGCAGGAG	TACGGGC-GC	ACGCTCTGGC
	*****	* * *	*** ** *	* * ** *	*****	*****	* ** ** *	** * ** *
P (-193):	CCCCTGCCCG	C--CAC----	CCCGAG-GA	GCACACAG-C	GGCCGG----	CGGCAGGAG	CGCAGGCGGC	ACAC-CCCGC
	MUSCLE	NRSF						
R (-120):	GGAGGCAGCG	AGGTAGCTGA	TCCC--AGC	GCGGCGGGCC	C CGCGGTGGC	CGGGAGCTAG	G-CGCGGTC-	-CTGCTCAGC
	**** **	* ** ** *	* ** *	** ** ** *	** ** ** *	* ** * ** *	* ** * ** *	* ** *
P (-126):	CCCGCA-CG	CCCT-GCCGA	GCTCAGGAGC	ACGCCGCGCG	CCACTGTTCC	CTCAGCCGAG	GACGCCCGC	GGGGCCGGG
								-1
R (-46):	CCGCTCGCTG	CGCGCCGGTC	CTGGGTG--C	CAGGGTTCTG	CGGATCAGG	gtgggtcccgcagggcgccaggtcccc-ggag		
	* * **	* ** ** *	* * ** *	* * ** ** *	* * ** ** *	*****	*****	*** * ** *
P (-48):	AGCCGAGGTG	TGGGCCATCC	CCGAGCGCAC	CCAGCTTCTG	CCGATCAGG	gtgggtcccgcagggcgccaggtcccc-ggag		

FIG. 4. Exon 1 vicinity of  $\alpha 1,3$ GT gene of rhesus (*R*) compared with pig (*P*). The rhesus exon 1 and its potential regulatory region (*capital letters*) are aligned with the corresponding porcine exon 1 sequence (12). The *lowercase letters* denote intron 1. An *asterisk* indicates that the nucleotide is identical between both species.

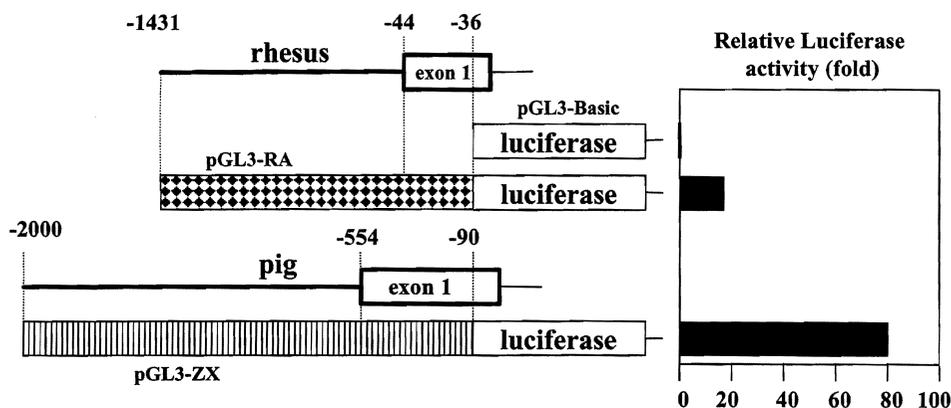


FIG. 5. Luciferase assay for promoter activity of the rhesus  $\alpha 1,3$ GT gene.

81, respectively, in marmoset, a single nucleotide deletion, G, from the triplet GAA (Glu) was identified, resulting in a frameshift. At position 96 (position [c]), a single nucleotide substitution resulted in a premature stop codon (from TGG (Trp) to TGA). At position 268 ([f]), a single nucleotide substitution resulted in an additional premature stop codon (TAT (Tyr) to TAG).

A single nucleotide deletion at position [g] was observed in both human and orangutan but not in rhesus. The premature stop codon at position [h] was observed only in human. Thus, five crucial mutations that might account for loss of  $\alpha 1,3$ GT expression were determined in human and rhesus and three

in orangutan. It is noteworthy that the mutations at position [b] and [f] were shared by the rhesus monkey, ape, and human.

#### DISCUSSION

The  $\alpha$ Gal epitopes are chemically related to the ABO blood type antigens. In the absence of expression of A or B epitopes or of  $\alpha$ Gal epitopes, antibodies develop against the respective antigens. Phylogenetic data on the emergence of the ABO antigen system (19) and especially molecular studies on ABO transferase (20) have been valuable in determining higher primate evolution. In contrast, little comparable information is

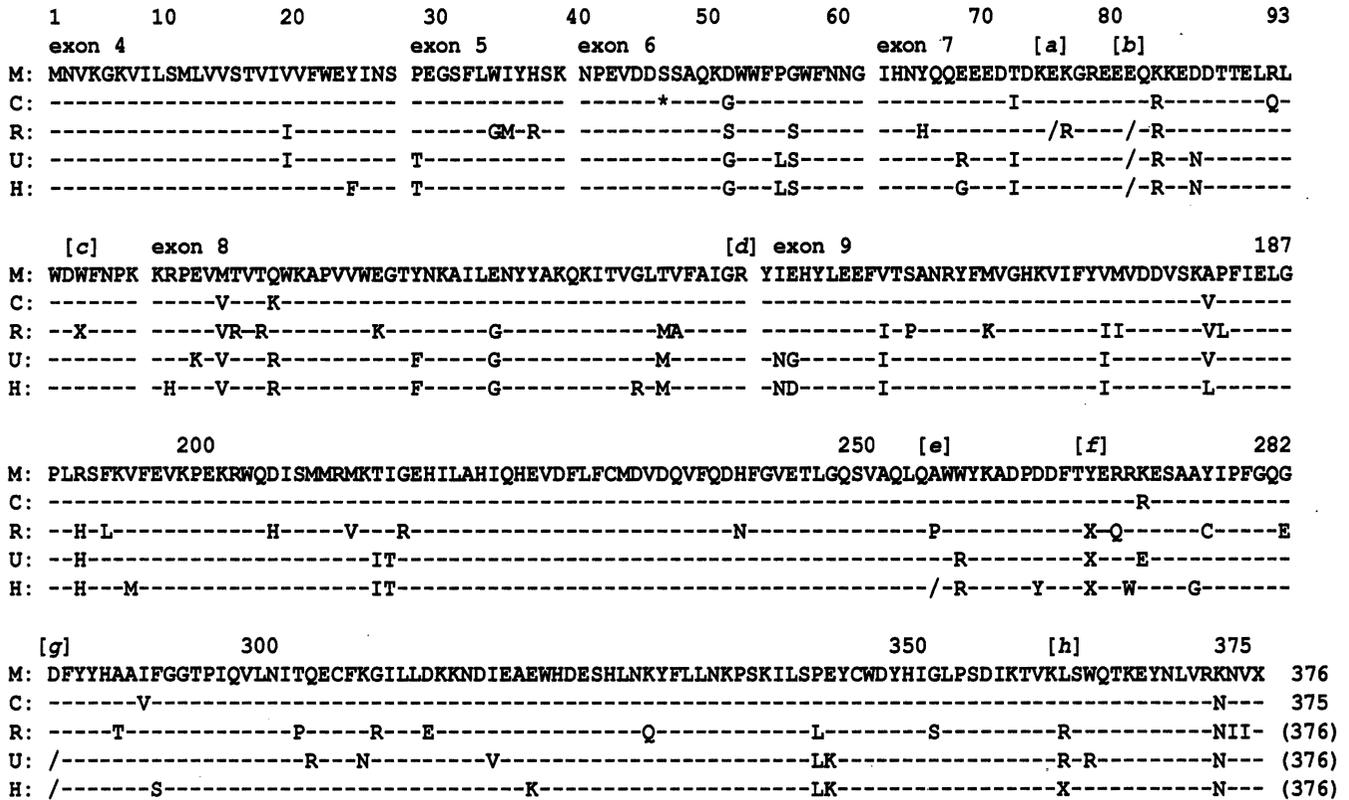


FIG. 6. Comparisons of deduced amino acid sequences of the  $\alpha 1,3$ GT gene in marmoset (M), cebus (C), rhesus (R), orangutan (O), and human (H). Nucleotide sequences are shown from the start codon to the stop codon of the marmoset monkey cDNA in one-letter abbreviations. Dashed horizontal lines indicate that each nucleotide is identical to the corresponding nucleotide of the marmoset base-line sequence. The single asterisk in cebus denotes a triplet nucleotide deletion not present in the marmoset. A single nucleotide deletion, depicted by a slash (/), was observed at positions [a], [b], [e], and [g]. Premature termination, depicted by an X, was observed at positions [c], [f], and [h].

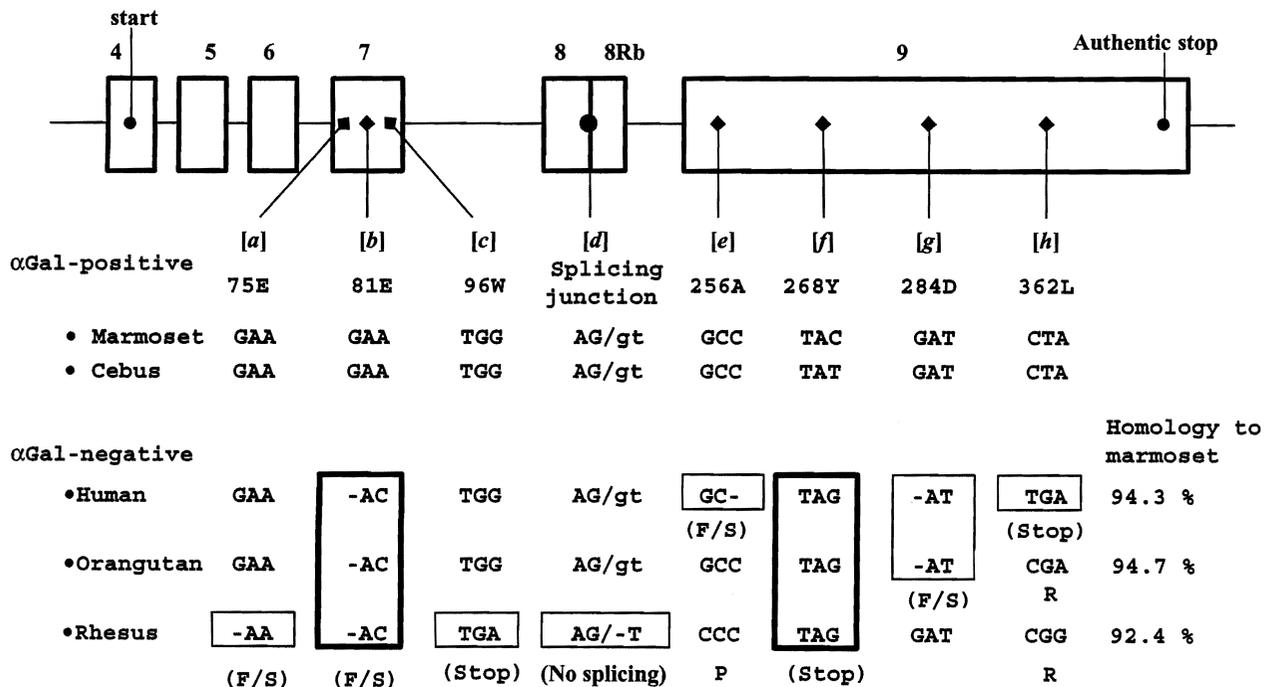


FIG. 7. Comparison of crucial point mutations of the  $\alpha 1,3$ GT gene in primates. The numbered open boxes (top) are exons in which crucial mutations in the primates are shown with the corresponding nucleotide sequence. No mutations were observed in  $\alpha$ Gal-positive primates. Crucial mutations were observed in exons 7 and 9 of the three higher primate species examined (lower boxes). Note that the mutations at position [b] and [f] (lower closed boxes with thick lines) are shared by rhesus, orangutan, and human. (F/S), frameshift; (Stop), stop codon.

known about the molecular basis of inactivation of the gene expressing the enzyme  $\alpha 1,3$ GT upon which synthesis of the  $\alpha$ Gal epitope depends.

The  $\alpha$ Gal system in lower mammals was not discovered until the 1980s (1, 2). Evidence of the presence of  $\alpha 1,3$ GT gene homologues in some  $\alpha$ Gal-negative higher primates was pub-

lished in 1989–1991 by three pioneer groups. Joziassse *et al.* (7) described an 89-bp sequence (HGT-10) in human cells that was similar to a fragment of the murine and bovine exon 4. Larsen *et al.* (5) characterized a 703-bp genomic sequence in human tissue with a high degree of homology to exon 9 of the mouse  $\alpha$ 1,3GT gene. Galili and Swanson (6) reported a 371-bp sequence in rhesus cells that resembled part of mouse and bovine exon 9.

All three groups concluded that point mutations in exon 9 at some time in the evolutionary past of the higher primates could be the cause of the loss of  $\alpha$ Gal epitope expression. We confirmed the point mutations that had been described earlier in exon 9 of the rhesus (6), orangutan (6), and human  $\alpha$ 1,3GT gene (5, 6), and demonstrated four additional point mutations in exons 7 and 8. The crucial mutations in the different species that result in frameshift and premature termination of translation include a deletion mutation at position [b] in exon 7 and a premature termination at position [f] in exon 9 that are shared by all three of these higher primates.

Because no  $\alpha$ 1,3GT transcripts could be found in the early investigations of the African green monkey (COS) and human cell lines (4), the possibility could not be dismissed that inactivation of the gene in the higher primates is due to mutations in the upstream regulatory regions of the  $\alpha$ 1,3GT gene (16). This potential explanation for loss of the  $\alpha$ Gal epitope was eliminated in our study using sensitive PCR methods for detection of mRNA transcripts that were not available at the time of the seminal investigations (4–7). A variety of transcripts with more than 10 different splicing patterns of the  $\alpha$ 1,3GT gene were observed in the full coding region in all rhesus, orangutan, and human tissues examined. Their presence was confirmed with Northern blot of total RNA derived from rhesus spleen.

Joziassse *et al.* (7, 16) estimated that inactivation of the  $\alpha$ 1,3GT gene occurred between 40 and 25 millions years ago, following the divergence between higher primates (catarrhines) and the New World monkey but prior to the further divergence of the individual higher primates. The hypothesis was based on their discovery of an intron-less processed cDNA homologue (termed HGT-2) in the human genome (7, 16). In contrast, Galili and Swanson (6) postulated that inactivation occurred at a much later time and independently in Old World monkeys, apes, and humans (6). Our finding of crucial point mutations shared by the rhesus, orangutan, and human at positions [b] and [f] of exons 7 and 9, respectively, appears to be more consistent with Joziassse's hypothesis (16) that the  $\alpha$ 1,3GT gene inactivation occurred in a common ancestor than with the suggestion by Galili and Swanson (6) that it occurred independently in the three different lineages. However, the time and the process of the inactivation remain to be determined (6, 7). The high homology of the  $\alpha$ 1,3GT gene in the  $\alpha$ Gal-negative species to the shared sequence of the  $\alpha$ Gal-positive marmoset and cebus (Fig. 7) suggests that the inactivation may have occurred more recently than previously thought.

The development of preformed antibodies directed against a wide range of microorganisms or macroparasites that express

$\alpha$ Gal-like substance has been the conventional explanation for the evolution of the  $\alpha$ Gal-negative state in higher primates (1, 2). The possibility cannot yet be excluded, however, that some other survival advantage drove the gene inactivation and that the protection from such infecting agents afforded by the "natural" anti- $\alpha$ Gal antibodies was fortuitous. Whatever the reason, antibodies against the  $\alpha$ Gal epitope are responsible for the immediate (hyperacute) rejection by  $\alpha$ Gal-negative recipients of tissues and organs from  $\alpha$ Gal-positive donors, precluding successful clinical xenotransplantation from  $\alpha$ Gal-positive animals (3). Detailed information about the  $\alpha$ 1,3GT gene may help in mapping strategies for transgenic modification of the  $\alpha$ Gal-positive species.

If higher primates were to spontaneously recover the expression of  $\alpha$ Gal epitopes, as has been described in normal breast and MCF7 human breast carcinoma cell lines (21), the potential for the development of  $\alpha$ Gal antibody-mediated autoimmune disease is implicit. Galili and co-workers (22) observed the presence of  $\alpha$ Gal epitopes in thyroid cells of a human with autoimmune Grave's disease, but no transcripts could be detected in these cells with Northern blot. A search for  $\alpha$ 1,3GT mRNA transcripts in tissues from such "autoimmune suspect" patients may be fruitful with more sensitive PCR technology used in our study.

*Acknowledgments*—We thank Rickquel Tripp for technical assistance and Terry Mangan for manuscript preparation.

#### REFERENCES

- Galili, U., Shohet, S. B., Kobrin, E., Stults, C. L. M., and Macher, B. A. (1988) *J. Biol. Chem.* **263**, 17755–17762
- Galili, U., Clark, M. R., Shohet, S. B., Buehler, J., and Macher, B. A. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 1369–1373
- Cooper, D. K. C., Koren, E., and Oliol, R. (1993) *Lancet* **342**, 682–683
- Joziassse, D. H., Shaper, J. H., Van den Eijnden, D. H., Van Tunen, A. J., and Shaper, N. L. (1989) *J. Biol. Chem.* **264**, 14290–14297
- Larsen, R. D., Rivera-Marrero, C. A., Ernst, L. K., Cummings, R. D., and Lowe, J. B. (1990) *J. Biol. Chem.* **265**, 7055–7061
- Galili, U., and Swanson, K. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 7401–7404
- Joziassse, D. H., Shaper, J. H., Jabs, E. W., and Shaper, N. L. (1991) *J. Biol. Chem.* **266**, 6991–6998
- Sandrin, M. S., Dabkowski, P. L., Henning, M. M., Mouhtouris, E., and McKenzie, I. F. C. (1994) *Xenotransplant* **1**, 81–88
- Strahan, K. M., Gu, F., Andersson, L., and Gustafsson, K. (1995) *Immunogenetics* **41**, 101–105
- Vanhove, B., Goret, H., Soulillou, J. F., and Pourcel, C. (1997) *Biochim. Biophys. Acta* **1356**, 1–11
- Katayama, A., Ogawa, H., Kadomatsu, K., Kurosawa, N., Kobayashi, T., Kaneda, N., Uchimura, K., Yokoyama, I., Muramatsu, T., and Takagi, H. (1998) *J. Glycoconj.* **15**, 583–589
- Koike, C., Friday, R. P., Nakashima, I., Luppi, P., Fung, J. J., Rao, A. S., Starzl, T. E., and Trucco, M. (2000) *Transplantation* **70**, 1275–1283
- Joziassse, D. H., Shaper, N. L., Kim, D., Van den Eijnden, D. H., and Shaper, J. H. (1992) *J. Biol. Chem.* **267**, 5534–5541
- Larsen, R. B., Rajan, V. P., Ruff, M. M., Kukowska-Latallo, J., Cummings, R. D., and Lowe, J. B. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8227–8231
- Koike, C., Friday, R. P., Fung, J. J., Starzl, T. E., and Trucco, M. (2001) *Transplant. Proc.* **33**, 710–711
- Joziassse, D. H., and Oriol, R. (1999) *Biochem. Biophys. Acta* **1455**, 403–418
- Mang, R., Goudsmit, J., and van der Kuyf, A. C. (1999) *J. Virol.* **73**, 7021–7026
- Henion, T. R., Macher, B. A., Anaraki, F., and Galili, U. (1994) *Glycobiology* **4**, 193–201
- Oliol, R., Le Pendu, J., and Mollicone, R. (1986) *Vox Sang.* **51**, 161–171
- Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345**, 229–233
- Petryniak, J., Varani, J., Ervin, P. R., and Goldstein, I. J. (1991) *Cancer Lett.* **60**, 59–65
- Etienne-Decerf, J., Winand, R., and Galili, U. (1991) *Autoimmunity* **10**, 81–87