

## The N342S MYLIP polymorphism is associated with high total cholesterol and increased LDL receptor degradation in humans

Daphna Weissglas-Volkov, ... , Peter Tontonoz, Päivi Pajukanta

*J Clin Invest.* 2011;121(8):3062-3071. <https://doi.org/10.1172/JCI45504>.

Research Article

Genetics

Atherosclerotic cardiovascular disease (ASCVD) affects more than 1 in 3 American adults. Hypercholesterolemia is a major treatable risk factor for ASCVD, yet many individuals fail to reach target levels of LDL-cholesterol (LDL-C) through the use of statins and lifestyle changes. The E3 ubiquitin ligase myosin regulatory light chain–interacting protein (MYLIP; also known as IDOL) is a recently identified regulator of the LDL receptor (LDLR) pathway. Genome-wide association studies (GWASs) in populations of mixed European descent have identified noncoding variants in the *MYLIP* region as being associated with LDL-C levels, but no underlying functional variants were pinpointed. In order to fine-map actual susceptibility variants, we studied a population demographically distinct from the discovery population to ensure a different pattern of linkage disequilibrium. Our analysis revealed that in a Mexican population, the nonsynonymous SNP rs9370867, which encodes the N342S amino acid substitution, is an underlying functional variant that was associated with high total cholesterol and accounted for one of the previous significant GWAS signals. Functional characterization showed that the Asn-encoding allele was associated with more potent LDLR degradation and decreased LDL uptake. Mutagenesis of residue 342 failed to affect intrinsic MYLIP E3 ligase activity, but it was critical for LDLR targeting. Our findings suggest that modulation of MYLIP activity can affect LDL-C levels and that pharmacologic inhibition of MYLIP activity might be [...]

Find the latest version:

<https://jci.me/45504/pdf>





# The N342S MYLIP polymorphism is associated with high total cholesterol and increased LDL receptor degradation in humans

Daphna Weissglas-Volkov,<sup>1</sup> Anna C. Calkin,<sup>2,3</sup> Teresa Tusie-Luna,<sup>4</sup> Janet S. Sinsheimer,<sup>1</sup> Noam Zelcer,<sup>3</sup> Laura Riba,<sup>4</sup> Ana Maria Vargas Tino,<sup>5</sup> Maria Luisa Ordoñez-Sánchez,<sup>4</sup> Ivette Cruz-Bautista,<sup>5</sup> Carlos A. Aguilar-Salinas,<sup>5</sup> Peter Tontonoz,<sup>2,3</sup> and Päivi Pajukanta<sup>1</sup>

<sup>1</sup>Department of Human Genetics, <sup>2</sup>Howard Hughes Medical Institute, and <sup>3</sup>Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, California, USA. <sup>4</sup>Molecular Biology and Genomic Medicine Unit, Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubiran (INCMNSZ), and Instituto de Investigaciones Biomédicas de la UNAM, Mexico City, Mexico. <sup>5</sup>Department of Endocrinology and Metabolism, Instituto Nacional de Ciencias Médicas y Nutrición, Salvador Zubiran (INCMNSZ), Mexico City, Mexico.

**Atherosclerotic cardiovascular disease (ASCVD) affects more than 1 in 3 American adults. Hypercholesterolemia is a major treatable risk factor for ASCVD, yet many individuals fail to reach target levels of LDL-cholesterol (LDL-C) through the use of statins and lifestyle changes. The E3 ubiquitin ligase myosin regulatory light chain-interacting protein (MYLIP; also known as IDOL) is a recently identified regulator of the LDL receptor (LDLR) pathway. Genome-wide association studies (GWASs) in populations of mixed European descent have identified noncoding variants in the MYLIP region as being associated with LDL-C levels, but no underlying functional variants were pinpointed. In order to fine-map actual susceptibility variants, we studied a population demographically distinct from the discovery population to ensure a different pattern of linkage disequilibrium. Our analysis revealed that in a Mexican population, the nonsynonymous SNP rs9370867, which encodes the N342S amino acid substitution, is an underlying functional variant that was associated with high total cholesterol and accounted for one of the previous significant GWAS signals. Functional characterization showed that the Asn-encoding allele was associated with more potent LDLR degradation and decreased LDL uptake. Mutagenesis of residue 342 failed to affect intrinsic MYLIP E3 ligase activity, but it was critical for LDLR targeting. Our findings suggest that modulation of MYLIP activity can affect LDL-C levels and that pharmacologic inhibition of MYLIP activity might be a useful strategy in the treatment of dyslipidemia and ASCVD.**

## Introduction

Atherosclerotic cardiovascular disease (ASCVD) affects more than 1 in 3 American adults, and the burden arising from this in terms of mortality and financial cost is great (1). Modifiable risk factors for ASCVD include hypercholesterolemia, hypertension, type 2 diabetes mellitus, obesity, and smoking (2). Statin drugs, which lower plasma LDL-cholesterol (LDL-C) levels via inhibition of the rate-limiting enzyme in cholesterol production, 3-hydroxy-3-methylglutaryl-CoA reductase, have a substantial impact on ASCVD incidence and mortality. Despite efforts to modify lifestyle and the availability of statins, however, the prevalence of ASCVD continues to increase (1). Many individuals are unable to reach their target lipid levels on statins alone or experience unwanted side effects (3). Thus, there is a need to identify novel targets that might permit further lowering of LDL-C and reduction of ASCVD risk.

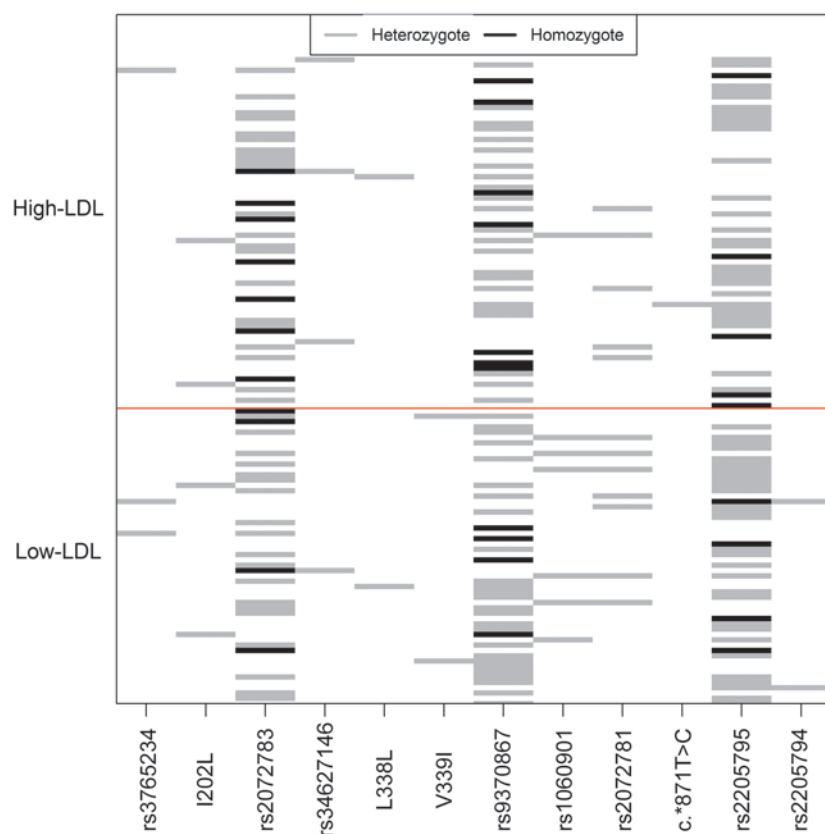
We previously identified the E3 ubiquitin ligase myosin regulatory light chain-interacting protein (MYLIP; also known as IDOL) as a regulator of the LDL receptor (LDLR) pathway for cellular cholesterol uptake (4). MYLIP is a direct target for regulation by the nuclear receptor liver X receptor (LXR), and its expression is coordinately

regulated with the ATP-binding cassette transporter in multiple cell types (4). In response to cellular cholesterol loading, activation of LXR leads to rapid induction of MYLIP expression. MYLIP stimulates ubiquitination of the LDLR on its cytoplasmic tail, thereby directing its degradation. Expression of *Myliip* in mouse liver by means of an adenoviral vector dramatically reduces LDLR protein levels and raises LDL-C levels (4). Thus, the LXR-MYLIP-LDLR pathway provides a complementary pathway to sterol regulatory element-binding proteins for the feedback inhibition of cholesterol uptake (4, 5).

The LXR-MYLIP pathway is conserved in human cells, but its role in human cholesterol metabolism remains to be addressed. Knock-down of MYLIP expression in human cultured cells increases LDLR protein levels (4); however, the effect of loss of MYLIP expression *in vivo* has not been investigated. Furthermore, given that MYLIP is expressed in many tissues and cell types, it is difficult to predict how gain- or loss-of-function mutations might affect plasma lipid levels or susceptibility to ASCVD. Recent genome-wide association studies (GWASs) in cohorts of mixed European descent have identified noncoding SNPs in the region of the MYLIP gene that are associated with LDL-C levels (6–8), thus demonstrating for the first time genetic association between the MYLIP gene and lipid levels in humans. However, the underlying causal variants and the mechanistic basis of these associations have not been established.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Citation for this article:** *J Clin Invest.* 2011;121(8):3062–3071. doi:10.1172/JCI45504.

**Figure 1**

Sequencing results of the *MYLIP* gene. Mexican subjects with high and low LDL-C (see Supplemental Table 1 for clinical characteristics) were sequenced for the coding region and exon-intron boundaries of the *MYLIP* gene. Each line represents the genotype of an individual; each column represents an identified sequencing variant. Individuals are ranked according to their LDL-C levels; low-LDL individuals fall below the red line, and high-LDL individuals above. Gray, heterozygote genotypes; black, homozygote rare genotypes; white, homozygote common genotypes. See Supplemental Table 2 for major and minor alleles and positions of the variants.

As yet, the function of most variants identified in GWASs is unknown, and most of these variants reside outside of coding sequence (9). As patterns of linkage disequilibrium (LD) vary among populations with different demographic histories (10), studies in a different population than the discovery population, such as Mexican subjects, can assist in fine-mapping the actual susceptibility variants (i.e., trans-ethnic fine-mapping). Furthermore, thus far Mexicans have not been included in any of the GWASs for lipids, although several epidemiological studies have clearly demonstrated that the Mexican population has an increased predisposition to combined hyperlipidemia and ASCVD (11). Therefore, it is important to investigate the loci conferring to dyslipidemias in Mexicans.

Here, we investigated the *MYLIP* region in the Mexican population in order to fine-map the actual susceptibility variants. We identified a nonsynonymous SNP, rs9370867 (which encodes the N342S amino acid substitution), as the underlying functional variant accounting for one of the previous genome-wide significant signals. We demonstrated that the N342 amino acid substitution altered the ability of *MYLIP* to degrade the LDLR and that the Asn-encoding allele (N342) was associated with higher total cholesterol (TC) levels in Mexican dyslipidemic individuals. These results suggest that modulation of *MYLIP* activity can affect human plasma cholesterol and support further investigation into whether pharmacologic inhibition of *MYLIP* activity might be a useful strategy in the treatment of dyslipidemia and ASCVD.

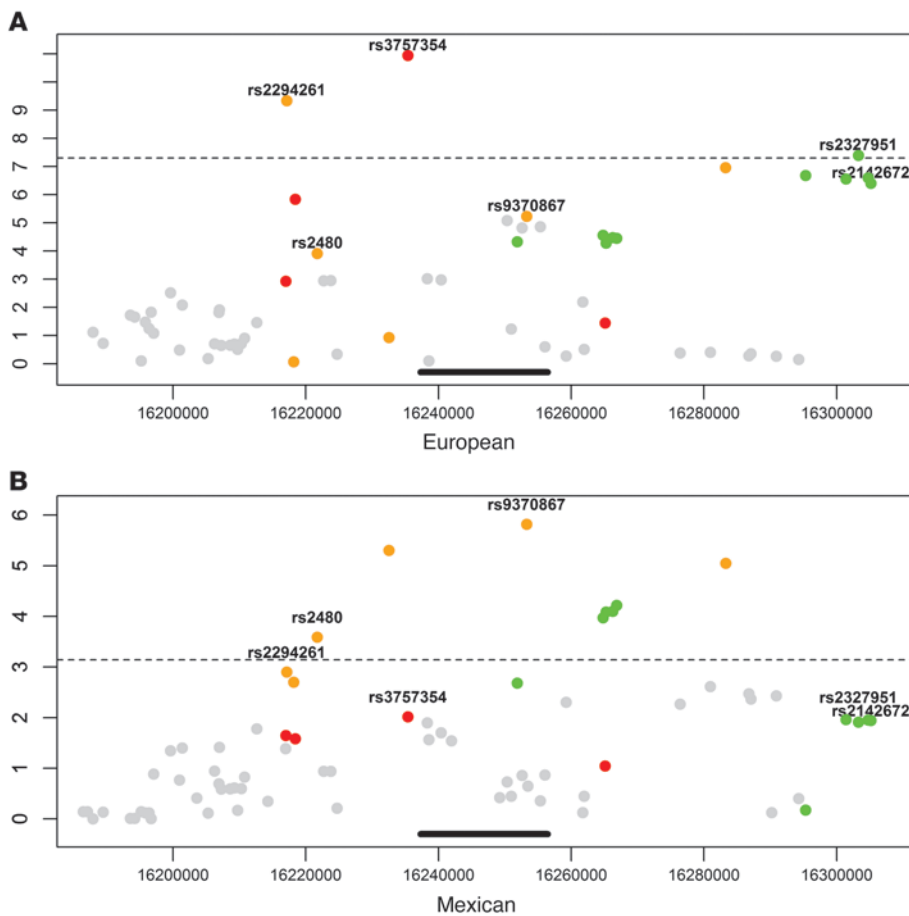
## Results

**Resequencing of the *MYLIP* gene.** To identify DNA sequence variants that influence *MYLIP* activity and to screen for population-specific variants, we sequenced all exons and exon-intron borders of

the *MYLIP* gene in Mexican individuals from each extreme of the LDL-C distribution. This discovery study sample consisted of 56 subjects with low LDL-C ( $\leq 10$ th age/sex-specific Mexican population percentile) and 66 subjects with high LDL-C ( $\geq 90$ th age/sex-specific Mexican population percentile). The clinical characteristics of the discovery study sample and the distribution of serum LDL-C levels are shown in Supplemental Table 1 and Supplemental Figure 1, A and B (supplemental material available online with this article; doi:10.1172/JCI45504DS1).

By sequencing, we identified 12 DNA variants, of which 6 were noncoding (i.e., intron boundary or untranslated regions [UTRs]): 3 variants were synonymous, and 3 nonsynonymous (Figure 1 and Supplemental Table 2). Of the 3 nonsynonymous variants, we believe I202L and V339I to be novel and rare (minor allele frequency [MAF]  $\sim 1\%$ ), and N342S is common (MAF 25%) and known in public databases as rs9370867. The I202L variant was observed in both low- and high-LDL groups of the discovery study sample, whereas the V339I variant was unique to the low-LDL group (Figure 1 and Supplemental Table 2) and therefore investigated in additional study samples. However, all 3 nonsynonymous variants were further investigated in functional analyses. Of the remaining variants, the 3' UTR variant rs2205794 and the synonymous variant rs1060901 were also investigated in additional subjects, because rs2205794 was unique to 1 tail of the LDL-C distribution and rs1060901 was observed in 6 subjects of the low-LDL group versus 1 subject of the high-LDL group (Figure 1 and Supplemental Table 2).

The low-frequency SNPs rs1060901 and rs2205794 (MAF 1%–5%) were genotyped in additional 450 normotriglyceridemic subjects, and the rare variant V339I (MAF  $< 1\%$ ) was genotyped in 1,100 normotriglyceridemic subjects. The rationale for this genotyping

**Figure 2**

Association results in the *MYLIP* gene region ( $\pm 50$  kb) compared between European and Mexican ethnicities. **(A)** GWAS meta-analysis results of subjects of mixed European descent (6). The colored circles represent  $-\log_{10}$  of the *P* values for LDL-C levels; dashed line indicates the genome-wide significance threshold ( $\log_{10}[5 \times 10^{-8}]$ ). Only SNPs that were also investigated in the Mexican population are shown. **(B)** Association results of the Mexican dyslipidemic study sample. Colored circles represent  $-\log_{10}$  of the *P* values for the high TC affection status. Red circles indicate SNPs in LD ( $r^2 \geq 0.5$ ) with rs3757354; orange circles indicates SNPs in LD with rs9370867; green circles indicate SNPs in LD with rs2327951 in the Mexican controls. The location of the SNPs is shown on the x axis in relation to the physical location of the *MYLIP* gene on chromosome 6.

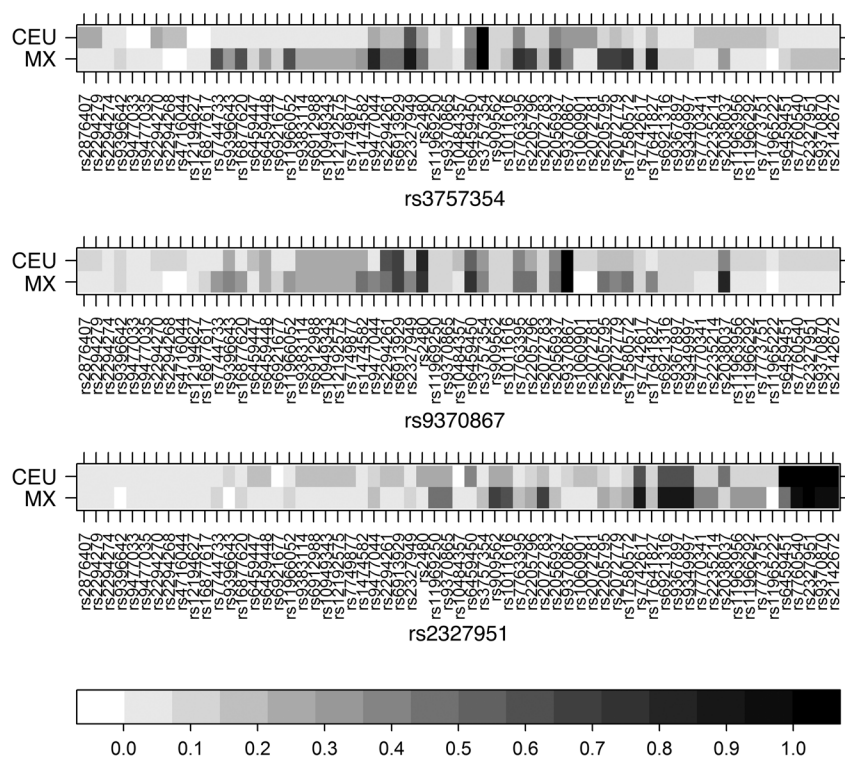
strategy was to achieve a high probability (>98%) of seeing at least 5 heterozygous individuals. In these larger study samples, all 3 variants appeared in similar frequencies in subjects with decreased (<100 mg/dl) and elevated (>130 mg/dl) LDL-C (Supplemental Table 3), which suggests that the initial findings resulted from the smaller size of the discovery sample. Furthermore, it is worth noting that we did not detect any rare coding mutations in the Mexican familial hypercholesterolemia (FH) patients or in dyslipidemic probands with LDL-C levels above the 90th age/sex-specific Mexican percentiles. Thus, rare coding variation in *MYLIP* did not account for FH in these Mexican subjects with unknown etiology of FH (12).

**Association analysis of the *MYLIP* gene.** To identify common variants that influence *MYLIP* activity and contribute to variation in plasma cholesterol levels, we investigated the *MYLIP* gene region (chr6, 16187296–16306457) in a Mexican dyslipidemic study sample ( $n = 2,310$ ) (Supplemental Table 1). We tested a total of 34 directly genotyped and 38 imputed SNPs surrounding the *MYLIP* gene ( $\pm 50$  kb) for association with TC and non-HDL-C levels in the Mexican dyslipidemic study sample. The SNPs were imputed with high certainty (>0.9), using HapMap MEX (Mexican-American) data as a reference panel. Together, the 72 analyzed SNPs tagged and captured 100% of the genetic variation (MAF  $\geq 1\%$ ;  $r^2 > 0.8$ ) in the HapMap MEX samples. We examined the TC and non-HDL-C traits rather than LDL-C, because LDL-C values cannot be reliably calculated using the Friedewald equation if triglyceride (TG) levels exceed 400 mg/dl. However, because the correlations between the LDL-C concentration and the TC and non-

HDL-C concentrations were greater than 0.9 ( $r = 0.93$  and  $r = 0.98$ , respectively) in this Mexican study sample, these traits were therefore excellent surrogates for LDL-C levels. TC and non-HDL-C levels were normally distributed and adjusted for age, sex, and hypertriglyceridemia affection status (Supplemental Figure 1, C and D). Furthermore, we also analyzed the dichotomized TC trait (see Methods), since the power of the quantitative trait analysis may be somewhat reduced by the focus on hypertriglyceridemia case-control status.

In the Mexican dyslipidemic study sample, the common nonsynonymous rs9370867 variant that we also observed by resequencing was the SNP most significantly associated with high TC and non-HDL-C levels for the A allele (MAF 25%), encoding Asn ( $P = 1.5 \times 10^{-6}$  to  $1.81 \times 10^{-3}$ ; Figure 2B and Supplemental Table 4). This association was gene-wide significant, as it surpassed the multiple-testing correction threshold for 70 tests (see Methods). However, as we tested 2 correlated traits, the Bonferroni correction for multiple testing (which presumes that tests are independent) is conservative for our analyses. Additionally, 8 other SNPs passed the multiple-testing threshold (Supplemental Table 4). These 8 SNPs belonged to 3 different LD groups ( $r^2 \leq 0.5$  between groups,  $r^2 \geq 0.7$  within groups; Figure 3). To further evaluate whether these SNPs represent independent association signals from the associated nonsynonymous rs9370867 variant, we performed a conditional association analysis while including rs9370867 genotype counts as a covariate. None of the other SNPs remained gene-wide significant in these conditional analyses (Supplemental Table 4).



**Figure 3**

Pairwise LD of the SNPs rs3757354, rs9370867, and rs2327951 with all other SNPs in the *MYLIP* gene region ( $\pm 50$  kb) in the European and Mexican populations. Shown is pairwise LD in  $r^2$  in 120 Mexican controls (MX) and in the CEU sample. SNPs shown on the x axis are ordered according to their bp position.

To further validate the association of the most strongly associated variant, rs9370867, we extended the genotyping of the SNP to 512 additional Mexican dyslipidemic samples available for study (total  $n = 2,822$ ), which further strengthened the association signal for the A risk allele (N342) with both higher TC and non-HDL-C residual values ( $P = 2.2 \times 10^{-05}$  and  $P = 2.5 \times 10^{-04}$ ), as well as with high-TC affection status ( $P = 7 \times 10^{-07}$ ; Tables 1 and 2). The adjusted means of TC by genotype were 5.49 mmol/l for the Asn/Asn homozygotes, 5.31 mmol/l for the Asn/Ser heterozygotes, and 5.20 mmol/l for the Ser/Ser homozygotes (Table 2).

Importantly, we also tested whether the N342 allele was significantly associated with the simultaneous elevation of both TC and TG levels, because combined hyperlipidemia is highly common among Mexican individuals and carries the highest risk to develop ASCVD in this population (13). We obtained a  $P$  value of  $8.1 \times 10^{-05}$  for the combined hyperlipidemia status (Table 1). In this analysis, we did not adjust for the hypertriglyceridemia affection status; rather, we tested whether rs9370867 also contributes to high TC in the presence of elevated TG levels.

The frequency of the Asn (risk) allele varies considerably among populations (0%–60%) (10, 14). In African (YRI), Asian (CHB and CHD), and Amerindian (Mayan) ancestries, the frequency is relatively low at 2%–8%, whereas in European (CEU; Northern and Western European ancestry from the CEPH collection), Sardinian, and Basque ancestries, the frequency is much higher at 49%–60% (10, 14). The Asn allele is considered the reference allele and the Ser (protective) allele the variant in public databases.

However, Ser at position 342 is the ancestral allele in cross-species comparisons.

To reduce the possibility of spurious associations caused by the admixed Mexican ancestry, especially when the allele frequencies of rs9370867 are known to vary among populations, we also performed association analyses while including individual ancestry (IA) estimates as a covariate in the regression model. We obtained  $P$  values of  $5.0 \times 10^{-04}$  and  $2.2 \times 10^{-03}$  for continuous TC and non-HDL-C levels, respectively, and  $3.8 \times 10^{-05}$  for the high-TC affection status in these adjusted analyses, which suggests that the associations were not confounded by population admixture.

*Comparison with previous GWAS results.* Recently, the *MYLIP* gene region has been implicated for LDL-C levels in 2 large-scale GWASs (7, 8) and in a meta-analysis of these studies and other GWASs of populations of European descent (6). Common variants (MAF >20%) from 3 different LD groups (pairwise LD in  $r^2 < 0.3$ ) in the *MYLIP* gene region (rs3757354, rs2294261/rs2480 [ $r^2 = 0.75$ ], and rs2327951/rs2142672 [ $r^2 = 1$ ]) surpassed the genome-wide significant level ( $P \leq 5 \times 10^{-08}$ ; Figure 2A and refs. 6–8). All of these GWAS SNPs resided outside of the *MYLIP* gene (Figure 2), and it is likely that these GWAS SNPs do not represent the actual functional variants, but rather are in LD with the underlying causal variants. Studies in diverse ethnicities, such as Mexican

individuals, may assist in fine-mapping the actual susceptibility variants (i.e., trans-ethnic fine-mapping), as allele frequencies and patterns of LD vary across populations (10).

We examined the association results in the Mexican dyslipidemic study sample in relation to the *MYLIP*-associated region (chr6, 16187296–16306457) from the mixed European descent GWASs in order to restrict the associated region using trans-ethnic LD comparisons and identify the actual disease-causing variants in the *MYLIP* gene region. First, using the HapMap MEX and CEU samples, we indeed observed considerable differences in allele frequencies and LD in the *MYLIP* gene region between Europeans and Mexicans (Supplemental Figure 2), which suggests that trans-ethnic association analysis may help fine-map the GWAS signals (15). Furthermore, the 72 SNPs we analyzed tagged and captured 100% and 94% of the genetic variation with MAF 1% or greater and  $r^2 > 0.8$  in the HapMap MEX and CEU samples, respectively.

Variants from the 3 associated LD groups were gene-wide significant in the Mexican dyslipidemic study sample (Figure 2 and Supplemental Table 4), thus extending the GWAS signals of the mixed European descent population to the Mexican population. Importantly, our LD analyses also demonstrated that the most strongly associated variant in Mexicans, rs9370867, was in strong LD with the GWAS associated variants rs2480 (7) and rs2294261 (6) in the CEU founders with European ancestry and, to a lesser extent, in the Mexican controls ( $r^2 = 0.73$  and  $r^2 = 0.49$ , respectively; Figure 3). Accordingly, in Mexicans, the SNPs rs2480 and rs2294261 resulted in less significant associations than the variant rs9370867 (Figure 2B



**Table 1**  
Association of rs9370867 in the Mexican study sample

	N	Association frequency		Effect (SE) <sup>A</sup>	P	Adjusted P <sup>B</sup>
		High TC	Normal TC			
TC status	2,661	31.40%	23.60%	1.49 (0.08)	7.02 × 10 <sup>-07</sup>	3.84 × 10 <sup>-05</sup>
Combined hyperlipidemia	1,656	31.65%	24.08%	1.44 (0.09)	8.13 × 10 <sup>-05</sup>	1.50 × 10 <sup>-03</sup>

<sup>A</sup>Effect size represents the odds ratio of each copy of the risk allele for the TC affection status and the proportion of 1 SD change in standardized TC residual values for each copy of the risk allele. <sup>B</sup>P value adjusted for admixture ancestry using IA estimates as a covariate in the regression analysis.

and Supplemental Table 4). We also examined the direct association evidence for rs9370867 in the large GWAS meta-analysis of cohorts of mixed European descent (6): rs9370867 was associated with LDL-C levels for the same A risk allele (MAF 50%) as in Mexicans ( $P = 5.8 \times 10^{-06}$ ; Figure 2A). Furthermore, the association signal for rs2480 in this GWAS meta-analysis was lower than the one for rs9370867 using the same sample size ( $P = 1.2 \times 10^{-04}$ ;  $n = 34,733$ ), and only about one-third of the total sample size was available for the analysis of rs9370867 (34,733 of 100,000) versus 95,443 available for rs2294261. We statistically estimated based on the Z score of rs9370867 that given a sample size of 95,443 subjects, the P value of rs9370867 would also have surpassed the genome-wide significance level (estimated  $P = 5.9 \times 10^{-14}$ ), as did the P values of rs2294261 ( $P = 4.6 \times 10^{-10}$ ), rs3757354 ( $P = 1.16 \times 10^{-11}$ ), and rs2327951 ( $P = 4.06 \times 10^{-08}$ ) from the other significant LD groups in the actual GWAS meta-analysis that indeed included more than 90,000 subjects (6). Hence, this calculation demonstrated that the sample size of rs9370867 may not have been sufficiently powered to meet the genome-wide significance level in the meta-analysis GWAS (6).

Because the frequency of the A allele and the pattern of pairwise LD of rs9370867 were considerably different between Europeans and Mexicans (50% versus 25%; ref. 10 and Figure 3), the consistency of these genetic association signals suggests that this nonsynonymous variant may be the actual susceptibility variant as opposed to being in strong LD with the causal variant. Furthermore, all the SNPs that were in LD ( $r^2 \geq 0.5$ ) with rs9370867 in the Mexican controls (i.e., rs2480, rs2294261, rs6913929, rs6459450, and rs2038037) reside in nonconserved intergenic regions. Their regulatory potential scores computed based on alignments of 7 species (7X Reg Potential; UCSC Genome Browser; ref. 16) were below 0.1, the default limit, indicating resemblance to regulatory elements, whereas the score for rs9370867 exceeded the limit at 0.4. Most importantly, the strength of their association was in full accordance with the strength of their pairwise LD with rs9370867 in Mexicans (Supplemental Table 4). Taken together, these trans-

ethnic LD and association data suggest that the nonsynonymous SNP rs9370867 is more likely the true underlying functional variant accounting for the genome-wide significant signals rs2480 and rs2294261 observed in the previous GWASs (6, 7).

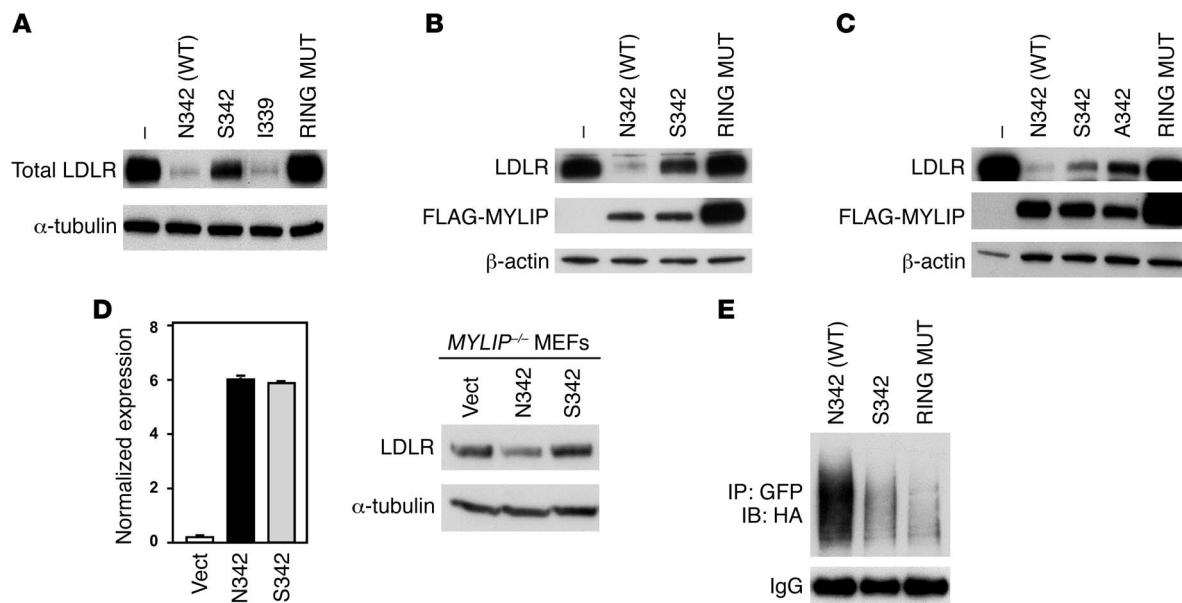
The GWAS variants of the other 2 associated LD groups, rs3757354 (6) and the SNPs rs2327951 (6) and rs2142672 (8), which are in full LD ( $r^2 \geq 0.97$ ) in both Europeans and Mexicans, were nominally associated with high TC in the Mexican dyslipidemic study sample for the same risks allele as in Europeans (Figure 2B and Supplemental Table 4). However, a cluster of 4 SNPs – rs7742617, rs6921316, rs9367897, and rs9349997, which were in strong LD with rs2327951 in the Mexican controls and, to a lesser extent, in the CEU founders of HapMap ( $r^2 = 0.87$  and  $0.58$ , respectively; Figure 3) – was substantially more strongly associated than rs2327951 with higher TC and non-HDL-C levels in Mexicans ( $P = 6.0 \times 10^{-05}$  to  $8.3 \times 10^{-03}$ ; Figure 2B and Supplemental Table 4). These SNPs reside closer to the MYLIP gene, 8–30 kb away, as opposed to rs2327951 and rs2142672, which are 48–49 kb downstream of MYLIP (Figure 2). Likewise, rs2327949, which was in stronger LD with rs3757354 in the Mexican controls than in the CEU founders of HapMap ( $r^2 = 0.87$  and  $0.60$ , respectively; Figure 3), was slightly more significant than rs3757354 with continuous TC and non-HDL-C levels in the Mexican dyslipidemic study sample ( $P = 6.5 \times 10^{-04}$  and  $P = 3.0 \times 10^{-03}$ , respectively; Supplemental Table 4). Taken together, these trans-ethnic LD data and association data in the Mexican dyslipidemic study sample support the associations of the other GWAS-associated LD groups as well, although overall the strongest signal for the MYLIP region was observed with rs9370867 in Mexicans.

*GWAS variants and gene expression analysis in adipose tissue.* The GWAS variants or SNPs in LD with these variants might influence plasma lipid concentrations by modifying regulatory elements. Therefore, we used cis-expression quantitative trait locus (cis-eQTL) analysis to determine whether the SNPs rs3757354, rs2142672 ( $r^2 \geq 0.97$  with rs2327951 in both Mexicans and Europeans), rs9367897, rs2480, and rs9370867 influence the expression levels of MYLIP

**Table 2**  
TC and non-HDL-C values by genotype in the Mexican study sample

	Genotype			Effect (SE) <sup>A</sup>	P	Adjusted P <sup>B</sup>
	Asn/Asn	Asn/Ser	Ser/Ser			
TC (mmol/l)	5.49 ± 0.07	5.31 ± 0.03	5.20 ± 0.02	0.13 (0.03)	2.18 × 10 <sup>-05</sup>	5.03 × 10 <sup>-04</sup>
Non-HDL-C (mmol/l)	4.30 ± 0.06	4.13 ± 0.03	4.05 ± 0.02	0.11 (0.03)	2.50 × 10 <sup>-04</sup>	2.23 × 10 <sup>-03</sup>

Shown are marginal mean ± SEM evaluated at the average age, sex, and hypertriglyceridemia affection status ( $N = 2,618$ ). <sup>A</sup>Effect size represents the proportion of 1 SD change in standardized values for each copy of the risk allele. <sup>B</sup>P value adjusted for admixture ancestry using IA estimates as a covariate in the regression analysis.

**Figure 4**

MYLIP S342 is associated with reduced ubiquitination and degradation of the LDLR and increased LDL uptake compared with N342. (A) Immunoblot analysis of HEK293T whole-cell lysates after cotransfection with LDLR and WT MYLIP (N342/V339), S342, I339, or C387A ring domain mutant (RING MUT) expression plasmids. (B) Immunoblot of HEK293T whole-cell lysates after cotransfection with LDLR and FLAG-tagged MYLIP N342, S342, or C387A ring domain mutant expression plasmids. (C) Immunoblot of HEK293T whole-cell lysates after cotransfection with LDLR and FLAG-tagged MYLIP N342, S342, A342, or C387A ring domain mutant expression plasmids. (D) Human MYLIP mRNA and endogenous LDLR protein expression in *MYLIP*<sup>-/-</sup> MEFs stably expressing empty vector (Vect; pBabe), MYLIP N342, or MYLIP S342. Data are mean  $\pm$  SEM. (E) Immunoblot of HEK293T cells cotransfected with LDLR-GFP; MYLIP N342, S342, or C387A ring domain mutant; and HA-ubiquitin expression plasmids followed by overnight immunoprecipitation of GFP with anti-GFP antibody (1  $\mu$ g).

in adipose, a relevant tissue for lipid metabolism (17). Although the noncoding GWAS SNP rs2480 was in LD with the rs9370867 variant, we also explored the potential regulatory function of these variants as an alternative functional mechanism for this genome-wide significant signal. We examined the correlation between these SNPs and the expression levels of the *MYLIP* gene in subcutaneous fat biopsy samples from 70 Mexican dyslipidemic cases and controls. The expression levels of the *MYLIP* probes were not differentially expressed by the SNP genotypes at a nominal significance level (i.e.,  $P > 0.05$ ). It should also be noted that there are no other genes within  $\pm 100$  kb of the analyzed SNPs.

**Functional characterization of the N342S amino acid substitution.** The N342S substitution (encoded by rs9370867) is located in the FERM domain of MYLIP, a putative protein-protein interaction motif. We explored the possibility that the N342S substitution alters the ability of MYLIP to degrade LDLR. In parallel, we also assessed the effect of the nonsynonymous variants I202L and V339I, identified by sequencing. Our initial studies used native MYLIP expression vectors in order to avoid potential effects of epitope tags on MYLIP function. As there are no available antibodies capable of efficiently detecting native MYLIP, the expression of native MYLIP constructs was monitored by mRNA levels (data not shown). However, we also conducted all of the experiments using tagged proteins to assure that the mutant proteins were expressed at similar levels to WT MYLIP (see below). Since MYLIP is a very potent inhibitor of LDLR, we tested several different MYLIP/LDLR ratios in order to maximize our ability to detect modest changes in MYLIP activity. In agreement with the negative genetic results, the L202 and I339 MYLIP constructs were as effective as the WT MYLIP construct at degrading

LDLR (Supplemental Figure 3, A and B). These observations suggest that I202 and V339 are not critical for MYLIP-induced degradation of LDLR. In contrast, MYLIP S342 demonstrated a reduced ability to degrade LDLR compared with the WT MYLIP with the N342 allele (Figure 4A). This differential activity was apparent over a range of MYLIP/LDLR ratios (Supplemental Figure 3C). Interestingly, this differential activity was consistent with the decreased levels of TC observed in individuals with Ser at residue 342 (Table 2), as lower MYLIP activity is predicted to increase LDLR protein levels and clearance of circulating LDL particles from the plasma.

In order to exclude the possibility that the reduced ability of MYLIP S342 to degrade LDLR was caused by altered stability of the mutant itself, we generated tagged MYLIP constructs. As previously demonstrated (18), the C387A RING domain mutant, which lacks E3 ligase activity, exhibited markedly increased protein stability (Figure 4B) as a result of its inability to undergo autoubiquitination and degradation. In contrast, protein levels resulting from transfection of S342 MYLIP were similar to that of N342 MYLIP. These data indicate that the intrinsic E3 ligase activity of S342 MYLIP and, consequently, its ability to autoubiquitinate is indeed intact. Furthermore, these data also suggest that the observed differences in LDLR degradation activity between S342 and N342 MYLIP cannot be explained by differences in protein levels. Studies using tagged MYLIP constructs also confirmed that I339 and L202 were expressed at comparable levels and showed comparable activity to WT MYLIP (Supplemental Figure 3D).

To further confirm the importance of amino acid 342, we also mutated this residue to an Ala. A342 also showed reduced ability to degrade the LDLR compared with WT N342 (Figure 4C). To





verify that amino acid 342 is important for the ability of MYLIP to degrade endogenous LDLR, we examined endogenous LDLR expression in *MYLIP*<sup>-/-</sup> MEFs stably overexpressing N342 or S342 MYLIP (Figure 4D). The results were in agreement with the experiments in 293T cells transiently transfected with LDLR and MYLIP constructs, further reinforcing our previous findings.

Sequence alignment of the *MYLIP* homologs across various species showed that the Ser residue at amino acid 342 was highly conserved across a large number of species (Supplemental Figure 3E). This raises the possibility that humans have evolved an MYLIP allele, N342, that is more efficient at degrading LDLR, perhaps creating a more tightly regulated LXR-MYLIP-LDLR axis. To explore this idea, we mutated the Ser of mouse *Mylip* at position 342 to an Asn to reflect the human *MYLIP* polymorphism at this residue. As predicted, mouse N342 MYLIP was more efficient than S342 at degrading total LDLR protein (Supplemental Figure 3F).

As the presence of LDLR at the cell surface is essential for the high affinity uptake of LDL particles, we investigated whether the observed effects of the MYLIP mutants on total LDLR protein (Figure 4A) was indicative of protein levels at the cell surface. Using a biotinylation strategy to specifically label plasma membrane proteins, we found that the effects of the MYLIP mutants on surface LDLR levels were similar to that of the total cell LDLR (Supplemental Figure 3G). Ubiquitination assays also confirmed that the S342 MYLIP protein was not able to stimulate LDLR ubiquitination as efficiently as did N342 MYLIP (Figure 4E), consistent with the reduced ability of S342 to degrade LDLR.

Taken together, these functional analyses demonstrated that amino acid 342 is important for the degradation of LDLR by MYLIP. The presence of a Ser at this site resulted in attenuated ubiquitination of LDLR, increased expression of surface LDLR, and a concomitant increase in LDL uptake. This altered activity was in line with the observation that individuals with the G allele of amino acid 342 that translates to a Ser had lower TC levels (Table 2).

## Discussion

In recent GWASs and a GWAS meta-analysis of populations of mixed European descent (6–8), there were 3 independent ( $r^2 < 0.3$ ) genome-wide significant signals in the *MYLIP* region, none of which resides within the *MYLIP* gene. It has been very challenging to unambiguously identify the risk variants and demonstrate function for these common variants implicated in GWASs of complex traits. Consequently, conversion of the GWAS associations into mechanistic insights and potential clinical applications has been a slow process. In the current study, we investigated the *MYLIP*-associated region in Mexican individuals in order to restrict the associated region and identify the actual susceptibility variants (i.e., trans-ethnic fine-mapping). We demonstrated that a common amino acid substitution (N342S; encoded by the SNP rs9370867 variant) located in a critical functional domain of the protein was associated with cholesterol levels in a Mexican dyslipidemic study sample and is most likely the actual susceptibility variant underlying one of the GWAS signals. Importantly, these genetic data were supported by our functional data, which demonstrated that rs9370867 influences plasma cholesterol levels by modifying the degradation of LDLR.

Mexicans descend from a recent mix of Amerindian and European ancestry, which has caused marked differences in allelic frequencies and patterns of LD across their genome (19). Accordingly, the pairwise LD and the allele frequency of N342S are significantly different

between European and Mexican populations (10). Nevertheless, as in the mixed European descent GWAS (6), the same A allele of N342 was associated with high TC in the Mexican dyslipidemic study sample, which implies that this common nonsynonymous variant may actually be the functional variant as opposed to being in strong LD with the causal variant. It should be noted that our sample size was less than 10% of this population-based GWAS (6), yet we observed comparable evidence for rs9370867 in the Mexican dyslipidemic study sample ( $P = 7.0 \times 10^{-07}$ , versus  $P = 5.8 \times 10^{-06}$  in GWAS). Thus, the current work also emphasizes the importance of investigating variants in a different population than the discovery population.

We observed that the nonsynonymous rs9370867 variant was in strong LD with 2 upstream genome-wide significant GWAS variants, rs2294261 (6) and rs2480 (7), in Europeans. The association evidence for rs2294261 and rs2480 was less significant than for rs9370867 in the Mexican dyslipidemic study sample, while their pairwise LD was diminished in Mexicans compared with Europeans, statistically demonstrating that the nonsynonymous variant rs9370867 is more likely to be the causal variant than rs2294261 or rs2480. Importantly, none of these variants exhibited a potential regulatory effect using cis-eQTL analysis, whereas our functional data showed that rs9370867 increased expression of surface LDLR and LDL uptake. Thus, our study demonstrated, for the first time to our knowledge, that the amino acid change rs9370867 is most likely the functional mechanism underlying these previous GWAS variants, rs2294261 and rs2480 (6, 7). Furthermore, our study also exemplifies how populations with a different demographic history, such as Mexicans, can assist in fine-mapping the actual susceptibility variant because of the different underlying LD structure.

Another GWAS signal was detected by the SNPs rs2327951 (6) and rs2142672 (8) that mapped and tagged variants 48–49 kb downstream from *MYLIP*. Using trans-ethnic fine-mapping in Mexicans, we were able to restrict the downstream association signal to a proximate cluster of SNPs (+8–30 kb) that were more strongly associated with cholesterol levels than were rs2327951 and rs2142672 in the Mexican dyslipidemic study sample. Regarding the upstream GWAS signal detected by the SNP rs3757354, we identified another SNP in LD with it, rs2327949, which was slightly more significant in the Mexican dyslipidemic study sample. Although our data still support the possibility that the rs3757354 promoter variant (-2 kb) may be the functional variant, it also provides another plausible underlying candidate for this LD block. It is worth noting that the expression levels of *MYLIP* were not significantly different with respect to the rs3757354 genotypes in the fat biopsies of Mexican subjects or in fat or liver biopsies from subjects of mixed European descent (6). Nevertheless, it is evident that additional association and functional studies are necessary to identify the actual causal variants of these other GWAS signals.

Many of the GWAS loci for lipids are already known to harbor mutations with larger effects in dyslipidemic families (20), which suggests that the newly identified loci, such as *MYLIP*, may also harbor both common variants with weak effects and rare variants with large effects. Thus, although the major focus of this study was to identify the functional *MYLIP* variants underlying the GWAS signals using trans-ethnic fine-mapping, we resequenced Mexican subjects with either low or high levels of LDL-C for the coding sequence of the *MYLIP* gene. Furthermore, subjects with unknown etiology of FH and hypercholesterolemia were also included in order to examine whether mutations in *MYLIP* may contribute to these lipid disorders. We did not observe a greater burden of rare





and low-frequency variants in either tail of the LDL distribution or detect any rare coding mutations in the Mexican FH patients or hypercholesterolemic probands. However, extensive resequencing of the *MYLIP* gene region in larger numbers of subjects from each extreme of the LDL-C is warranted in order to definitively conclude whether rare and low-frequency variants in *MYLIP* account for low and/or high LDL-C levels.

To assure that the associations were not confounded by population admixture, we also performed the test of association including adjustment for ancestry estimates as a covariate. Importantly, although the allele frequency of rs9370867 is substantially different between Amerindian and European populations (10, 14), the adjusted *P* value remained highly significant, which confirmed that the association is not spurious. Furthermore, because the Mexican dyslipidemic study samples were ascertained based on hypertriglyceridemia affection status, we adjusted the association analyses of TC and non-HDL-C for the hypertriglyceridemia case-control status in order to avoid confounding effects of this sampling. Yet a substantial proportion of the hypertriglyceridemic study sample also had elevated TC concentrations. Combined hyperlipidemia, the simultaneous elevation of both TC and TG levels, is highly common among Mexican individuals (18% incidence; ref. 13). Combined hyperlipidemia also carries the highest risk to develop ASCVD in Mexicans (13); therefore, it is important to identify variants, such as rs9370867, that contribute to high TC in the presence of elevated TG levels.

Our genetic findings are supported by functional data demonstrating the importance of the amino acid substitution at residue 342 of *MYLIP*. Consistent with the observation that Mexican individuals with the S342 encoding allele had lower TC levels, we demonstrated that S342 was associated with reduced ubiquitination and degradation of the *LDLR* and enhanced LDL uptake *in vitro* compared with the N342 encoding allele. A predicted consequence of this reduction in *MYLIP* activity *in vivo* is enhanced expression of *LDLR* and subsequent uptake of circulating LDL particles, thus resulting in lower plasma cholesterol levels. However, the specific contribution of the different tissues and cell types exhibiting *MYLIP* activity warrants investigation in future studies.

In conclusion, genetic modulation of *MYLIP* activity via the common amino acid substitution N342S was associated with changes in plasma cholesterol levels in Mexican individuals, supporting a role for *MYLIP* in whole-body cholesterol metabolism. Moreover, our results support further investigation of the possibility that inhibition of *MYLIP* activity may provide a strategy for the treatment of dyslipidemia and ASCVD.

## Methods

**Study participants.** The study design was approved by the ethics committees of INCMNSZ and UCLA, and all subjects provided written informed consent. Clinical characteristics of the study samples described below are shown in Supplemental Table 1.

**Mexican dyslipidemic cases and controls.** A total of 2,822 Mexican dyslipidemic cases and controls were recruited at INCMNSZ, as described in detail previously (21). See Supplemental Table 1 and Supplemental Methods for a detailed description of the study sample. Measurements of fasting lipid levels were performed with commercially available standardized methods, as described previously (21). Serum LDL-C levels were calculated using the Friedewald formula. However, for subjects with TG in excess of 4.5 mmol/l (400 mg/dl; *n* = 947), LDL-C values were not calculated, as the Friedewald equation cannot be reliably used with such high TG values (22).

**Resequencing study sample.** For resequencing, 56 Mexican subjects with LDL-C levels at or below the age/sex-specific 10th Mexican population percentile (23) and 66 subjects with LDL-C levels at or above the age/sex-specific 90th Mexican population percentile (23) were selected from each extreme of the LDL-C distribution. The high-LDL-C group consisted of 35 unrelated individuals: 21 FH patients from 13 Mexican FH families without known mutations in the *LDLR*, *APOB*, and *PCSK9* genes (12); and 10 Mexican dyslipidemic probands from families that provided positive lod score for linkage with markers in the *MYLIP* region in our previous study (6p22.3) (21).

**Resequencing, genotyping, and imputation.** The exons, UTRs, and exon-intron boundaries of the *MYLIP* gene (NM\_013262) were resequenced. The samples were amplified by PCR for the Applied Biosystems 3730 Capillary DNA Analyzer (Applied Biosystems). PCR primers were designed using the Primer3 program. Sequence contigs were assembled with Sequencher software (GeneCodes Corp.). Novel variants that were not present in the dbSNP build 130 were sequenced on both strands.

We genotyped the variants identified by resequencing in larger study samples — tag SNPs in the *MYLIP* gene region in the Mexican dyslipidemic study sample and SNPs for cis-eQTL analysis in the fat biopsy study sample — using Pyrosequencing, TaqMan, and Illumina BeadArray technology genotyping platforms. All SNPs were in Hardy-Weinberg equilibrium ( $P > 0.05$ ) in the Mexican controls and had a genotyping call rate greater than 95%, with the 34 tag SNPs exhibiting a genotyping call rate greater than 98%.

A total of 34 tag SNPs in the *MYLIP* gene and its surrounding  $\pm 50$ -kb region were genotyped in 2,310 subjects from the Mexican dyslipidemic study sample. We used the IMPUTE2 program (24) and phased chromosomes from the phase III HapMap MEX sample as a reference to impute SNPs with MAF  $\geq 1\%$  cataloged in the HapMap in the *MYLIP* gene region ( $\pm 50$  kb). All 38 imputed SNPs exhibited good imputation quality (certainty,  $> 92\%$ ) and were in Hardy-Weinberg equilibrium ( $P > 0.1$ ) in the Mexican controls. After excluding imputed genotypes with low quality score (i.e.,  $< 0.7$ ), the genotype call rate exceeded 90%, and the estimated error rate based on a leave-one-out approximation was less than 2.5% (24).

The most significantly associated SNP, rs9370867, was also genotyped in 512 additional subjects from the Mexican dyslipidemic study sample (i.e., the entire Mexican dyslipidemic study sample; *n* = 2,822). In addition, the SNPs rs9370867, rs2480, rs2142672, rs3757354, and rs9367897 were genotyped in 70 Mexican dyslipidemic cases and controls who underwent subcutaneous fat biopsies (25).

**Association and LD analyses.** Multivariate linear regression using an additive genetic model with age and sex as covariates was used to assess the effects of the SNPs on continuous TC and non-HDL-C, and multivariate logistic regression was used to analyze the binary TC and combined hyperlipidemia status. The regression models of the TC and non-HDL-C traits were adjusted for hypertriglyceridemia affection status in order to correct for the sampling variable. The likelihood was maximized using an expectation-maximization algorithm, as implemented in SNPTEST version 2.1.1 software (26), in order to account for genotype uncertainty in the regression analyses of the imputed SNPs. We also examined whether the association results are independent of the top signal rs9370867 by including the minor allele counts (i.e., 0–2) of rs9370867 as a covariate in the regression analyses as described above. Subjects with trait levels more than 4 SD from the mean or on lipid-lowering therapy at the time of the blood drawing were excluded from the quantitative analyses. To correct for multiple testing, we adjusted for the number of independent SNPs ( $r^2 < 0.8$ ) in the region (*n* = 35) that we analyzed with 2 traits (TC and non-HDL-C), resulting in Bonferroni correction for 70 independent tests ( $P \leq 7.2 \times 10^{-4}$ ).

To adjust for the admixed Mexican ancestry, IA estimates were used as a covariate in the regression analyses (see Supplemental Methods).



In order to predict the association strength of rs9370867 in the GWAS meta-analysis using a larger sample size of 95,443 subjects, we calculated the estimated  $P$  as  $2 \times (1 - \text{standard normal distribution}(|Z_{\text{obs}}|/\text{SQRT}[n_{\text{obs}}/n_{\text{est}}]))$ , in which the  $Z$  score obtained in the GWAS meta-analysis ( $Z_{\text{obs}}$ ) was 4.5, the actual sample size in the study ( $n_{\text{obs}}$ ) was 34,733, and the estimated sample size ( $n_{\text{est}}$ ) was 95,443.

We extracted from the CEU and MEX populations of HapMap (10) all SNPs polymorphic in both the Mexican and European populations ( $n = 65$ ) in the surrounding regions of the *MYLIP* gene ( $\pm 50$  kb). Pairwise LD (in  $r^2$ ) of all SNP pairs (i.e., triangular grid of LD relationships) was calculated in the CEU and MEX HapMap (10) founder samples ( $n = 60$ ) and plotted using the LDheatmap library in the R software package (Supplemental Figure 2).

Pairwise LD of the SNPs rs3757354, rs9370867, and rs2327951 with all other SNPs in the *MYLIP* gene region ( $\pm 50$  kb) were calculated using PLINK version 1.07 software (27) and plotted using the Lattice graphics library in the R software package (Figure 3). For this LD analysis, we used the extracted genotype data ( $n = 65$ ) of the CEU founder sample ( $n = 60$ ) of HapMap and genotyped ( $n = 34$ ) and imputed ( $n = 38$ ) data of 120 Mexican controls (TGs,  $<150$  mg/dl; TC,  $<240$  mg/dl) with complete genotype data and certainty greater than 0.92 for all imputed genotypes. The estimated error rate for the imputed genotypes in this sample was less than 1.3%, based on a leave-one-out approximation (24). Hence, to account for this genotype uncertainty, we used a larger sample size in the LD analyses of Mexicans (240 chromosomes, versus 120 available for the CEU sample).

The coverage of the 72 analyzed SNPs in the MEX and CEU populations of HapMap and the number of independent SNPs analyzed ( $n = 35$ ) were calculated with  $r^2$  threshold  $>0.8$  and MAF  $\geq 1\%$  using the tagger algorithm as implemented in Haploview version 4.2 software (28).

**cis-eQTL analysis.** A total of 70 Mexican dyslipidemic cases and controls underwent subcutaneous fat biopsies. Clinical characteristics, fat biopsy procedure, RNA extraction, and RNA hybridization to the Human Genome U133 Plus 2.0 oligonucleotide microarray of the 70 samples have been described previously (25). Data processing and quality control procedures of the microarrays were performed as described previously (25). A total of 14,942 probesets passed quality control procedures (25), of which only the *MYLIP* probesets ( $n = 3$ ) were used in the cis-eQTL analysis. Microarray data are available at the NCBI Gene Expression Omnibus (GEO) database (accession no. GSE17170). We examined the association of expression levels between *MYLIP* probesets 220319\_s\_at, 223130\_s\_at, and 228098\_s\_at and SNPs rs9370867, rs2480, rs2142672, rs3757354, and rs9367897. The normalized relative expression adjusted for age, sex, kinship, and affection status was used as a dependent variable in linear regression with the SNP genotypes coded for the additive model in R software package.

**Cells, reagents, constructs, transfections.** The FLAG-tagged construct was a gift of E. Saez (Scripps Research Institute, La Jolla, California, USA). All other constructs were obtained as previously described (4). Mutant *MYLIP* constructs were generated using the Quickchange site-directed mutagenesis kit (Stratagene). DNA sequencing was used to verify mutant constructs. HEK 293T cells were maintained in DMEM supplemented with 10% FBS (Omega Scientific). Transfections were performed using Fugene (Roche Diagnostics) according to the manufacturer's instructions, with a LDLR/*MYLIP* ratio of 2:1 or 4:1 unless otherwise indicated. Cells were harvested approximately 24–48 hours after transfection. *MYLIP*<sup>-/-</sup> MEFs were immortalized by stable expression of SV40 Large-T Antigen retrovirus and subsequent selection by hygromycin B. Stable expression of control retrovirus (pBabe), *MYLIP*-N342, or *MYLIP*-S342 was performed as previously described (4), and expressing cells were

selected with puromycin. Cells were maintained in DMEM supplemented with 10% FBS and MEM nonessential amino acids (Gibco; Invitrogen) unless otherwise indicated.

**Immunoprecipitation and immunoblotting.** HEK293T cells were washed with PBS, then harvested in RIPA buffer supplemented with protease inhibitors. Lysates were clarified by centrifugation, then quantified using the Bradford assay (Biorad) with BSA as a reference. Proteins were separated by 4%–12% Nupage Bis-Tris gels, then transferred to PVDF (GE Osmonics). Membranes were probed with the following antibodies:  $\beta$ -actin (Sigma-Aldrich; 1:5,000), FLAG (Sigma-Aldrich; 1:1,000), HA (Covance; 1:1,000), LDLR (Cayman Chemical Co.; 1:1,000), pan-cadherin (Santa Cruz; 1:1,000), and  $\alpha$ -tubulin (Calbiochem; 1:3,000). Appropriate secondary HRP-conjugated antibodies were used (Invitrogen, Biorad) and visualized with chemiluminescence (Amersham). To assess cell surface expression, samples were biotinylated (Thermo Scientific; 250  $\mu\text{g}/\text{ml}$ ) for 30 minutes at 4°C, then subsequently quenched and washed with PBS prior to harvesting in RIPA buffer with protease inhibitors. Equal amounts of lysate were incubated with neutravidin agarose resin (Pierce) overnight with rotation, then washed and heated to 95°C with 2 $\times$  sample buffer for 5 minutes. For immunoprecipitation, equal amounts of lysate were incubated with anti-GFP antibody (Abcam; 1  $\mu\text{g}$ ) overnight with rotation followed by the addition of protein G beads (Santa Cruz). Samples were washed, then heated to 70°C with 2 $\times$  sample buffer prior to immunoblotting.

**RNA isolation and quantitative PCR.** Total RNA was extracted from cells using TRIzol (Invitrogen). 500 ng RNA was reverse transcribed using the iScript cDNA synthesis kit (Biorad) according to the manufacturer's instruction. Sybergreen (Diagenode) real-time quantitative PCR assays were performed on an Applied Biosystems 7900HT Fast Real-Time PCR system. Data are expressed as the average of duplicate samples normalized to 36B4.

**Statistics.** Multivariate linear and logistic regression for the additive genetic model was used to assess the effect of the SNPs on continuous and dichotomous lipid traits, respectively. 2-sided  $P$  values surpassing the Bonferroni multiple-comparison correction threshold for 70 independent tests ( $P \leq 7.2 \times 10^{-04}$ ) were considered significant. For cis-eQTL analysis, we performed multivariate linear regression for the additive genetic model. 2-sided  $P$  values less than 0.05 were considered significant.

## Acknowledgments

We thank the Mexican individuals who participated in the study. We also thank Elina Nikkola, Cindy Montes, Salvador Ramírez, and Maribel Rodríguez for laboratory technical assistance. This research was supported by grants HL095056, HL28481, HL066088, and HL090553 from the NIH. A.C. Calkin was funded by a National Heart Foundation of Australia Overseas Fellowship (O 08M 3934). P. Tontonoz is an Investigator of the Howard Hughes Medical Institute.

Received for publication January 20, 2011, and accepted in revised form May 25, 2011.

Address correspondence to: Päivi Pajukanta, Department of Human Genetics, David Geffen School of Medicine at UCLA, Gonda Center, Room 6335B, 695 Charles E. Young Drive South, Los Angeles, California 90095-7088, USA. Phone: 310.267.2011; Fax: 310.794.5446; E-mail: ppajukanta@mednet.ucla.edu.

Noam Zelcer's present address is: Department of Medical Biochemistry, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands.



1. Lloyd-Jones D, et al. Executive summary: heart disease and stroke statistics--2010 update: a report from the American Heart Association. *Circulation*. 2010;121(7):948-954.
2. Yusuf S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*. 2004;364(9438):937-952.
3. Golomb BA, Evans MA. Statin adverse effects: a review of the literature and evidence for a mitochondrial mechanism. *Am J Cardiovasc Drugs*. 2008;8(6):373-418.
4. Zelcer N, Hong C, Boyadjian R, Tontonoz P. LXR regulates cholesterol uptake through MYLIP-dependent ubiquitination of the LDL receptor. *Science*. 2009;325(5936):100-104.
5. Scotti E, et al. Targeted disruption of the MYLIP gene alters cellular regulation of the LDLR by sterols and LXR agonists. *Mol Cell Biol*. 2011;31(9):1885-1893.
6. Teslovich TM, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466(7307):707-713.
7. Chasman DI, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet*. 2009;5(11):e1000730.
8. Waterworth DM, et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2010;30(11):2264-2276.
9. McCarthy MI, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9(5):356-369.
10. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature*. 2005;437(7063):1299-1320.
11. Meaney E, et al. Cardiovascular risk factors in the urban Mexican population: the FRIMEX study. *Public Health*. 2007;121(5):378-384.
12. Robles-Osorio L, et al. Genetic heterogeneity of autosomal dominant hypercholesterolemia in Mexico. *Arch Med Res*. 2006;37(1):102-108.
13. Aguilar-Salinas CA, Gómez-Pérez FJ, Rull J, Villalpando S, Barquera S, Rojas R. Prevalence of dyslipidemias in the Mexican National Health and Nutrition Survey 2006. *Salud Pública Mex*. 2010;52 suppl 1:S44-S53.
14. Pickrell JK, et al. Signals of recent positive selection in a worldwide sample of human populations. *Genome Res*. 2009;19(5):826-837.
15. Lanktree MB, Anand SS, Yusuf S, Hegele RA. Replication of genetic associations with plasma lipoprotein traits in a multiethnic sample. *J Lipid Res*. 2009;50(7):1487-1496.
16. King DC, Taylor J, Elnitski L, Chiaromonte F, Miller W, Hardison RC. Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences. *Genome Research*. 2005;15(8):1051-1060.
17. Hegele RA. Plasma lipoproteins: genetic influences and clinical implications. *Nat Rev Genet*. 2009;10(2):109-121.
18. Bornhauser BC, Johansson C, Lindholm D. Functional activities and cellular localization of the ezrin, radixin, moesin (ERM) and RING zinc finger domains in MIR. *FEBS Lett*. 2003;553(1-2):195-199.
19. Price AL, et al. A genomewide admixture map for Latino populations. *Am J Hum Genet*. 2007;80(6):1024-1036.
20. Weissglas-Volkov D, Pajukanta P. Genetic causes of high and low serum HDL-cholesterol. *J Lipid Res*. 2010;51(8):2032-2057.
21. Weissglas-Volkov D, et al. Identification of two common variants contributing to serum apolipoprotein B levels in Mexicans. *Arterioscler Thromb Vasc Biol*. 2010;30(2):353-359.
22. Sniderman AD, Blank D, Zakarian R, Bergeron J, Frohlich J. Triglycerides and small dense LDL: the twin Achilles heels of the Friedewald formula. *Clin Biochem*. 2003;36(7):499-504.
23. Aguilar-Salinas CA, et al. High prevalence of low HDL cholesterol concentrations and mixed hyperlipidemia in a Mexican nationwide survey. *J Lipid Res*. 2001;42(8):1298-1307.
24. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009;5(6):e1000529.
25. Plaisier CL, et al. A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial combined hyperlipidemia. *PLoS Genet*. 2009;5(9):e1000642.
26. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39(7):906-913.
27. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575.
28. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263-265.