The acid lipase gene family: three enzymes, one highly conserved gene structure

Peter Lohse,[1] Pia Lohse, Soheyla Chahrokh-Zadeh, and Dietrich Seidel

Department of Clinical Chemistry, Grosshadern Clinic, University of Munich, D-81366 Munich, Germany

Abstract Human gastric lipase (HGL; triacylglycerol lipase; EC 3.1.1.3) plays an important role in the digestion of dietary triglycerides in the gastrointestinal tract, especially in patients suffering from pancreatic lipase deficiencies. The enzyme is secreted by the fundic mucosa of the stomach and hydrolyzes the ester bonds of triglycerides under acidic pH conditions, while cholesteryl esters are not attacked. The 379-amino acid protein is highly homologous to two other acidic lipases, rat lingual lipase (RLL; triacylglycerol lipase; EC 3.1.1.3) and human lysosomal acid lipase (HLAL; cholesteryl esterase; EC 3.1.1.13). To determine whether this remarkable similarity is also present at the genomic level, we have elucidated the respective gene structures by screening three bacteriophage lambda libraries and by polymerase chain reaction-based intron amplification. The genes encoding HGL, RLL, and HLAL are composed of 10 exons interrupted by nine introns and span about 14 kb, 18.7 kb, and 38.8 kb of genomic DNA, respectively. The HGL and RLL gene organizations are identical, suggesting that RLL is the rat gastric lipase expressed in the serous von Ebner glands of the tongue. The positions of the HLAL intervening sequences are also absolutely conserved, except for the location of intron 1. Our results support the concept that HLAL and HGL/RLL are members of a gene family of lipases that most likely have evolved by duplication of an ancestral gene and subsequently assumed distinct roles in neutral lipid metabolism due to sequence divergence and different expression patterns.—**Lohse, P., P. Lohse, S. Chahrokh-Zadeh, and D. Seidel.** The acid lipase gene family: three enzymes,one highly conserved gene structure. *J. Lipid Res.* 1997. **38:** 880–891.

Supplementary key words lipid metabolism • human lysosomal acid lipase • human gastric lipase • rat lingual lipase • structure-function relationships • lipase evolution

Human gastric lipase (HGL) is an acid-stable lipolytic enzyme involved in the digestion and absorption of dietary triglycerides in the stomach, duodenum, and jejunum (for review see refs. 1–3). The enzyme is synthesized primarily by the chief cells of the fundic mucosa (4, 5) and secreted into the gastric lumen where it initiates the hydrolysis of about 10 to 20% of the triglycerides ingested, with diacylglycerol and fatty acids being the main end products (4, 6). The limited lipolysis in the stomach appears to be critical for the subsequent hydrolysis by colipase-dependent pancreatic lipase and bile salt-stimulated lipase (1, 7, 8). This enzyme system is obviously geared to the complete absorption of dietary lipids and their rapid delivery to peripheral tissues and to the liver. The key metabolic role of HGL is accentuated in certain pathological states that lead to exocrine pancreatic insufficiency such as pancreatitis, cystic fibrosis, or alcohol abuse (9–12).

Gastric lipases have been purified from human (13), rabbit (14), and canine (15) stomach. In addition, a cDNA clone for HGL has been isolated and the protein's complete amino acid sequence has been predicted (16). HGL consists of 379 residues plus a signal peptide of 19 amino acids, with a molecular weight of 43.16 kDa. The deduced sequence is highly homologous to that of two other lipases with acidic pH optimum, rat lingual lipase (RLL; 377 amino acids; about 76% identity; 17) and human lysosomal acid lipase/cholesteryl ester hydrolase (HLAL; 378 amino acids; approx. 58% identity; 18).

RLL has the same physiological function as HGL, i.e., to initiate triglyceride digestion in the stomach. The enzyme is secreted by the serous glands of von Ebner in the back of the tongue (19) and hydrolyzes medium- and long-chain triglycerides in the gastrointestinal tract at a broad pH range of 2.2–6.0 (for review see refs. 2, 3, 20, 21). HLAL, in contrast, catalyzes the intralysosomal breakdown of cholesteryl esters and triglycerides at acid pH in virtually all cells except erythrocytes, thereby ensuring intracellular lipid homeostasis (for review see refs. 22 and 23).

As a prerequisite for structural studies and in order to obtain insights into the regulation of their expres-

Abbreviations: HGL, human gastric lipase; HLAL, human lysosomal acid lipase; PCR, polymerase chain reaction; RLL, rat lingual lipase.

[1]To whom correspondence should be addressed.

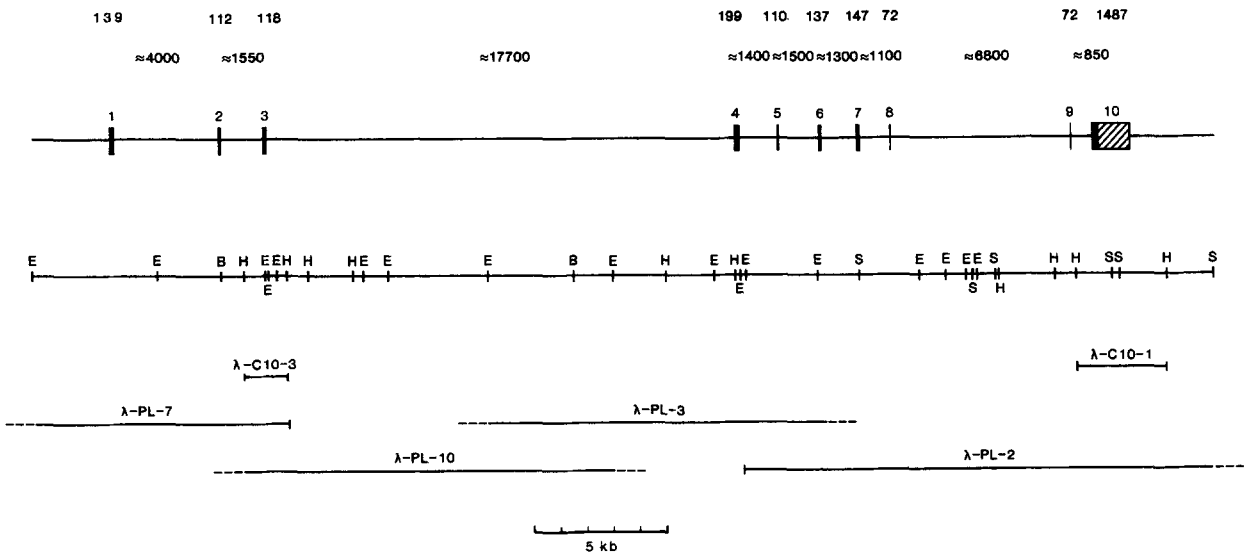

Fig. 1. Exon/intron organization and restriction enzyme cleavage map of the human lysosomal acid lipase/cholesteryl ester hydrolase gene derived from nine overlapping bacteriophage lambda clones obtained from two human genomic libraries. Open, solid, and hatched boxes represent 5′-nontranslated, coding, and 3′-nontranslated sequences, respectively. Introns are indicated by lines. The numbers above the gene structure refer to the size (in bp) of the exons and introns. The restriction map for BamH I (B), EcoR I (E), Hind III (H), and Sac I (S) is illustrated below. The regions spanned by six representative bacteriophage lambda clones used to characterize the gene are also shown.

sion, we have determined the organizations of the HGL, RLL, and HLAL genes and the sequences of the 5′-flanking and 5′-nontranslated regions.

MATERIALS AND METHODS

Isolation of HGL and RLL cDNAs

First strand cDNA synthesis of 10 μg of total RNA from human stomach and rat tongue (strain Sprague-Dawley) tissue, respectively, was performed with Superscript Plus reverse transcriptase (Bethesda Research Laboratories) utilizing an oligo $(dT)_{18}$ primer that contained, in addition, restriction sites for Sma I, BamH I, and Hpa I. The newly generated HGL and RLL cDNAs were further amplified by the PCR technique using a sense primer in exon 1 and an anti-sense oligonucleotide which was either located in the 3′-nontranslated region or identical to the restriction sites incorporated into the oligo $(dT)_{18}$ primer.

Isolation of HLAL cDNA

Total cellular RNA was isolated from human leukocytes and first strand cDNA synthesis was performed with an antisense primer located at the 3′-end of the 3′-untranslated region of the published HLAL cDNA sequence (18).

Genomic clone isolation

In the case of the HLAL gene which has been previously mapped to human chromosome 10 (24, 25), two bacteriophage lambda libraries, a genomic hamster (CHO)/human lymphocyte 762-8A hybrid library specific for human chromosome 10 (complete Hind III-digest; American Type Culture Collection; 26) and a human genomic placenta DNA library in Lambda Fix II (Stratagene) were screened by the filter hybridization technique (27) using as randomly primed, ^{32}P-labeled probes the human full-length HLAL cDNA and fragments thereof as well as DNA fragments of the human HLAL gene. Positive clones were plaque-purified and characterized by restriction endonuclease mapping and DNA sequencing. To determine the 5′-flanking and 5′-nontranslated sequences of the HGL and RLL genes whose chromosomal locations have not been reported, genomic fragments containing the 3′-end of exon 1 and the 5′-end of intron 1 were used as probes to screen the human placenta DNA library and a rat genomic library (strain Sprague-Dawley; Stratagene), respectively.

Intron amplification by PCR

Oligonucleotide primers identical and complementary to HLAL, HGL, and RLL cDNA sequences were synthesized by the phosphoramidite method on a DNA synthesizer (Applied Biosystems, Inc., Foster City, CA, model 381A) and desalted on NAP-5 columns (Pharmacia LKB Biotechnology Inc.). Each primer contained a native or incorporated restriction enzyme site for Bam HI, Eco RI, or Hind III for subsequent cloning purposes. The primers were used without further purification to amplify human and rat genomic DNA isolated from white blood cells by the polymerase chain reaction (28) using an automated DNA Thermal Cycler (Hybaid

TABLE 1. Exon and intron sizes and sequences at exon/intron boundaries of the HLAL gene

Exon	3′ Splice Acceptor	5′ Splice Donor	Exon Length, Base Pairs	Amino Acids Encoded	Intron Length	Codon Phase
		39				
1	CTGAAC	CTCCAGgtgaga	139	0	~4.00 kb	
	40	151				
2	atacagAATGAA	AATGTGgtaagt	112	37	~1.55 kb	0
	152	269				
3	ctacagAGTGAA	ACAAAGgtatgg	118	40	~17.7 kb	I
	270	468				
4	tttcagGTCCCA	TTTCAGgtatat	199	66	~1.40 kb	II
	469	578				
5	ttacagTTATGA	CTATAGgtatgt	110	37	~1.50 kb	I
	579	715				
6	tcacagGTTTTA	ATTAAGgtactt	137	45	~1.30 kb	0
	716	862				
7	ttgtagGACTTA	AATATGgtatgc	147	49	~1.10 kb	0
	863	934				
8	ttgcagTCTAGA	AGCCAGgtaggc	72	24	~6.80 kb	0
	935	1006				
9	ttttagGCTGTT	AACCAGgtaaag	72	24	~0.85 kb	0
	1007	2493				
10	ttgtagAGTTAT	AGGCCC	1487	77		

Exon and intron sequences are shown in uppercase and lowercase letters, respectively. The locations of the splice sites with respect to the published cDNA sequence (18) are indicated by the numbers above. Intron phases I, II, and 0 refer to introns that interrupt the coding sequence after the first, second, or third nucleotide of a codon, respectively.

Limited, UK, model OmniGene). A typical reaction contained 1 μg each of two oligonucleotides in a mixture of 50 mM KCl, 10 mM Tris-HCl, pH 9, 1.5 mM $MgCl_2$, 0.1% Triton X-100, and 200 μM each dATP, dCTP, dGTP, and dTTP. The cycle profile used was denaturation at 95°C for 1 min, annealing at 50° or 55°C for 1 min, and extension at 74°C for 1–5 min for 35–40 cycles.

Subcloning and sequencing

DNA inserts of recombinant clones and amplified PCR products containing HLAL, HGL, and RLL sequences were digested with the restriction enzymes BamH I, EcoR I, Hind III, Sac I, and Xho I (Bethesda Research Laboratories and New England Biolabs, Beverly, MA) according to the recommendations of the manufacturers. Restriction fragments were isolated by low melting point agarose (Bethesda Research Laboratories) gel electrophoresis and ligated into the Bluescript II KS phagemid (Stratagene). Nucleotide sequences of the exons, the exon/intron junctions, and the 5′-flanking regions were determined by the dideoxy chain termination method (29) with T7 DNA polymerase (Sequenase; United States Biochemical Corp.,

TABLE 2. Comparison of the HLAL intronic sequences published by Anderson et al. (31) and Aslanidis et al. (32) with our sequence data

Intron	Nucleotide(s)	Anderson et al.	Aslanidis et al.	Lohse et al.
Intron 1 5′-end	+11	C	C missing	C
Intron 1 5′-end	+23 to +24	C missing	C	C
Intron 1 5′-end	+35	G	not reported	G/A polymorphism
Intron 1 3′-end	−65	G	not reported	T
Intron 1 3′-end	−54	C	not reported	C missing
Intron 2 5′-end	+12 to +18	AAAGTTA	TTAAAAG	AAAGTTA
Intron 2 3′-end	−12 to −17		1 add. nucleotide	1 add. nucleotide
Intron 4 3′-end	−11 to −35		partial sequence	5 add. nucleotides
Intron 5 3′-end	−5	C	C	C/T polymorphism
Intron 6 3′-end	−42	G	not reported	G/A polymorphism
Intron 7 5′-end	+13 to +49		partial sequence	15 add. nucleotides
Intron 8 5′-end	+54	N	not reported	T
Intron 8 5′-end	+68	A	not reported	A missing
Intron 9 3′-end	−8 to −9		add. T	add. T
Intron 9 3′-end	−21 to −22		add. T	add. T

The numbers indicate nucleotide positions relative to the intron/exon (−) and exon/intron (+) borders according to Anderson et al. (31).

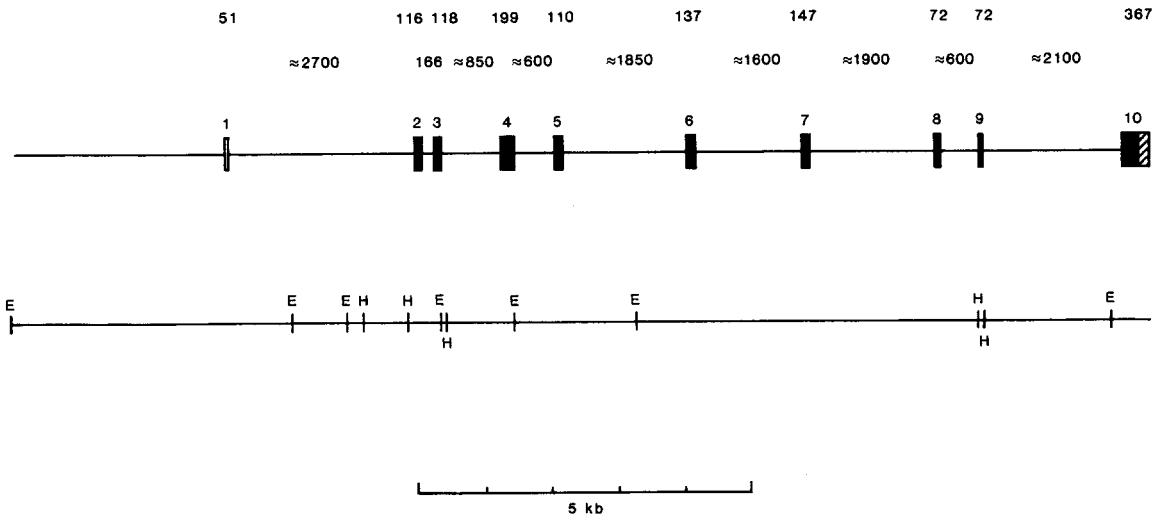

Fig. 2. Exon/intron organization and restriction enzyme cleavage map of the human gastric lipase gene derived from PCR-amplified genomic DNA fragments and a bacteriophage lambda clone. 5′-Nontranslated, coding, and 3′-nontranslated sequences are represented by open, solid, and hatched boxes, respectively, while introns are indicated by lines, and their respective sizes (in bp) are given above the gene structure. Recognition sites for the restriction enzymes Hind III (H) and EcoR I (E) are shown below.

Cleveland, OH) using T3 and T7 primers as well as oligonucleotides corresponding to known sequences within the HLAL, HGL, and RLL genes.

Determination of the 5′-end of the HGL and RLL mRNA

An aliquot of the first strand cDNA was dG-tailed with terminal deoxynucleotidyl transferase (Bethesda Research Laboratories) by a modified 5′-RACE technique (30). It was then PCR-amplified using an oligo $(dC)_{12}$ primer containing additional restriction sites for Hind III and BamH I and a HGL specific oligonucleotide complementary to bases 222 to 245 (16) and a RLL specific primer complementary to nucleotides 224 to 247 (17), respectively. One tenth of the reaction was used for a second round of amplification with the oligo $(dC)_{12}$ primer and an internal oligonucleotide complementary to the region 61 to 84 of the HGL cDNA (16) and to bases 71 to 94 of the RLL cDNA (17), respectively. The resulting fragments were purified by low melting point agarose gel electrophoresis, subcloned, and sequenced.

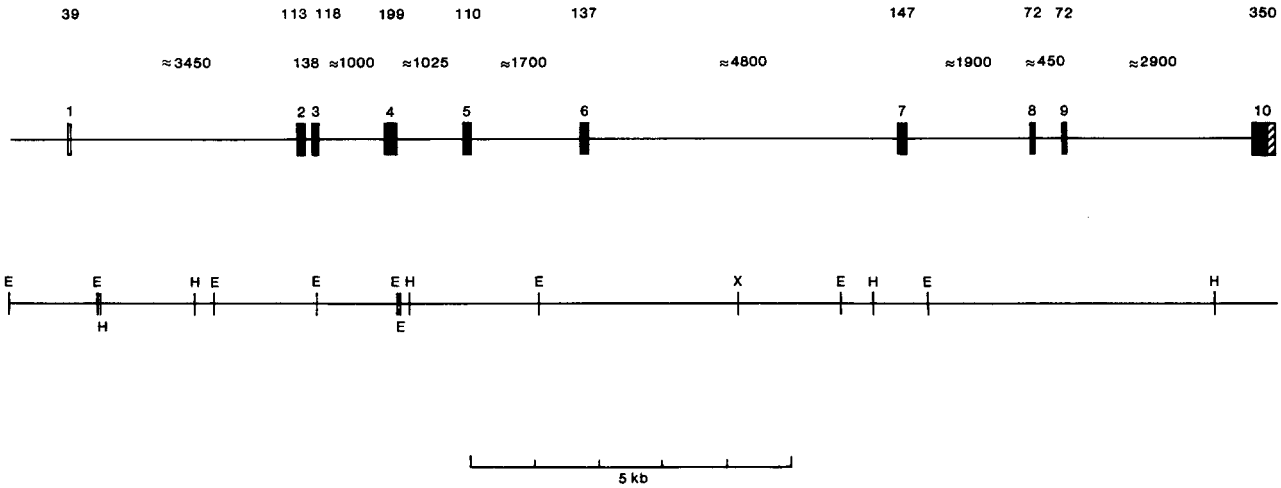

Fig. 3. Exon/intron organization and restriction enzyme cleavage map of the rat lingual lipase gene derived from PCR-amplified genomic DNA fragments and a bacteriophage lambda clone. Exons are indicated by solid boxes, 5′-noncoding parts are left open, and the 3′-noncoding region is hatched. Introns are shown as lines. The respective sizes of the RLL exons and introns are illustrated above the gene structure, while a restriction map for the enzymes Hind III (H), EcoR I (E), and Xho I (X) is depicted below.

TABLE 3. Exon and intron sizes and sequences at exon/intron boundaries of the HGL gene

Exon	3′ Splice Acceptor	5′ Splice Donor	Exon Length, Base Pairs	Amino Acids Encoded	Intron Length	Codon Phase
		33				
1	AAAATA	GAAACTgtaagt	51	0	~2.70 kb	
	34	149				
2	tttcagGCAGGT	AACATTgtaagt	116	35	166 bp	0
	150	267				
3	cctcagAGTCAG	ATACAGgtatat	118	40	~850 bp	I
	268	466				
4	ccttagGCCAGA	TTTCAGgtaaac	199	66	~600 bp	II
	467	576				
5	tttcagCTTTGA	CCATTGgtaagt	110	37	~1.85 kb	I
	577	713				
6	tttcagGTTTTA	TTCAAGgtatgc	137	45	~1.60 kb	0
	714	860				
7	tttcagTTTATA	AACACGgttagt	147	49	~1.90 kb	0
	861	932				
8	ttctagAGTCGC	ACCCAGgtattc	72	24	~600 bp	0
	933	1004				
9	tgttagGCTGTT	GATCAGgtaagc	72	24	~2.10 kb	0
	1005	1371				
10	ttctagTCCCAA	ATCAAC	367	78		

Exon and intron sequences are given in uppercase and lowercase letters, respectively. The positions of the splice sites with respect to the cDNA sequence (16) are indicated by the numbers above. Intron phases I, II, and 0 refer to introns that interrupt the coding sequence after the first, second, or third nucleotide of a codon, respectively.

RESULTS AND DISCUSSION

The high degree of similarity of the nucleotide and amino acid sequences of HGL, HLAL, and RLL led to the conclusion that these enzymes are members of an acid lipase family (16, 18). To investigate whether the respective genes have evolved from a common ancestor via duplication or triplication, we first elucidated the HLAL gene organization using two different methods. HLAL introns smaller than 2 kb were identified by polymerase chain reaction-based intron amplification of human genomic DNA. The larger introns of the HLAL gene were isolated by conventional plaque hybridization of two bacteriophage lambda libraries.

The HLAL gene consists of 10 exons spread over more than 38.8 kb on human chromosome 10 and is interrupted by 9 introns (**Fig. 1**). The first exon contains the 5′-nontranslated region, while exon 2 encodes the last nucleotide of the 5′-nontranslated region, the entire signal peptide, and the first 16 residues of the mature protein (**Table 1**). Exons 3 through 9, with lengths ranging from 72 bp to 199 bp, code for amino acids 17–301. Exon 10 is in comparison unusually large and contains the coding sequence for residues 302–378, the translation stop codon, and the long 3′-untranslated region, encompassing 1253 nucleotides. The intron lengths vary considerably, the smallest intron 9 being approximately 850 bp long, while intron 3 spans about 17.7 kb and intron 8 about 6.8 kb of DNA.

The exon sequences exactly matched with the published HLAL cDNA sequence (18). Within the signal peptide, however, we observed a common DNA polymorphism due to the substitution of an A for a C, resulting in the replacement of proline (CCC), residue 16, by threonine (ACC) and in the loss of a cleavage site for the restriction enzyme Hae III (5′-GG/CC-3′). A second polymorphic site, also located in exon 2, leads to the substitution of amino acid 2 of the mature enzyme, glycine (GGA), with arginine (AGA), thereby creating a new restriction enzyme cleavage site for Xba I (5′-T/CTAGA-3′). This conversion adds one positive charge unit to the HLAL-1 isoenzyme (pI 6.27), thus creating the more basic HLAL-2 isoenzyme (pI 6.4).

The intron sequences were also found to be polymorphic (**Table 2**). Sequence analyses of Wolman and cholesteryl ester storage disease (CESD) patients revealed a G/A polymorphism at the 5′-end of intron 1 (position +36 relative to the exon/intron border), a C/T substitution at the 3′-end of intron 5 (position −5 relative to the intron/exon border), an A/C replacement at the 5′-end of intron 6 (position +62), and an A/G polymorphism at the 3′-end of intron 6 (position −42). Together with the Pro/Thr and Gly/Asp amino acid exchanges, these sites have been successfully used for haplotype analysis in affected families (P. Lohse, unpublished results).

Overall, our data are in agreement with the HLAL genomic structures published by Anderson et al. (31) and Aslanidis et al. (32). In our study, however, the gene

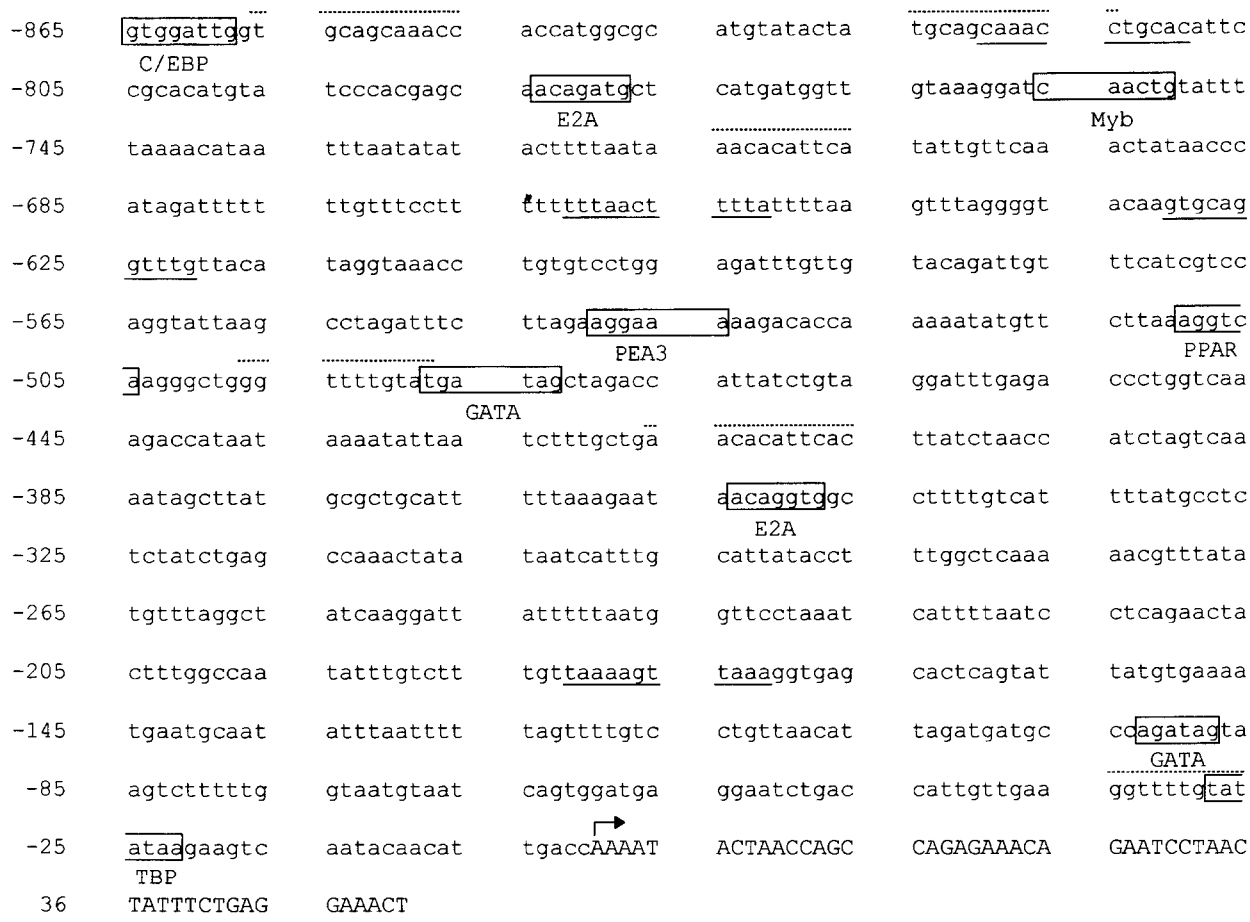

Fig. 4. 5′-Flanking and 5′-nontranslated sequences of the HGL gene and potential regulatory elements of the HGL promoter. The sequences of the first exon and of the nontranscribed 5′-flanking region, determined on both strands, are shown in upper- and lowercase letters, respectively. The arrow identifies the major transcription initiation site. The numbers refer to nucleotide positions relative to the start site of transcription. Putative binding sequences for TATA box-binding factor (TBP) and the transcription factors C/EBP, E2A, GATA, Myb, PEA3, and PPAR (43) are boxed. Direct repeats are overdotted and indirect repeats are underlined.

locus was determined to be slightly larger than the size reported by Anderson et al. (38.8 kb versus 36 kb). Homology analysis of the intron sequences (Table 2) demonstrated a rather high degree of divergence considering the fact that the same bacteriophage lambda library was screened by all three groups. These differences can only partially be explained as a result of DNA polymorphisms.

The structures of the HGL and RLL genes were determined solely by intron amplification using cDNA primers designed to flank the putative exon/intron borders, based on the expected homology with the previously characterized HLAL gene organization. This method appears to be especially suited for the structural analysis of closely related genes. The only limitation is the size of the introns, as we were unable to amplify intervening sequences in excess of 5.5 kb. Shortening the length of the PCR denaturation step and the use of high concentrations of an exonuclease-free, aminoterminal deletion mutant of Taq DNA polymerase (Klentaq I) in combination with a DNA polymerase with 3′-exonuclease proofreading activity in very low concentrations (33) may help to circumvent this problem.

Subsequent restriction site mapping and DNA sequence analysis established the linear order of the PCR-amplified HGL- and RLL-specific genomic fragments (**Fig. 2** and **Fig. 3**). The predicted structures of exons 1–4 and their intervening sequences were confirmed by results obtained from bacteriophage lambda clones isolated for the sequence analysis of the HGL and RLL 5′-nontranslated and 5′-flanking regions.

The HGL gene spans approximately 14 kb of genomic DNA and includes ten exons and nine introns (Fig. 2) that are bounded by sequences matching the splice site GT/AG consensus sequences (34; **Table 3**). The major transcriptional start site of HGL mRNA was determined by a modified 5′-RACE technique (30) as an A 62 bp upstream of the translation initiation codon ATG

TABLE 4. Exon and intron sizes and sequences at exon/intron boundaries of the RLL gene

Exon	3′ Splice Acceptor	5′ Splice Donor	Exon Length, Base Pairs	Amino Acids Encoded	Intron Length	Codon Phase
		37				
1	AGGCAA	GGCACTgtaagt	39	0	~3.45 kb	
	38	150				
2	tttcagAGCAGT	AATATTgtgagt	113	34	138 bp	0
	151	268				
3	cctcagAGTCAG	ATATAGgtacaa	118	40	~1.00 kb	I
	269	467				
4	ccctagGCAAGA	TTTCAGgtaagg	199	66	~1.03 kb	II
	468	577				
5	ttacagCTTTGA	CTATTGgtaagt	110	37	~1.70 kb	I
	578	714				
6	tttcagGTTTCA	TTCAAGgtatgt	137	45	~4.80 kb	0
	715	861				
7	tttcagCTTATG	AATGTGgtgtgt	147	49	~1.90 kb	0
	862	933				
8	ttctagAGTCGT	GCACAGgtatca	72	24	~450 bp	0
	934	1005				
9	tattagCTTGTT	AACCAGgtatgt	72	24	~2.90 kb	0
	1006	1355				
10	tttcagAAAACG	ATCAGC	350	76		

Exon and intron sequences are presented in uppercase and lowercase letters, respectively. The intron positions with respect to the RLL cDNA sequence (17) are indicated by the numbers above. Intron phases I, II, and 0 refer to introns that interrupt the coding sequence after the first, second, or third nucleotide of a codon, respectively.

(**Fig. 4**). Exon 1 consists of 51 bp of the 5′-nontranslated region, while exon 2 codes for the remaining 11 bp of the 5′-nontranslated sequence, the 19 residues of the signal peptide, and the first 16 amino acids of the mature protein. The largest exon in the HGL gene, exon 10, encodes the carboxyl-terminal residues 302–379, the translation stop codon TAG, and the complete 3′-nontranslated region of 130 bp. Our sequence data agreed with the published HGL cDNA sequence except for six additional nucleotides (ATCAAC) at the 3′-end of exon 10.

The RLL gene is spread over approximately 18.7 kb and also contains 10 exons interrupted by nine introns (Fig. 3). Each intron was found to begin with the dinucleotide GT and to end with the dinucleotide AG (34; **Table 4**). Reverse transcription of the cap structure (30) identified the mRNA 5′-end as an A 50 bp upstream of the ATG start codon (**Fig. 5**). The genomic sequences agreed with the published RLL cDNA sequence (17) with the exception of six differences in exon 1. These include a C instead of a G at position 4 of the RLL cDNA, a missing T at 14, a T in between nucleotides 21 and 23, a G instead of an A at 29, a C instead of an A at 34, and an A instead of a C at 35 (nucleotide numbering according to ref. 17; Fig. 5). In each case, these differences may be due to DNA polymorphism or possible sequencing errors. The source of the mRNA used for cDNA synthesis and of the genomic DNA used for intron amplification as well as genomic DNA library construction, however, was the same Sprague-Dawley rat strain.

Comparison with the HGL gene organization revealed that both genes are identical in structure. The only differences were a shorter 5′-nontranslated region of 39 instead of 51 base pairs contained in RLL exon 1, the deletion of one amino acid triplet, corresponding to residue 7 of the HGL signal peptide, in RLL exon 2, and the smaller size of the 3′-nontranslated region coded for by RLL exon 10, encompassing 119 instead of 130 nucleotides.

The HGL/RLL gene structure is also identical with that of the HLAL gene, except for the position of HLAL intron 1. Exons 3–9 have the same length and code for the same number of amino acids. The second exon of the RLL and HGL genes is 1 bp and 4 bp longer, respectively, but encodes only 34 and 35 instead of 37 amino acids, respectively, because of the shorter signal peptides of RLL (18 instead of 21 residues in HLAL) and HGL (19 amino acids). Exon 10 of the HLAL gene (1487 bp), in contrast, is much larger in size than the corresponding exons of the HGL gene (367 bp) and the RLL gene (350 bp) due to the presence of a long and A/T rich 3′-nontranslated region encompassing 1253 bp, and codes for 77 instead of 78 and 76 amino acids in HGL and RLL, respectively.

All introns interrupt the coding sequences of the HGL, RLL, and HLAL genes in identical codon phases (**Fig. 6**). The intron lengths, in contrast, show much

Fig. 5. 5′-Flanking and 5′-nontranslated sequences of the RLL gene and potential regulatory elements of the RLL promoter. The RLL exon 1 sequence is depicted in uppercase letters, while the 5′-flanking sequences of the promoter region are shown in lowercase letters, based on the sequence analysis of the complementary DNA strands. The transcription initiation site is indicated by an arrow. Potential binding sites for transcription factors GATA, GHF, HNF-5, Myb, and TBP (43) as well as the putative CCAAT-element are boxed. A direct repeat is overdotted and an indirect repeat is underlined.

greater variability. HGL and RLL introns 1, 2, 3, 4, and 8 are in part considerably smaller and HGL and RLL introns 5, 7, and 10 are slightly larger than the corresponding introns of the HLAL gene, resulting in much smaller HGL and RLL genes. Intron 6 is the largest intron of the RLL gene, containing about 4.8 kb compared to approximately 1.6 kb and 1.3 kb in the HGL and HLAL genes, respectively, and mainly accounts for the overall size difference of the RLL and HGL gene loci.

The precise conservation of the exon/intron structures and the similarity of the coding sequences indicate that HGL/RLL and HLAL are most likely products of a gene duplication event and have subsequently evolved in parallel, assuming distinct roles in lipid metabolism due to sequence divergence.

With regard to exon sizes and codon phases, the structural identity of the HGL, RLL, and HLAL genes is even more pronounced than that of the human lipoprotein lipase (35, 36) and human hepatic lipase (37, 38) genes (Fig. 6) which, together with the gene for human pancreatic lipase (39), form another branch of the mammalian lipase gene family without structural homology to the acidic lipases except for the G–X–S–X–G lipase consensus sequence. This suggests either that the gene-duplication event leading to HGL/RLL and HLAL occurred later in evolution than the duplication producing the lipoprotein lipase and hepatic lipase genes or that the evolutionary pressure applied to maintain the integrity of the HLAL and HGL/RLL gene organizations is higher. The results also indicate that the ability of HLAL to hydrolyze cholesteryl esters, in addition to triglycerides, is due to amino acid substitutions within the conserved exon structure. Alternatively, HGL and RLL may have lost this property as a consequence of mutations affecting a site necessary for interaction with cholesteryl esters. The latter assumption would be in agreement with our hypothesis that the ancestral acid

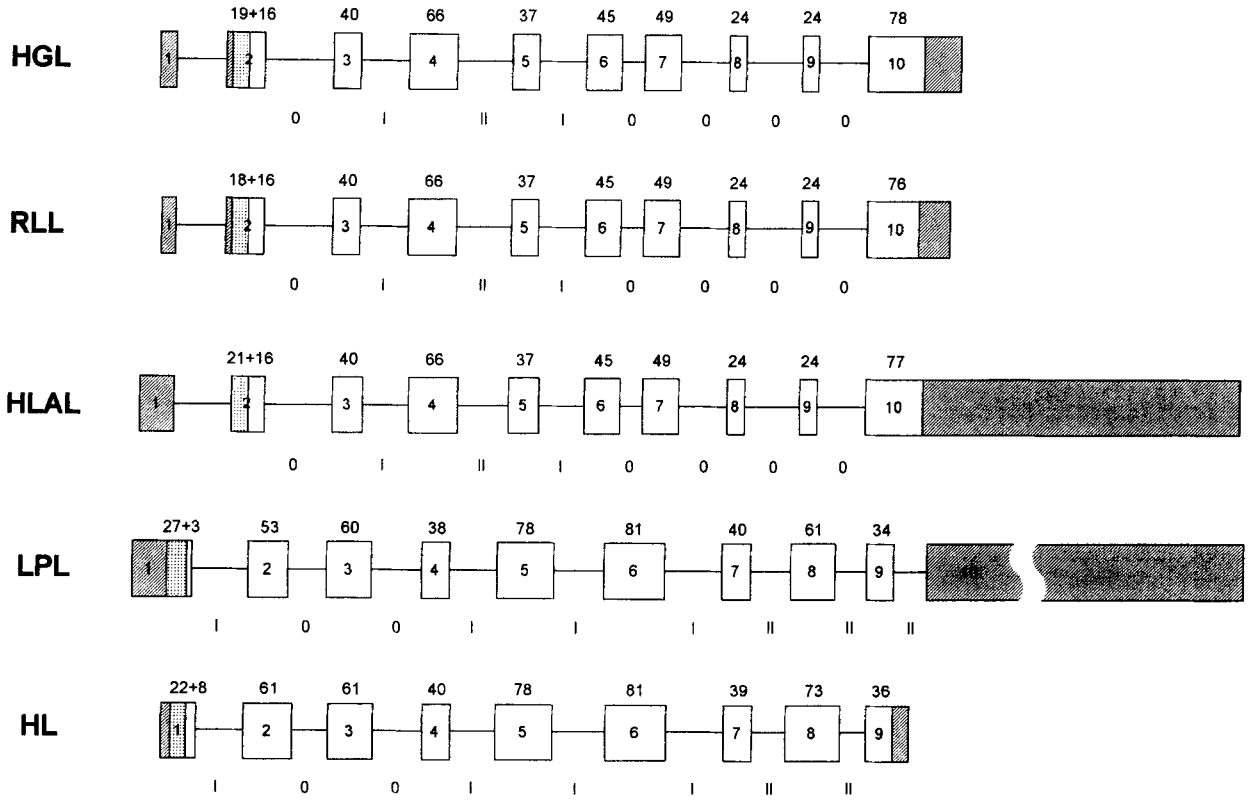

Fig. 6. Comparison of the HGL, RLL, HLAL, human lipoprotein lipase (LPL), and human hepatic lipase (HL) gene organizations. Signal peptide, mature protein, and nontranslated sequences are indicated by stippled, open, and hatched boxes, respectively. Introns (not drawn to scale) are represented by lines. The numbers above the gene structures refer to the amino acids encoded by an exon. Numerals I, II, and 0 below indicate whether an intron interrupts the coding sequence after the first, second or third nucleotide of a codon, respectively. Note that exon 10 of the lipoprotein lipase gene is not drawn to scale.

lipase gene has evolved as a consequence of exon shuffling that combined lipase-like and cholesteryl esterase-like domains, as has been suggested previously for the evolution of the bile-salt stimulated lipase/cholesteryl esterase gene (40).

Our studies may also help to solve the question of whether or not there is a gastric lipase, typical of omnivores, in addition to lingual lipase in rodents and to pregastric lipase in ruminants (21). Faced with the choice of one enzyme produced at two different anatomic sites or two different enzymes with similar characteristics, our results demonstrating the complete identity of the HGL and RLL gene organizations strongly suggest that RLL is the gastric lipase of rats produced in the lingual serous glands.

In this context, it is interesting to note that the 5′-regulatory sequences for the three genes differ markedly, in accordance with the very different pattern of tissue-specific expression. Screening of the 5′-nontranslated region and of 750 bp of the 5′-flanking sequence of the HLAL gene (**Fig. 7**) for eukaryotic promoter elements by computer scan (PC/GENE; IntelliGenetics, Mountain View, CA; 41, 42) demonstrated structural features typical of a housekeeping gene constitutively expressed in all cells, consistent with the enzyme's role in maintaining intracellular lipid homeostasis. First, the sequence lacks a TATA-box structure, but does contain putative regulatory protein-binding sites commonly observed in promoters of housekeeping genes, including four transcription control elements recognized by the ubiquitous transcription factor Sp1 (43). Second, the promoter region was found to have a high GC content within the 350 bp sequence immediately upstream of the cap site; it averaged 67.5% compared with approximately 41% for the entire human genome.

The HLAL exon 1 and 5′-flanking sequences reported here are completely homologous to those determined by Aslanidis et al. (32) with two exceptions at positions −106 to −111 and −488 relative to the transcription initiation site (numbering according to ref. 32 with nucleotide −100 corresponding to +1; **Table 5**). In addition, there are two polymorphic sites in the HLAL promoter region immediately upstream of the cap site. A C/A substitution at nucleotide −59 can be

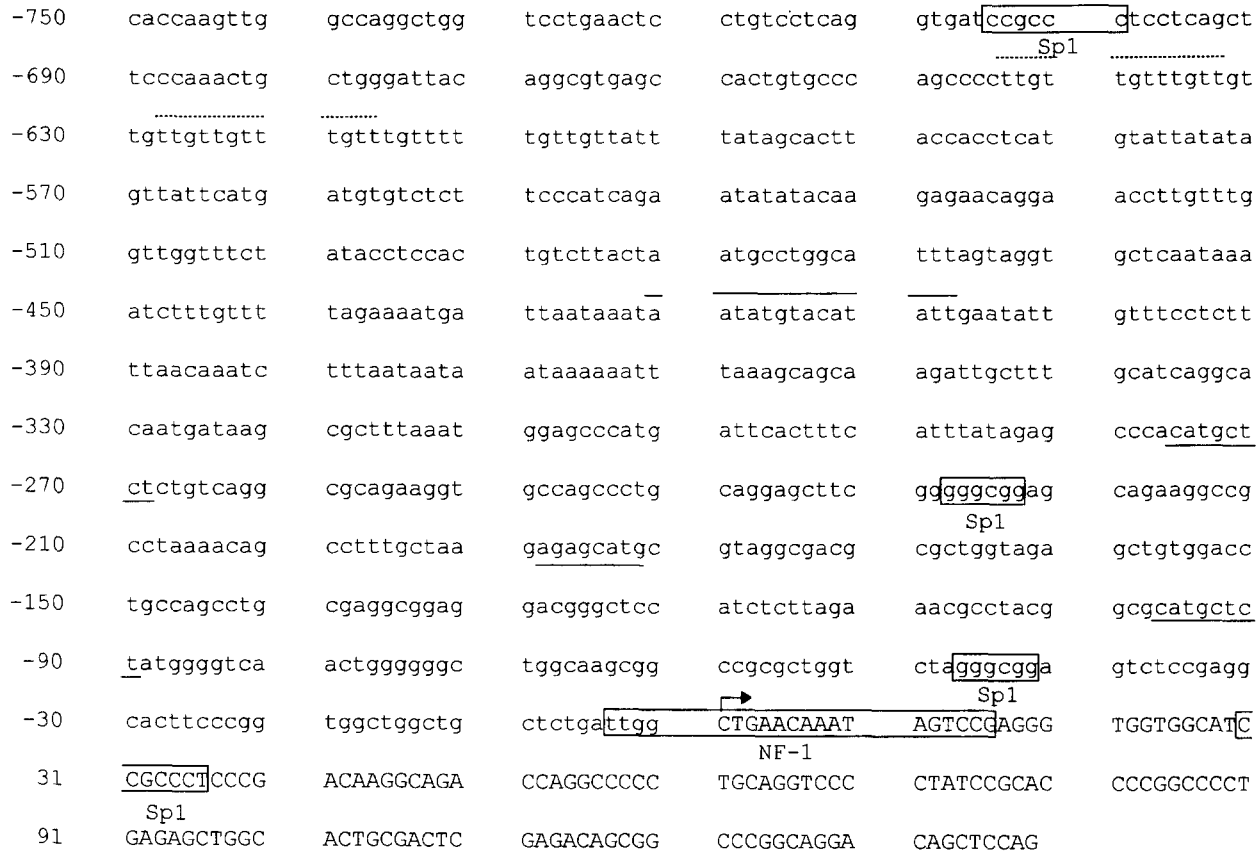

Fig. 7. 5′-Flanking and 5′-nontranslated sequences of the HLAL gene and potential regulatory elements of the HLAL promoter. The sequences of each DNA strand were determined by the Sanger dideoxy chain termination method (29). Transcribed regions of exon 1 are shown in uppercase letters and flanking sequences are given in lowercase letters. The arrow indicates the major start site of transcription (31). The numbers refer to nucleotide positions relative to the transcription initiation site. Consensus sequences recognized by transcription factors Sp1 and CTF/NF-1 (43) are boxed. A palindromic sequence is overlined, a direct repeat is overdotted, and inverted repeats are underlined.

detected by the presence (C) or absence (A) of a restriction enzyme cleavage site for Hae III (5′-GG/CC-3′). The second G/A polymorphism at position −48 destroys (G) or creates (A) a site for the restriction endonuclease Mae I (5′-C/TAG-3′). The 5′-flanking sequence determined by us differs from the 5′-region published by Anderson et al. (31) in 28 nucleotides and the HLAL exon 1 sequences are at variance at positions +24, +25, +27, +28, +119, and +120 (Table 5).

In contrast to the HLAL promoter, the 5′-flanking sequence of the HGL gene is AT-rich (66.7%), has a potential TATA box element 21 bp upstream of the cap

TABLE 5. Comparison of the HLAL 5′-flanking region and exon 1 sequences presented in this paper (Fig. 7) with those reported by Anderson et al. (31) and Aslanidis et al. (32)

	Nucleotide(s)	Anderson et al.	Aslanidis et al.	Lohse et al.
5′-Flanking region	−488	not reported	T	T missing
	−111 to −106	AACG	TACACA	AAACGC
	−59	A	A	A/C polymorphism
	−48	G	G	G/A polymorphism
Exon 1	+24	C	T	T
	+25	T	G	G
	+27	A	C	C
	+28	G	A	A
	+119	A	G	G
	+120	A	G	G

The nucleotides of the 5′-flanking sequence are numbered according to ref. 32, position −100 corresponding to the putative transcription initiation site +1 (31).

site, and contains multiple putative binding sites for sequence-specific DNA-binding proteins such as C/EBP, E2A, GATA, Myb, and PEA3 (43; Fig. 4). The RLL 5′-flanking region of 898 nucleotides upstream of the mapped transcriptional start site also has a high A/T content of 67.8% and contains a TATA box sequence 25 bp 5′ of the cap site (Fig. 5). In addition, a CCAAT element is present 342 bp upstream of the TATA-box, outside of the region of preferential occurrence of this motif in vertebrate promoter sequences (−212 to −57; 42). Comparison with consensus sequences for transcription factor binding sites (43) provided matches for two GATA sequences, one GHF sequence, one HNF-5 sequence, and one Myb sequence.

The characterization of the HLAL, HGL, and RLL promoter regions may provide the information necessary to define the regulatory elements that control developmental and tissue-specific expression of the respective genes.

In summary, we have elucidated the HLAL, HGL, and RLL gene structures and determined the sequences of exon 1 and of 750, 865, and 898 bp of the 5′-flanking regions, respectively. The three genes are organized into ten exons separated by nine introns and span approximately 38.8 kb (HLAL), 14 kb (HGL), and 18.7 kb (RLL) of the human and rat genome. The size differences are mainly due to variations in intron length. With regard to exon sizes and intron phase class, the HGL and RLL gene structures are identical, while the HLAL gene differs in the location of intron 1. These results strongly suggest that the HGL/RLL and HLAL genes arose by gene duplication and assumed different roles in neutral lipid metabolism due to protein sequence divergence as well as tissue-specific expression.

We would like to thank Johanna Brandecker and Christian Schmidt for preparing the figures. This work was supported by research grant Lo 391/2-1 from the Deutsche Forschungsgemeinschaft.

Manuscript received 21 November 1996 and in revised form 10 February 1997.

REFERENCES

1. Gargouri, Y., H. Moreau, and R. Verger. 1989. Gastric lipases: biochemical and physiological studies. *Biochim. Biophys. Acta.* **1006:** 255–271.
2. Hamosh, M. 1990. Lingual and Gastric Lipases. Their Role in Fat Digestion. M. Hamosh, editor. CRC Press, Boca Raton. 1–239.
3. Hamosh, M. 1990. Lingual and Gastric Lipases. *Nutrition.* **6:** 421–428.
4. Moreau, H., R. Laugier, Y. Gargouri, F. Ferrato, and R. Verger. 1988. Human preduodenal lipase is entirely of gastric fundic origin. *Gastroenterology.* **95:** 1221–1226.
5. Moreau, H., A. Bernadac, Y. Gargouri, F. Benkouka, R. Laugier, and R. Verger. 1989. Immunocytolocalization of human gastric lipase in chief cells of the fundic mucosa. *Histochemistry.* **91:** 419–423.
6. Carriere, F., J. A. Barrowman, R. Verger, and R. Laugier. 1993. Secretion and contribution to lipolysis of gastric and pancreatic lipases during a test meal in humans. *Gastroenterology.* **105:** 876–888.
7. Bernbäck, S., L. Bläckberg, and O. Hernell. 1989. Fatty acids generated by gastric lipase promote human milk triacylglycerol digestion by pancreatic colipase-dependent lipase. *Biochim. Biophys. Acta.* **1001:** 286–293.
8. Bernbäck, S., L. Bläckberg, and O. Hernell. 1990. The complete digestion of human milk triacylglycerol in vitro requires gastric lipase, pancreatic colipase-dependent lipase, and bile salt-stimulated lipase. *J. Clin. Invest.* **85:** 1221–1226.
9. Abrams, C. K., M. Hamosh, V. S. Hubbard, S. K. Dutta, and P. Hamosh. 1984. Lingual lipase in cystic fibrosis. Quantitation of enzyme activity in the upper small intestine of patients with exocrine pancreatic insufficiency. *J. Clin. Invest.* **73:** 374–382.
10. Abrams, C. K., M. Hamosh, S. K. Dutta, V. S. Hubbard, and P. Hamosh. 1987. Role of nonpancreatic lipolytic activity in exocrine pancreatic insufficiency. *Gastroenterology.* **92:** 125–129.
11. Moreau, J., M. Bouisson, D. Balas, A. Ravaud, S. Stupnik, L. Buscail, N. Vaysse, and A. Ribet. 1990. Gastric lipase in alcoholic pancreatitis. Comparison of secretive profiles following pentagastrin stimulation in normal adults and patients with pancreatic insufficiency. *Gastroenterology.* **99:** 175–180.
12. Balasubramanian, K., P. L. Zentler-Munro, J. C. Batten, and T. C. Northfield. 1992. Increased intragastric acid-resistant lipase activity and lipolysis in pancreatic steatorrhoea due to cystic fibrosis. *Pancreas.* **7:** 305–310.
13. Tiruppathi, C., and K. A. Balasubramanian. 1982. Purification and properties of an acid lipase from human gastric juice. *Biochim. Biophys. Acta.* **712:** 692–697.
14. Moreau, H., Y. Gargouri, D. Lecat, J. L. Junien, and R. Verger. 1988. Purification, characterization and kinetic properties of the rabbit gastric lipase. *Biochim. Biophys. Acta.* **960:** 286–293.
15. Carriere, F., H. Moreau, V. Raphel, R. Laugier, C. Benicourt, J. L. Junien, and R. Verger. 1991. Purification and characterization of dog gastric lipase. *Eur. J. Biochem.* **202:** 75–83.
16. Bodmer, M. W., S. Angal, G. T. Yarranton, T. J. R. Harris, A. Lyons, D. J. King, G. Pieroni, C. Riviere, R. Verger, and P. A. Lowe. 1987. Molecular cloning of a human gastric lipase and expression of the enzyme in yeast. *Biochim. Biophys. Acta.* **909:** 237–244.
17. Docherty, A. J. P., M. W. Bodmer, S. Angal, R. Verger, C. Riviere, P. A. Lowe, A. Lyons, J. S. Emtage, and T. J. R. Harris. 1985. Molecular cloning and nucleotide sequence of rat lingual lipase. *Nucleic Acids Res.* **13:** 1891–1903.
18. Anderson, R. A., and G. N. Sando. 1991. Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesteryl ester hydrolase. *J. Biol. Chem.* **266:** 22479–22484.
19. Roberts, I. M., and R. Jaffe. 1986. Lingual lipase: immunocytochemical localization in the rat von Ebner gland. *Gastroenterology.* **90:** 1170–1175.
20. Hamosh, M. 1984. Lingual lipase. *In* Lipases. B. Borgström and H. L. Brockman, editors. Elsevier Science Publishers, Amsterdam. 49–81.

21. Hamosh, M. 1986. Lingual lipase. *Gastroenterology.* **90:** 1290–1297.
22. Fowler, S. D., and W. J. Brown. 1984. Lysosomal acid lipase. *In* Lipases. B. Borgström and H. L. Brockman, editors. Elsevier Science Publishers, Amsterdam. 329–364.
23. Assmann, G., and U. Seedorf. 1995. Acid lipase deficiency: Wolman disease and cholesteryl ester storage disease. *In* The Metabolic and Molecular Bases of Inherited Disease. C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, editors. McGraw-Hill, Inc., New York. 2563–2587.
24. Van Cong, N., D. Weil, M. C. Hors-Cayla, M. S. Gross, S. Heuertz, C. Foubert, and J. Frezal. 1980. Assignment of the genes for human lysosomal acid lipases A and B to chromosomes 10 and 16. *Hum. Genet.* **55:** 375–381.
25. Anderson, R. A., N. Rao, R. S. Byrum, C. B. Rothschild, D. W. Bowden, R. Hayworth, and M. Pettenati. 1993. In situ localization of the genetic locus encoding the lysosomal acid lipase/cholesteryl esterase (LIPA) deficient in Wolman disease to chromosome 10q23.2–q23.3. *Genomics.* **15:** 245–247.
26. Deaven, L. L., M. A. Van Dilla, M. F. Bartholdi, A. V. Carrano, L. S. Cram, J. C. Fuscoe, J. W. Gray, C. E. Hildebrand, R. K. Moyzis, and J. Perlman. 1986. Construction of human chromosome-specific DNA libraries from flow-sorted chromosomes. *Cold Spring Harb. Symp. Quant. Biol.* **51:** 159–167.
27. Benton, W. D., and R. W. Davis. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. *Science.* **196:** 180–182.
28. Mullis, K. B., and F. A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155:** 335–350.
29. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74:** 5463–5467.
30. Hirzmann, J., D. Luo, J. Hahnen, and G. Hobom. 1993. Determination of messenger RNA 5′-ends by reverse transcription of the cap structure. *Nucleic Acids Res.* **21:** 3597–3598.
31. Anderson, R. A., R. S. Byrum, P. M. Coates, and G. N. Sando. 1994. Mutations at the lysosomal acid cholesteryl ester hydrolase gene locus in Wolman disease. *Proc. Natl. Acad. Sci. USA.* **91:** 2718–2722.
32. Aslanidis, C., H. Klima, K. J. Lackner, and G. Schmitz. 1994. Genomic organization of the human lysosomal acid lipase gene (LIPA). *Genomics.* **20:** 329–331.
33. Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci. USA.* **91:** 2216–2220.
34. Mount, S. M. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res.* **10:** 459–472.
35. Deeb, S. S., and R. Peng. 1989. Structure of the human lipoprotein lipase gene. *Biochemistry.* **28:** 4131–4135.
36. Kirchgessner, T. G., J-C. Chuat, C. Heinzmann, J. Etienne, S. Guilhot, K. Svenson, D. Ameis, C. Pilon, L. D'Auriol, A. Andalibi, M. C. Schotz, F. Galibert, and A. J. Lusis. 1989. Organization of the human lipoprotein lipase gene and evolution of the lipase gene family. *Proc. Natl. Acad. Sci. USA.* **86:** 9647–9651.
37. Cai, S-J., D. M. Wong, S-H. Chen, and L. Chan. 1989. Structure of the human hepatic triglyceride lipase gene. *Biochemistry.* **28:** 8966–8971.
38. Ameis, D., G. Stahnke, J. Kobayashi, J. McLean, G. Lee, M. Büscher, M. C. Schotz, and H. Will. 1990. Isolation and characterization of the human hepatic lipase gene. *J. Biol. Chem.* **265:** 6552–6555.
39. Sims, H. F., M. L. Jennens, and M. E. Lowe. 1993. The human pancreatic lipase-encoding gene: structure and conservation of an Alu sequence in the lipase gene family. *Gene.* **131:** 281–285.
40. Fontaine, R. N., C. P. Carter, and D. Y. Hui. 1991. Structure of the rat pancreatic cholesterol esterase gene. *Biochemistry.* **30:** 7008–7014.
41. Bucher, P., and E. N. Trifonov. 1986. Compilation and analysis of eukaryotic POL II promoter sequences. *Nucleic Acids Res.* **14:** 10009–10026.
42. Bucher, P. 1990. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **212:** 563–578.
43. Faisst, S., and S. Meyer. 1992. Compilation of vertebrate-encoded transcription factors. *Nucleic Acids Res.* **20:** 3–26.