Identification of a novel lipase gene mutated in lpd mice with hypertriglyceridemia and associated with dyslipidemia in humans

Xiao-Yan Wen1, 7, *, Robert A. Hegele2, Jian Wang2, Ding Yan Wang1, Joseph Cheung3, Michael Wilson4, Maryam Yahyapour1, Yahong Bai1, Lihua Zhuang1, Jennifer Skaug3, T. Kue Young5, Philip W. Connelly6, Ben F. Koop4, Lap-Chee Tsui3, 7 and A. Keith Stewart1, 7

1Department of Experimental Therapeutics, Toronto General Research Institute, Toronto General Hospital, University Health Network, Toronto, Ontario, Canada, 2Robarts Research Institute, London, Ontario, Canada, 3The Hospital for Sick Children Research Institute, Toronto, Ontario, Canada, 4University of Victoria, Victoria, British Columbia, Canada, 5Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada, 6Departments of Laboratory Medicine and Pathobiology, University of Toronto and St Michael’s Hospital, Toronto, Ontario, Canada and 7McLaughlin Center for Molecular Medicine, Toronto, Ontario, Canada

Received January 3, 2003; Revised February 28, 2003; Accepted March 12, 2003

Triglyceride (TG) metabolism is crucial for whole body and local energy homeostasis and accumulating evidence suggests an independent association between plasma TG concentration and increased atherosclerosis risk. We previously generated a mouse insertional mutation lpd (lipid defect) whose phenotype included elevated plasma TG and hepatic steatosis. Using shotgun sequencing (~500 kb) and bioinformatics, we have now identified a novel lipase gene lpd1 (lpd lipase) within the lpd locus, and demonstrate the genetic disruption of exon 10 of lpd1 in the lpd mutant locus. lpd1 is highly expressed in the testis and weakly expressed in the liver of 2-week old mice. Human LPDL cDNA was subsequently cloned, and was found to encode a 460AA protein with 71% protein sequence identity to mouse lpd1 and ~35% identity to other known lipases. We next sequenced the human LPDL gene exons in hypertriglyceridemic subjects and normal controls, and identified seven SNPs within the gene exons and six SNPs in the adjacent introns. Two hypertriglyceridemic subjects were heterozygous for a rare DNA variant, namely 164G > A (C55Y), which was absent from 600 normal chromosomes. Two other coding SNPs were associated with variation in plasma HDL cholesterol in independent normolipidemic populations. Using bioinformatics, we identified another novel lipase designated LPDLR (for ‘LPDL related lipase’), which had 44% protein sequence identity with LPDL. Together with the phospholipase gene PSPLA1, LPDL and LPDLR form a new lipase gene subfamily, which is characterized by shortened lid motif. Study of this lipase subfamily may identify novel molecular mechanisms for plasma and/or tissue TG metabolism.

INTRODUCTION

Lipases hydrolyze a wide range of esterified FA species within triglyceride (TG), and are often active against other substrates, such as phospholipids (PLs). At least 20 lipases or lipase-like molecules have been given names and accession numbers in OMIM (www.ncbi.nlm.nih.gov/entrez/query). These lipases have been characterized based upon factors such as their anatomical distribution, localization intra- or extra-cellularly, substrate specificity, and or homology with other lipases (1). For instance, lipases that function within the plasma compartment, anchored to endothelium by heparan sulfate proteoglycans, include, in order from most-to-least-potent TG lipase activity, and least-to-most-potent PL lipase activity, lipoprotein

*To whom correspondence should be addressed at: Department of Experimental Therapeutics, Toronto General Research Institute, Toronto General Hospital, Room 410, 67 College Street, Toronto, Ontario, Canada M2M 1E1. Tel: +1 4163403713; Fax: +1 4163403453; Email: x.wen@utoronto.ca

Present address: The University of Hong Kong, Pokfulam Road, Hong Kong.

Human Molecular Genetics, 2003, Vol. 12, No. 10 DOI: 10.1093/hmg/ddg124

Human Molecular Genetics, Vol. 12, No. 10 © Oxford University Press 2003; all rights reserved
lipase (LPL), hepatic lipase (HL), and endothelial lipase (EL) (2–5). Other lipases are non-secreted and have predominantly intracellular hydrolytic activity, such as hormone sensitive lipase (HSL) and lysosomal acid lipase (LAL) (6,7). The activity of other lipases is extracorporeal, such as that of pancreatic lipase (PNLIP) within the intestine.

Naturally occurring loss-of-function mutations in LPL cause chylomicronemia (8,9); some LPL single-nucleotide polymorphisms (SNPs) are fairly consistently associated with metabolic and cardiovascular phenotypes (10) and LPL knock-out and transgenic mice have instructive phenotypes involving the expected alterations in plasma TG and HDL (11). Similarly, naturally occurring loss-of-function mutations in HL cause a complex hyperlipidemia with early atherosclerosis (12); some HL SNPs, especially −514C>T, are consistently associated with metabolic and cardiovascular disease phenotypes (9,13), and HL knock-out and transgenic mice have instructive phenotypes that reflect the human phenotypes (14). In contrast, neither naturally occurring human mutations, nor induced murine mutations in EL have yet been reported, although several common SNPs have been associated with variation in plasma concentrations of HDL cholesterol (15). Mild to moderate hypertriglyceridemia, we hypothesized that the murine phenotype, we next confirmed that the transgene insertion in the lpdl locus disrupted the lpdl lipase gene. We mapped the transgene junction clones relative to the gene structure of the mouse lpdl gene. One junction clone (D3) was mapped before mouse exon 10 while the other junction clone (3A) mapped after exon 10 (Fig. 2), indicating that exon 10 of the lpdl gene was deleted in the mutant lpdl locus. It is perhaps of interest that ~7 kb upstream of the lpdl gene, there were five conserved peaks (with >75% identity) designated as conserved nucleotide sequences (CNS), which may represent another gene (Fig. 2). A

RESULTS

Sequencing of the lpdl locus and bioinformatic identification of a novel lipase (lpdl)

Because mutant lpd homozygotes had hepatic steatosis and hypertriglyceridemia, we hypothesized that the murine lpd locus would encode a TG lipase. Since analysis of the junction sequences of the lpdl transgene insertion locus did not yield any lipase-related sequences and the mouse genetic sequencing database was not available at that time, we chose to clone the entire wild-type lpdl locus with bacteria artificial chromosomes (BACs) that encompassed the affected region. By screening a wild-type mouse genomic BAC library with two probes that flanked each side of the transgene, three BAC clones containing the lpdl locus were identified. Using shotgun strategy, we sequenced one BAC clone (BAC no. O16) with ~3-fold redundancy. Seven hundred and nineteen sequences for a total of 449 373 bp were randomly sequenced. These sequences formed 93 contigs for a total contig length of 163 372 bp. We used the BLASTX engine (www.ncbi.nlm.nih.gov) to interrogate all non-redundant GenBank sequences with the 93 contigs and identified a fragment in contig no. 6 that had significant homology to a portion of both human and rat phosphatidylserine-specific phospholipase A1 (PS-PLA1) with ~50% identity at the protein sequence level.

Following the identification of this lipase-related sequence, bioinformatic gene prediction tools were used to identify five putative exons from four contigs (nos 6, 28, 86 and 98), which translated as a continuous putative protein sequence of 261 amino acid (designated as lpdl lipase) that includes the highly conserved Gly-Met-Ser-Leu-Gly lipase consensus sequences (Fig. 1).

Genomic structure and expression of human/mouse LPDL/lpdl and genetic disruption of the lpdl gene in the lpd insertion locus

Using the mouse lpdl exon sequences to BLAST, we identified a genomic sequence of 340 kb (AP001660) on human chromosome 21q with significant homology to the mouse lpdl gene. Ten DNA fragments from this genomic sequence were further characterized as exons of the human LPDL gene. The exon–intron boundaries were determined using a combination of analysis with exon/intron consensus sequences, bioinformatic gene prediction tools and alignment with the cloned human cDNA sequences (Table 1). The exon sizes of human LPDL gene range from 90 to 386 bp and they span a genomic region >100 kb. The largest intron, intron 9, spans 35 kb and the smallest intron, intron 5, spans ~1 kb (Table 1). Start and stop codons are located in exons 1 and 10, respectively, and lipase consensus sequence G × S × G is encoded from exon 3. Exons 4, 5 and 6 span the most conserved lipase regions, including the lid sequences and two of three active residues within the triad structure for catalytic activity.

Using 5′ RACE, sequencing, searching mouse genetic databases and bioinformatic gene prediction, we identified a continuous mouse genomic fragment of 110 kb, which contained all 10 putative exons of the mouse lpdl gene. When aligning the mouse lpdl sequences with the human genomic LPDL sequences, nine of 10 exons were within the conserved peaks with >75% sequence identity (Fig. 2). Within the ~5 kb region before exon 1, there was another cluster of peaks with sequence identity of 50–75% which may represent the conserved promoter sequences for the gene and regulatory elements.

Since the identified lipase-like gene was a logical candidate for the lpd phenotype, we next confirmed that the transgene insertion in the lpdl locus disrupted the lpdl lipase gene. We mapped the transgene junction clones relative to the gene structure of the mouse lpdl gene. One junction clone (D3) was mapped before mouse exon 10 while the other junction clone (3A) mapped after exon 10 (Fig. 2), indicating that exon 10 of the lpdl gene was deleted in the mutant lpdl locus. It is perhaps of interest that ~7 kb upstream of the lpdl gene, there were five conserved peaks (with >75% identity) designated as conserved nucleotide sequences (CNS), which may represent another gene (Fig. 2). A
TBLASTX search against GenBank with the CNS sequences identified a novel putative protein with high sequencing homology to human putative RNA-binding protein 11 (P57052) (data not shown).

With the availability of murine lpdl exon sequences, we next sought to identify mouse and human ESTs in the published genetic databases, however no ESTs with significant homology were found. Northern blot analysis was then performed on mouse multitissue blots using probes generated from predicted exon sequences of the mouse lpdl genomic BAC clone. We detected a ~2 kb band in testis RNA but not in any other adult mouse tissues examined, including heart, brain, spleen, lung, liver, skeleton muscle and kidney (Fig. 3A-a). By RNA in situ hybridization, the lpdl gene was found to be highly expressed in the testis (Fig. 3B-b) and weakly expressed in the hepatocytes of 2-week-old mice (Fig. 3B-d). In the testis of adult mice, lpdl was...
expressed in the cytoplasm of primary spermatocytes but not in the matured sperm or in Leydig cells (Fig. 3B-f). Similar to murine $lpdl$ expression pattern, northern blotting showed that human $LPDL$ was expressed in testis (Fig. 3A-b) but was not detected in any other human tissues including brain, liver, heart, skeleton muscle, lung and intestines (data not shown).

**Identification of human $LPDL$ and murine $lpdlr$ ($lpdl$ related lipase) gene**

Using mouse $lpdl$ gene fragments as probes to screen a human testis large insert cDNA library (Clontech), four positive clones were identified. Sequencing of these clones revealed a cDNA of 1685 bp, with an open reading frame (ORF) of 1383 bp, a start codon (ATG) at nt 78 and a stop codon (TAG) at nt 1640. The ORF encodes a human $LPDL$ protein of 460 amino acids (Fig. 4). A hydrophobic leader sequence with a putative cleavage site after amino acid residue 15 was predicted by the SPScan program of SeqWeb Wisconsin GCG Package. The lipase consensus sequence $G/C2S/C2G$ was found with an active serine at amino acid residue 159. DNA sequence analysis suggested the existence of two additional active residues, Asp183 and His258, which are predicted to form a catalytic triad with Ser159 (19). A lipase lid sequence, which may determine substrate specificity (20), was identified between two cysteines at residues 238 and 251. Seven conserved cysteines at residues 55, 238, 251, 275, 286, 289 and 297 could participate in disulfide bridge formation (21). Other conserved cysteines include Cys11 and Cys455 (Fig. 4).

Using $lpdl$ protein sequence to BLAST-search against the translated EST database, we identified a mouse EST (BG868436, from salivary gland) which translated to another novel lipase with significant homology to $lpdl$. We named this novel lipase as $lpdl$-related lipase or $lpdlr$. Sequencing of BG868436 revealed a cDNA of 2155 bp in length with ORF starting from nucleotide 78 and stopping at nucleotide 1640. The ORF encodes a mouse $lpdlr$ protein of 451 amino acids (Fig. 4). The lipase consensus sequence $G \times S \times G$ was found with an active serine at amino acid residue 154. Alignment analysis suggested that, for the putative catalytic triad, Asp178 and Ser154 were conserved (19), but the normally conserved histidine residue within the triad was replaced by Tyr253.
A human EST (AW45952) harbored part of the LPDLR gene. Northern blot hybridization with labeled AW45952 fragment as probe demonstrated that the human LPDLR gene was highly expressed in lung, colon, prostate, kidney and testis with four different sized isoforms ranging from 2 to 4 kb (Fig. 3A-c and A-d).

(Fig. 4). A human EST (AW45952) harbored part of the LPDLR gene. Northern blot hybridization with labeled AW45952 fragment as probe demonstrated that the human LPDLR gene was highly expressed in lung, colon, prostate, kidney and testis with four different sized isoforms ranging from 2 to 4 kb (Fig. 3A-c and A-d).

With the availability of genomic and/or cDNA sequences of both human and mouse genes of LPDL and LPDLR, we...
Figure 4. Multiple sequence alignment of human LPDL and mouse LPDLR with other TG lipases and PS-PLA1. Triglyceride lipases and PS-PLA1 demonstrate significant sequence homology which includes the conserved lipase consensus sequence GxC2SxC2G, multiple cysteine residues that may be required for disulfide bridge formation (undermarked by an asterisk), and three residues of catalytic triad (undermarked by a hash). The lid sequences of hLPDL and mLPDLR are similar to hPS-PLA1 with 12 amino acids in length, which is much shorter than the other TG lipases. PS-PLA1 carries a modified phosphatidylserine-binding motif xxFxLXXXKxR which is not seen in TG lipases.
identified their chromosome locations by using the UCSC Genome Browser (http://genome.ucsc.edu). As shown, both mouse *lpdl* and *lpdlr* genes map to murine chromosome 16 residing in B1 and C3.1 region, respectively (Fig. 5A). Human *LPDL* maps to chromosome 21q11.2 (B) while *LPDLR* is located in the distal part of chromosome 3 (3q27.2) (C).

**Figure 5.** Chromosomal location and the adjacent gene map of mouse and human LPDL and LPDLR genes. Chromosomal locations of LPDLs and LPDLRs were identified by using the UCSC Genome Browser (http://genome.ucsc.edu). (A) Both mouse *lpdl* and *lpdlr* genes map to murine chromosome 16 in B1 and C3.1 region, respectively. Human *LPDL* maps to chromosome 21q11.2 (B) while *LPDLR* is located in the distal part of chromosome 3 (3q27.2) (C).
**Alignment of human LPDL and mouse LPDLR demonstrate higher sequence homology with PS-lipase family. Lipases and phospholipase PS-PLA1 sequences are compared**

**Figure 6.** The multiple sequence alignment dendrogram of members in the lipase family. Lipases and phospholipase PS-PLA1 sequences are compared by SeqWeb peptide PileUp software from Wisconsin GCG Package. Human LPDL and mouse LPDLR demonstrate higher sequence homology with PS-PLA1 and thus are clustered together and they form a subfamily.

**LPDL and LPDLR belong to a novel subfamily**

Alignment of human LPDL and mouse lpd1r protein sequences with other human lipases revealed significant structural conservation (Fig. 4). The ‘catalytic triad’, as well as the lipase consensus sequences G × S × G are conserved in all TG lipases and PS-PLA1. Among the highly conserved cysteine residues required in TG lipase for tertiary structure formation (21), seven appear to be conserved in both LPDL and lpd1r proteins (Fig. 4). Interestingly, the glycine G364 of LPDL is conserved in all TG lipases, but not in PS-PLA1 (Fig. 4). Using algorithm analysis (22), LPDL showed 36, 34, 32, 31 and 31% amino acid identity to human endothelial lipase (LIPG), PNLIP, HL, pancreatic lipase related protein 1 (LIP1) and LPL, respectively. Human LPDL protein exhibits 44% amino acid identity to mouse lpd1r protein, and 71% amino acid identity to its mouse homolog lpd1 (from partial lpd1 sequences). Interestingly, human LPDL also demonstrates relatively high (~34% amino acid) sequence identity to phospholipase PS-PLA1.

The lid domain plays a crucial role in determining lipase substrate specificity (20,23). The lid in both human LPDL and mouse lpd1r is composed of 12 amino acids, which is much shorter than those found in human PNLIP, LIP1, EDL, LPL and HL (23, 23, 19, 22 and 22 residues, respectively; Fig. 4). Interestingly, both hLPDL and mLPD1R lid sequences show higher homology to the lid of PS-PLA1, which is also 12 amino acids in length (Fig. 4). The multiple sequence alignment dendrogram shows that LPDL, LPDLR and PS-PLA1 share higher protein sequences homology and are clustered together (Fig. 6), suggesting they form a subfamily within the lipase gene family.

**LPDL SNPs and association with plasma lipoproteins**

Since the mouse lpd1 mutation had both disrupted lpd1 and high plasma TG, we considered that LPDL gene variation in humans might contribute to dyslipidemia. We addressed this hypothesis in two ways. First, from genomic DNA we directly sequenced LPDL exons of 60 non-diabetic Caucasians with moderate to severe hypertriglyceridemia (mean±SEM untreated TG 12.1±8.5 mmol/l; age 55±12 years) who had no obvious secondary cause of hyperlipidemia, and 10 matched normolipidemic Caucasian controls (untreated TG 1.1±0.3 mmol/l). The hypertriglyceridemic subjects had previously been shown to have no mutation in LPL, HL or EL. For newly discovered SNPs, allele frequencies were determined in 80 Caucasians. We found six non-transcribed and seven transcribed SNPs, including the nonsynonymous coding SNPs C55Y, G364E, E431K and D444E (Table 2). Genotype frequencies for each SNP did not deviate significantly from Hardy–Weinberg expectations in all samples. Mild to moderate pairwise linkage disequilibrium was observed for about half of the pairwise comparisons of LPDL SNP genotypes in Caucasians (data not shown). Two SNPs were further characterized in several additional samples of 80 individuals each: in African, East Indians, Chinese, Inuit and Amerindian, the frequencies for K431 were 0.57, 0.24, 0.05, 0.31 and 0.20, respectively, and the frequencies for E444 were 0.53, 0.51, 0.58, 0.69 and 0.51, respectively. Allele frequencies of coding SNPs in 186 hypertriglyceridemic Caucasian subjects (TG > 10 mmol/l) and 232 matched Caucasian controls (TG < 1 mmol/l) were compared, and none was found to be significantly different between samples. However, heterozygosity for C55Y was found only in the Caucasian hypertriglyceridemic patients (2/186 versus 0/232), suggesting that this might be a rare mutation associated with hypertriglyceridemia.

Next, we tested for associations of SNP genotypes with plasma lipoproteins in three unrelated samples using our established approach (24–26). Two independently ascertained, unrelated samples of healthy, normolipidemic Caucasians (174 and 161 individuals) and a well-characterized sample of healthy Inuit (208 subjects) (26) were studied. In ANOVA, dependent variables were the four plasma lipoprotein traits (TG, total, HDL and LDL cholesterol), appropriately transformed to give distributions that were not significantly different from normal. Correction was made for age, sex and body mass index by including these as independent covariates, along with the genotype for the coding SNPs only, assuming dominant, co-dominant and recessive models for each minor allele, as described (26). Seven significant associations were found with plasma lipoproteins (Table 3). At least one LPDL SNP genotype was associated with variation in HDL cholesterol in all three samples (Table 3). Also, LPDL SNP genotypes were associated with variation in LDL cholesterol in both Caucasian samples (Table 3).

**DISCUSSION**

Characterization of insertional mutations has been a useful strategy for identifying novel genes, since the mutated locus becomes tagged by the transgene, which provides a unique marker for cloning (27). The fortuitous observation of plasma and tissue TG disturbances in the insertional mutation of lpd1 transgenic mouse provided an important clue that the disrupted lpd1 locus on murine chromosome 16 contained a gene required...
in lipid metabolism. We report here the cloning of a novel lipase gene \( \text{lpdl} \) within the \( \text{lpd} \) locus. Mapping of the transgene insertion junction sequences with the wild-type \( \text{lpd} \) sequences revealed that exon 10 of \( \text{lpdl} \) is deleted in the \( \text{lpd} \) mutant locus.

Mutation screening of the human \( \text{LPDL} \) gene in patients with hypertriglyceridemia suggests that a rare missense mutation (C55Y) may be associated with elevated TG. Other \( \text{LPDL} \) SNPs were associated with variation in HDL and LDL cholesterol in three independent samples. Using bioinformatic tools, we also identified another novel lipase, called \( \text{LPDLR} \) (or \( \text{LIPH} \)). Based on their significant sequence homology and conservation of the functional domains, we propose that \( \text{LPDL} \), \( \text{LPDLR} \) and \( \text{PS-PLA1} \) form a gene subfamily.

The lipase consensus sequence \( \text{G/C2} \text{S/C2} \text{G} \) is conserved in most TG lipases and \( \text{PS-PLA1} \), as Ser forms a catalytic triad with His and Asp that mimics the catalytic triad of trypsin (19,23). In human \( \text{LPDL} \), sequence alignment suggests that the catalytic triad consists of Ser-159, Asp-183 and His-258, whereas in mouse \( \text{lpdlr} \), the triad consists of Ser-154, Asp-178 and Tyr-253, which replaces the histidine residue with tyrosine. The tyrosine in this functional site is also observed in mouse \( \text{lpdl} \) (data not shown). Interestingly, the \( \text{PNLIP} \) and \( \text{LIP1} \) utilize Lys(k) instead of His in this position, indicating His is less conserved than the Ser and Asp in the triad (Fig. 4).

Most TG lipases, like LPL and HL, also display some degree of phospholipase activity. The lid sequences which form a surface loop by a disulfide bridge of two cysteines are reported to determine substrate specificity (Fig. 4). While the lid of LPL confers preferential triglyceride hydrolysis, the lid of HL augments more phospholipase activity (20). \( \text{LPDL} \), \( \text{lpdlr} \) and

### Table 2. \( \text{LPDL} \) SNPs and allele frequencies

<table>
<thead>
<tr>
<th>Location</th>
<th>Amino acid</th>
<th>Nucleotide</th>
<th>Allele frequencies (Caucasian)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Non-transcribed ( \text{LPDL} ) SNPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 1</td>
<td></td>
<td>41 nt 5' to exon 2 C&gt;T</td>
<td>T: 0.14</td>
</tr>
<tr>
<td>Intron 2</td>
<td></td>
<td>74 nt 5' to exon 3 C&gt;T</td>
<td>T: 0.25</td>
</tr>
<tr>
<td>Intron 4</td>
<td></td>
<td>49 nt 3' to exon 4 A&gt;G</td>
<td>G: 0.32</td>
</tr>
<tr>
<td>Intron 5</td>
<td></td>
<td>16 nt 5' to exon 6 T&gt;C</td>
<td>C: 0.02</td>
</tr>
<tr>
<td>Intron 9</td>
<td></td>
<td>46 nt 3' to exon 9 G&gt;A</td>
<td>A: 0.35</td>
</tr>
<tr>
<td>3' to ORF</td>
<td></td>
<td>nt+146 G&gt;T</td>
<td>T: 0.18</td>
</tr>
<tr>
<td>B. Transcribed ( \text{LPDL} ) SNPs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 1</td>
<td></td>
<td>-54C&gt;T</td>
<td>T: 0.20</td>
</tr>
<tr>
<td>Exon 2</td>
<td>C55Y</td>
<td>164G&gt;A</td>
<td>Only in hypertriglyceridemic subjects</td>
</tr>
<tr>
<td>Exon 3</td>
<td>S159</td>
<td>477C&gt;T</td>
<td>T: 0.25</td>
</tr>
<tr>
<td>Exon 8</td>
<td>G364E</td>
<td>1091G&gt;A</td>
<td>A: 0.03</td>
</tr>
<tr>
<td>Exon 9</td>
<td>E431K</td>
<td>1291G&gt;A</td>
<td>A: 0.34</td>
</tr>
<tr>
<td>Exon 10</td>
<td>D444E</td>
<td>1332C&gt;A</td>
<td>A: 0.55</td>
</tr>
<tr>
<td>3' UTR</td>
<td></td>
<td>+80G&gt;A</td>
<td>A: 0.43</td>
</tr>
</tbody>
</table>

SNPs were identified by direct sequencing of \( \text{LPDL} \) exons in 60 hypertriglyceridemic Caucasian subjects. SNP allele frequencies were determined in 80 Caucasian subjects.

### Table 3. Summary of significant (\( P < 0.05 \)) quantitative lipoprotein associations with \( \text{LPDL} \) SNPs

<table>
<thead>
<tr>
<th>Sample</th>
<th>( n )</th>
<th>SNP name</th>
<th>Trait</th>
<th>Model</th>
<th>( P )-value</th>
<th>Adjusted mean ± SEM for genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian sample 1</td>
<td>174</td>
<td>477C&gt;T</td>
<td>log HDL-C</td>
<td>Dominant</td>
<td>0.019</td>
<td>C/C: 0.049 ± 0.012 mmol/l</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>root LDL-C</td>
<td>Dominant</td>
<td>0.020</td>
<td>C/T and T/T: 0.006 ± 0.015 mmol/l</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IVS2-74C&gt;T</td>
<td>log HDL-C</td>
<td>0.026</td>
<td>C/T and T/T: 1.64 ± 0.05 mmol/l</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>root LDL-C</td>
<td>Dominant</td>
<td>0.012</td>
<td>C/T and T/T: 0.007 ± 0.014 mmol/l</td>
</tr>
<tr>
<td>Caucasian sample 2</td>
<td>161</td>
<td>IVS2-74C&gt;T</td>
<td>log HDL-C</td>
<td>Dominant</td>
<td>0.019</td>
<td>C/T and T/T: 1.63 ± 0.05 mmol/l</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>root LDL-C</td>
<td>Dominant</td>
<td>0.005</td>
<td>C/T and T/T: 0.034 ± 0.016 mmol/l</td>
</tr>
<tr>
<td>Inuit</td>
<td>208</td>
<td>D444E</td>
<td>log HDL-C</td>
<td>Recessive</td>
<td>0.032</td>
<td>D/D and D/E: 0.23 ± 0.03 mmol/l</td>
</tr>
</tbody>
</table>

\( n \), number of subjects; SNP, single nucleotide polymorphism; adjusted mean, mean value adjusted for age, sex and obesity; SEM, standard error of the mean; log, natural logarithm, HDL-C, plasma high density lipoprotein cholesterol concentration; root, square root; LDL-C, plasma low density lipoprotein cholesterol concentration.
PS-PLA1 share a shortened lid structure (12 amino acids). Assuming LPDL and LPDLR are TG lipases, elements other than the lid must determine their TG substrate specificity, since PS-PLA1 was reported neither to hydrolyze TG (28) nor associate with plasma TG (29). However, PS-PLA1 contains a modified phosphatidylserine-binding peptide motif xxFxLXxxxx resembling the consensus sequence motif FxFxLxxxxKxR (30) which is not present in LPDL and lpdlr (Fig. 4).

Like LPDL, other lipases such as EL and HSL are highly expressed in the testis, which may reflect higher TG energy metabolism (31–34). Targeted disruption of HSL resulted in male sterility indicating the importance of HSL expression in testis tissue (35). Whether testicular expression of lpdl gene is important for testis function remains to be addressed by lpdlr gene knock-out or conditional knock-out study since the lpd mutation is perinatal lethal and the lpd mutant mice do not survive to mating age (16). Since the homozygous lpd mutants demonstrated fatty livers with extensive accumulation of TG (16), it is expected that lpdl gene would be expressed in the liver. Although northern blot hybridization with lpdl probe did not detect the expression signal in the liver, the more sensitive RNA in situ hybridization did detect weak expression in hepatic tissue in 2-week-old mice, supporting the hepatic phenotype in the lpdl mutant mice. By northern analysis, human LPDLR is expressed in lung, kidney, prostate, testis and colon, and ESTs of mouse lpdlr had also been identified from salivary gland and mammary gland suggesting a role in digesting exogenous dietary TG.

The finding that the exon 10 of lpdl is deleted in the lpd mutant suggests that C-terminal sequences may participate in the substrate specificity during TG hydrolysis. Besides deletion of exon 10, other genetic rearrangements could not be ruled out because of the complexity of gene mutations, especially transgene-induced mutations. For example, the most recently characterized mutation fld (fatty liver dystrophy) is characterized by a deletion of 2 kb sequences eliminating exon 2 and 3, and inversion of 40 genomic sequences plus a duplication of 0.5 kb segments in 3’UTR (36). Because lpd mutant mice also demonstrated runting and a potential neurological defect, we cannot rule out causative defects of more than one gene. An adjacent putative RNA binding protein 5’-flanking the lpdl gene is a potential candidate region contributing to the complexity of the phenotype (data not shown). Other genes in the 3’ region downstream of lpdl gene include a putative novel gene similar to spliceosome-associated protein 49 (AF317552), a nuclear receptor interacting protein 1 (Nrip1) and a nuclear protein HACSI (Samsn1, AF218085) (37). A genetic knock-out of lpdl lipase gene is in progress to clarify whether lpdl gene mutation alone could result in a phenotype as described in the lpd insertional mutation. If more than one genes are involved in the lpd mutation, the phenotype of the lpdl gene knock-out mice may be less severe and the resulting mice could be viable.

To identify potential mutations in human LPDL gene that contribute to human dyslipidemias, we conducted mutation screening of LPDL exons in the genomic DNA of Caucasian patients with moderate to severe hypertriglyceridemia. While seven coding SNPs were discovered, only one putative mutation was identified, namely C55Y, which was present only in hypertriglyceridemic subjects. C55 is an important residue in LPDL, which is predicted to participate disulfide bridge formation and in determining lipase tertiary structure (21,23). C55 is also conserved in both mouse lpdlr and human PS-PLA1 (Fig. 4). Therefore, the C55Y substitution may affect its function.

Results in three independent normolipidemic samples strengthen the case for association between LPDL SNPs and HDL cholesterol, although linkage disequilibrium with unmeasured variants at another gene remains possible. Variation in HDL cholesterol has previously been associated with SNPs in other lipases, specifically in LPL, HL and EL (10,13,15). The mechanisms underlying these associations are unknown, but LPDL appears to be a fourth lipase that is associated with variation in plasma HDL cholesterol. The absence of concomitant association of LPDL SNPs with plasma TG is compatible with observations from other experiments and model systems in which TG and HDL metabolism are uncoupled (9). The less consistent association of LPDL SNPs with LDL cholesterol will require validation in future studies. The association of LPDL SNPs with plasma HDL cholesterol rather than TG concentration in humans echoes the observations for EL, which despite strong sequence similarity to both the intravascular lipases LPL and HL was found by SNP analysis to be associated with HDL cholesterol rather than TG (15). Recent in vivo functional assessment has confirmed that EL plays a direct role in HDL metabolism, and these findings support the idea that members of the lipase family have different affinities for a range of lipoproteins (38–40). The apparently more pronounced association with TG in the mouse but HDL cholesterol in the human could also be related to species differences and the extent of functional impairment attributable to the induced murine mutation and the human SNPs, respectively. Clearly, functional assessment of LPDL and more intensive characterization of biochemical phenotypes in both species are required. Similarly, the less consistent association of LPDL SNPs with LDL cholesterol will require validation in future studies. Analyses using all LPDL SNPs described here may be helpful in future association studies with plasma lipoproteins, as advocated by some authors (41).

Since the molecular basis for moderate dyslipidemia is still unknown for the majority of subjects, identification of novel loci and genes remains and important approach to begin to fill this gap in understanding. Linkage studies (42,43) and systematic mouse mutagenesis (44) are complementary approaches to attain this goal. Characterizing LPDL and LPDLR may identify novel molecular mechanisms for plasma and/or tissue TG metabolism, perhaps leading to new therapies for dyslipidemias and atherosclerosis prevention.

**MATERIALS AND METHODS**

**BAC cloning, sequencing and sequence data analysis**

We identified BAC clones of lpd locus by screening a mouse 129 BAC library RPCI-22 with 32P-dCTP (Amersham Pharmacia Biotech) labeled probes generated from flanking sequences of the lpd transgene insertion site. Three positive clones were identified. One BAC clone, BAC no. 016,
hybridized to both flanking probes and was further sequenced. For sequencing the BAC clone, a sub-library in M13mp19 was first constructed. Briefly, BAC DNA was purified using the NucleoBond Plasmid Maxi kit (Clontech), randomly sheared by nebulization, and rendered blunt by treatment with mung bean nuclease, T4 polymerase and Klenow. Fragments from 1.5 to 3 kb were size-fractionated by agarose gel electrophoresis, ligated into Smal-cut M13mp19, and transformed into DH5α competent cells. Random clones were picked and single-stranded templates were purified with Qiagen M13 plates. Sequencing was carried out on ABI 377 and 373 automated DNA sequencers using fluorescently labeled dye-primer and dye-terminator chemistries (Amersham Pharmacia Biotech).

DNA sequencing was performed on ABI 377 and 373 automated DNA sequencers using fluorescently labeled dye-primer and dye-terminator chemistries (Amersham Pharmacia Biotech) and hybridized to 32P-dCTP labeled murine genomic DNA to monitor transfection. Mouse and human multiple tissue northern blots (nos. 7762-1 and 7766-1, Clontech, Palo Alto, CA, USA) were hybridized according to the manufacturer’s instructions using ExpressHyb solution (Clontech, Palo Alto, CA, USA). Probes were labeled with random priming using 32P-dCTP. Hybridized filters were exposed to X-ray film overnight or analyzed with a Bio-Rad phospho-imager system Molecular Imager FX using Quantity One software. RNA in situ hybridization was performed on sections as described (45). Briefly, sections of fresh-frozen mouse tissues were fixed in 4% paraformaldehyde (PFA), acetylated with 1.33% triethanolamine (Fluka), 15 mM HCl, 0.25% acetic anhydride (Fluka), and blocked using a prehybridization solution (50% formamide, 5× Denhardt’s, 250 g/ml baker’s yeast RNA, 500 g/ml herring sperm DNA). This procedure was followed by hybridization of the denatured DIG-labeled RNA anti-sense probes to complementary RNA molecules on the sections at 70°C. DIG-labeled RNA sense probes also served as experimental controls. Slides were washed and blocked with 10% heat-inactivated sheep serum (HISs) for 1 h at room temperature. DIG-labeled hybrids were detected using an anti-DIG antibody conjugated with alkaline phosphatase, anti-digoxigenin-AP Fab fragments (Roche Diagnostics). This strategy utilized a color reaction with 4-nitro blue tetrazolium chloride (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP; Roche Diagnostics) as substrates, with hybridizations appearing as dark blue. Both cytoplasm and nucleus were counterstained as red by Fast Red (DAKO, catalog no. S19863) and slides were mounted in Cytoseal 60 Mounting Medium ( Stephens Scientific, Riverdale, NJ, USA).

Bioinformatic analysis of DNA and protein sequences

BLAST searching of human GenBank of nucleotide sequences at the NCBI site (www.ncbi.nlm.nih.gov/BLAST/) with mouse lpdl sequences from BAC contigs identified a genomic sequence of 340 kb (AP001660) on human chromosome 21q which carries the human LPDL gene. Mouse lpdl genomic sequences, 110 kb, were retrieved from UCSC Genome Bioinformatic (http://genome.ucsc.edu). Comparison of 110 kb of mouse lpdl genomic sequences with orthologous human genomic sequences was performed using the program VISTA (www-gsd.lbl.gov/vista). The EST clones of mouse and human LPDLR were identified by TBLASTX against the GenBank EST database with lpdl protein sequences. Analysis of nucleotide and protein sequences of mouse and human LPDLR and other lipases was carried out using SeqWeb software (Wisconsin GCG Package). Multiple sequence alignment was performed with the SeqWeb PileUp program and shading of consensus sequence was performed with an internet BOXSHADE3.21 server (www.ch.embnet.org/software/BOX_form.html).
methods, such purification with serum alkaline phosphatase and exonuclease I, followed by ddNTP primer extension and analysis on an ABI Prism 377 DNA Sequencer (PE Biosystems, Mississauga, Ontario, Canada). For some non-
nonsynonymous SNPs, allele frequencies were determined in
additional subjects from various ethnic groups.

Estimates of pairwise linkage disequilibrium between the 12
SNP genotypes in 80 normal Caucasians were calculated as
described (46). To test for association with clinical and
quantitative traits, we aimed to select as few SNPs as possible
for genotyping, based upon the following prioritization strategy:
(1) the SNP changed the coding sequence; (2) the SNP minor
allele frequency was >0.10; (3) non-coding SNPs had high
information content (heterozygosity) and displayed strong
linkage disequilibrium with other non-coding SNPs (representing
‘haplotype blocks’). Based on this prioritization, the following
four LPDL SNPs were genotyped in all samples: E431K, D444E,
IVS2 [−74]C>T and 447C>T. Genotyping with these four SNPs
essentially accounted for genotypes for all 12 SNPs because of
linkage disequilibrium (data not shown).

We used our established approach to identify associations
with clinical phenotypes (24–26), namely transformed plasma
TG, total, LDL and HDL cholesterol in two independently
ascertained, unrelated samples of healthy, normolipidemic
Caucasians (Caucasian 1 and Caucasian 2) and a well-
characterized sample of healthy Inuit (208 subjects) (26). The
first sample of 174 Caucasians was 48.3% male and had mean
(±SEM) age 50.1±4.3 years. The second sample of 161
Caucasians was 42.0% male and had mean (±SEM) age
53.7±5.8 years. A total of eight ANOVAs was performed as
initial screening to identify significant associations: one
ANOVA for each lipoprotein trait as the dependent variable
in the Caucasian 1 and Inuit samples. Each ANOVA included
age, sex and body mass index as independent covariates. The
above four SNP genotypes were included as independent
categorical variables in each ANOVA, with significance taken
from the type III sums of squares (nominal P < 0.05). Type III
sums of squares analysis is applicable to unbalanced study
categorical variables in each ANOV A, with signi-
cance taken

REFERENCES
4. Jaye, M., Lynch, K.J., Krawiec, J., Marchadier, D., Maugeais, C., Doan, K.,
endothelial-derived lipase that modulates HDL metabolism. Nat. Genet.,
21, 424–428.
5. Jin, W., Marchadier, D. and Rader, D.J. (2002) Lipases and HDL
6. Holm, C., Kirchherr, T.G., Svenson, K.L., Fredrikson, G., Nilsson, S.,
Miller, C.G., Shively, J.E., Heinzmann, C., Sparkes, R.S. and Mohandas, T.
(1988) Hormone-sensitive lipase: sequence, expression, and chromosomal
Molecular and enzymatic analyses of lysosomal acid lipase in cholesterol
biochemical, and molecular genetic characteristics. Arterioscler. Thromb.,
13, 720–728.
insights from genetic and metabolic studies. Curr. Opin. Lipidol., 10,
259–267.
hepatic lipase and scavenger receptor BI in clearing phospholipid/free
cholesterol-rich lipoproteins in PLTP-deficient mice. Biochim. Biophys.
Acta, 1503, 133–140.
Rader, D.J. (2002) Identification of genetic variants in endothelial lipase in
persons with elevated high-density lipoprotein cholesterol. Circulation,
106, 1321–1326.
lpd (lipid defect): a novel mutation on mouse chromosome 16 associated
Murine phosphatidylserine-specific phospholipase A1 (Ps-pla1) maps to
chromosome 16 but is distinct from the lpd (lipid defect) locus.
a new member of the triglyceride lipase family synthesized by the intestine.
Genomics, 80, 268–273.
19. Emmerich, J., Beg, O.U., Peterson, J., Previato, L., Brunzell, J.D.,
Analysis of the catalytic triad by site-directed mutagenesis of Ser-132,
and lipoprotein lipase: the loop covering the catalytic site mediates lipase
Lipoprotein lipase. Molecular model based on the pancreatic lipase X-ray
structure: consequences for heparin binding and catalysis. J. Biol. Chem.,
269, 4626–4633.
and Cao, H. (2001) Polymorphisms in PNLP, encoding pancreatic lipase,
and associations with metabolic traits. J. Hum. Genet., 46, 320–324.
association between genetic variation in the CYP7 gene promoter and plasma
lipoproteins in three Canadian populations. Atherosclerosis, 154, 579–587.


