Circulation
Cardiovascular Genetics

American Heart Association

Putting Pleiotropy and Selection Into Context Defines a New Paradigm for Interpreting Genetic Data

Irene M. Predazzi, Antonis Rokas, Amos Deinard, Nathalie Schnetz-Boutaud, Nicholas D. Williams, William S. Bush, Alessandra Tacconelli, Klaus Friedrich, Sergio Fazio, Giuseppe Novelli, Jonathan L. Haines, Giorgio Sirugo and Scott M. Williams

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/6/3/299

Data Supplement (unedited) at:
http://circgenetics.ahajournals.org/content/suppl/2013/04/24/CIRCGENETICS.113.000126.DC1.html

Original Article

Putting Pleiotropy and Selection Into Context Defines a New Paradigm for Interpreting Genetic Data

Irene M. Predazzi, PhD; Antonis Rokas, PhD; Amos Deinard, PhD, DVM; Nathalie Schnetz-Boutaud, PhD; Nicholas D. Williams; William S. Bush, PhD; Alessandra Tacconelli, MSc; Klaus Friedrich, PhD, DVM; Sergio Fazio, MD, PhD; Giuseppe Novelli, PhD; Jonathan L. Haines, PhD; Giorgio Sirugo, MD, PhD; Scott M. Williams, PhD

Background—Natural selection shapes many human genes, including some related to complex diseases. Understanding how selection affects genes, especially pleiotropic ones, may be important in evaluating disease associations and the role played by environmental variation. This may be of particular interest for genes with antagonistic roles that cause divergent patterns of selection. The lectin-like low-density lipoprotein 1 receptor, encoded by *OLR1*, is exemplary. It has antagonistic functions in the cardiovascular and immune systems because the same protein domain binds oxidized low-density lipoprotein and bacterial cell wall proteins, the former contributing to atherosclerosis and the latter presumably protecting from infection. We studied patterns of selection in this gene, in humans and nonhuman primates, to determine whether variable selection can lead to conflicting results in cardiovascular disease association studies.

Methods and Results—We analyzed sequences from 11 nonhuman primate species, as well as single-nucleotide polymorphisms and sequence data from multiple human populations. Results indicate that the derived allele is favored across primate lineages (probably because of recent positive selection). However, both the derived and ancestral alleles were maintained in human populations, especially European ones (possibly because of balancing selection derived from dual roles of *LOX-1*). Balancing selection likely reflects response to diverse environmental pressures among humans.

Conclusions—These data indicate that differential selection patterns, within and between species, in *OLR1* render association studies difficult to replicate even if the gene is etiologically connected to cardiovascular disease. Selection analyses can identify genes exhibiting gene–environment interactions critical for unraveling disease association. (***Circ Cardiovasc Genet***. 2013;6:299-307.)

Key Words: evolution ◼ genetics ◼ immune system ◼ lipoproteins ◼ LOX-1 receptor ◼ Pleitropic gene

Genes that associate with complex disease are likely to be shaped by natural selection. However, the underlying pattern of selection may be confounded by multiple functions of a single gene or pleiotropy.[1] Specifically, if a gene has 2 functions and a variant improves function for 1 phenotype but has the opposite effect on the second, a phenomenon known as antagonistic pleiotropy, it will influence our ability to detect associations in a context-dependent fashion, potentially masking important biology. One way to explore this phenomenon is to compare inter- and intraspecific patterns of genetic variation in genes with multiple known functions and to assess patterns of selection.

Clinical Perspective p 307

We tested this hypothesis in the lectin-like low-density lipoprotein receptor 1 gene (*OLR1*, Online Mendelian Inheritance in Man [OMIM] 602601), which encodes lectin-like low-density lipoprotein receptor 1 (LOX-1), thought to increase cardiovascular disease (CVD) risk by acting as a scavenger receptor for oxidized low-density lipoprotein in endothelial cells.[2] LOX-1 also binds other ligands, including bacterial cell wall proteins, thus playing an important protective role in immune function.[3] These 2 biological functions likely operate independently and antagonistically. Although it is impossible to know with certainty what the dominant function was and is for *OLR1*, the immune function role is likely the ancestral one because ancient humans had a short life span and were not likely exposed to the selection pressure of atherosclerosis-based CVD.

Regarding the possible functional role of genetic variation in *OLR1*, studies have demonstrated that a haplotype containing single-nucleotide polymorphisms (SNPs) in and around intron 4 causes alternative splicing of exon 5 and leads to increased synthesis of a truncated isoform known as LOXIN.[4,5] The SNPs in this haplotype (rs3736232, rs3736234, rs3736233, rs3736235, rs3816844, and rs1050283) are in strong linkage disequilibrium (LD) with each other in all populations, and the region is therefore often studied using only 1 representative SNP.[4] Because exon 5 encodes a C-type lectin-binding domain, crucial for LOX-1 to bind its ligands, LOXIN is dysfunctional (Figure 1).[5–7] Despite compelling experimental data indicating loss of function because of this variation,[5] studies investigating association with disease have often yielded inconsistent and even contradictory results (Table I in the online-only Data Supplement).[8] We hypothesize that this is because LOX-1 is favored in some environments, whereas LOXIN is favored in others. Variation in the relative importance of the 2 functions and the subsequent divergent selection may confound results of studies addressing association with CVD. For example, validation of epidemiological results using different populations or different end points might lead to contrasting results if the 2 functions of the gene are both involved in the pathogenesis of the disease, as is the case for LOX-1 and LOXIN in atherosclerosis. In addition, because both immunity and lipid oxidation are important players in atherogenesis, it is reasonable to expect classical genetic association studies to be influenced by this effect. To determine whether this is likely, we performed analyses of *OLR1* with respect to its evolutionary history of both inter- and intraspecies.

Results

Evidence for Ancient Positive Selection in Intron 4 of the *OLR1* Locus

To examine whether and how ancient selection has shaped the *OLR1* locus, we conducted 2 selection tests on coding and noncoding regions in primate lineages. First, we evaluated the ζ ratio of nonsynonymous to synonymous substitution rate in coding regions and the ω ratio of noncoding to synonymous substitution rate for the intron 4 region across the 12 primate species (human+11 nonhumans). We found strong evidence of selection in intron 4 ($P=0.001$) but not in the coding sequence of the *OLR1* locus ($P=0.519$; Table 1). Furthermore, using either the Naive Empirical Bayes algorithm (which uses maximum likelihood estimates of model parameters to identify sites undergoing positive selection) or the Bayes Empirical Bayes algorithm (which assigns priors for model parameters and estimates averages by numeric integration over these priors to identify sites under positive selection),[9] we identified 3 sites (sites 112, 207, and 224 from the beginning of the amplicon) with posterior probability of positive selection ≥92% (Table II in the online-only Data Supplement).

Evidence for Intraspecies Balancing Selection in Intron 4 Within Humans and for Intraspecies Positive Selection Within Chimps

In humans, especially the HapMap CEU population, balancing selection is likely to be acting on *OLR1* as indicated by high Tajima's *D* from HapMap 2 data ($D=3.54$; $P<0.001$; Figure 2; Table 2). When we divided the gene into 2 distinct regions defined by their LD patterns, (1) the region of high LD between rs12822177 and rs11615002 that contains intron 4 and is marked in Figure 1 as including the C-type lectin-binding domain and (2) the rest of the gene, we observed a higher Tajima's *D* in the LD block region in Europeans ($D_{LD}=3.49$; $D_{Non-LD}=2.88$; Figure 2A; Table 2). Similarly, F_{st} for the region of LD was higher than for the rest of the gene ($F_{st-LD}=0.27$; $F_{st-non\ LD}=0.080$), supporting differential selection patterns in different parts of the gene (Table 2). Similar results were found for SNPs in the LD block derived from 1000 genome samples of majority of European ancestry, indicating
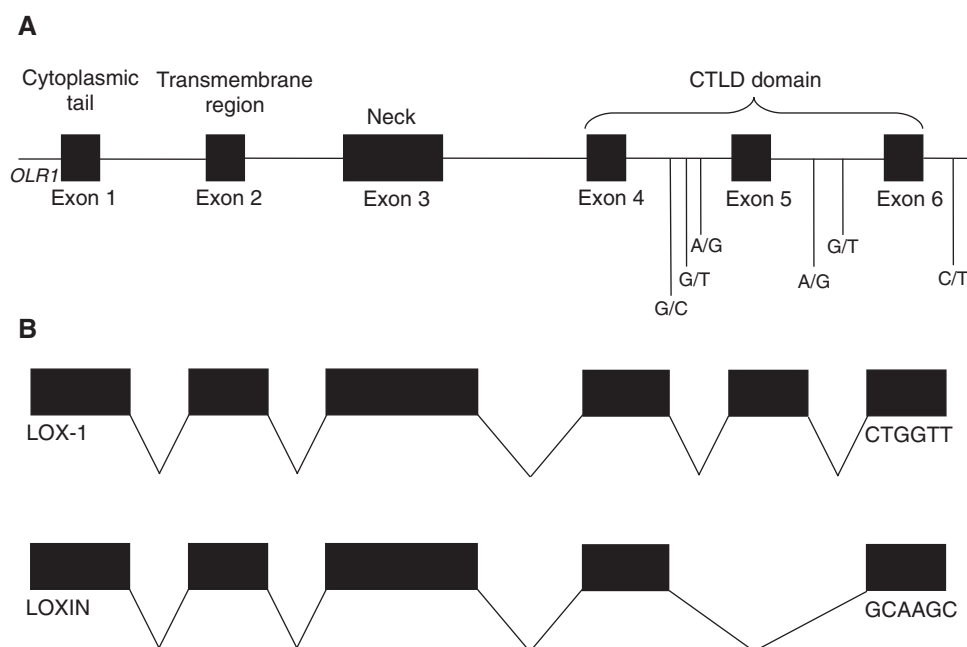

Figure 1. A, Structure of *OLR1* with location of variants (in order from **left** to **right**: rs3736232, rs3736234, rs3736233, rs3736235, rs3816844, and rs1050283, $r^2>0.95$) that modulate the ratio between the lectin-like low-density lipoprotein receptor 1 (LOX-1) and LOXIN mRNA. Bracket identifies region of strong linkage disequilibrium CTLD indicates c-type lectin domain.

Table 1.   Results of Ancient (Interspecific) Selection Analyses on Coding (ω) or Noncoding (ζ) Regions of *OLR1*

ω/ζ Ratio Tests	Coding Region				Intron 4			
	M1a		M2a		ncM1a		ncM2a	
Model (Data)	Frequency	ω	Frequency	ω	Frequency	ζ	Frequency	ζ
Class 0 (0<ω<1 or 0<ζ<1)	0.427	0.000	0.450	0.024	0.988	0.379	0.843	0.379
Class 1 (ω=1 or ζ=1)	0.573	1.000	0.503	1.000	0.014	1.000	0.012	1.000
Class 2 (ω>1 or ζ>1)	n/a	n/a	0.047	3.229	n/a	n/a	0.147	4.748
Ln likelihood	−1770.666		−1770.009		−2796.810		−2787.700	
Parameters		2.000				4.000		
Likelihood ratio test		1.313				18.214		
P value		0.519				0.001		

n/a indicates not applicable.

that for these continental populations, common and rare SNPs share patterns of selection (Table III in the online-only Data Supplement). In general, the European data provide the strongest evidence for selection in *OLR1*.

In contrast to the pattern of selection found among humans, several *OLR1* gene regions in *Pan troglodytes* (chimpanzee) showed evidence for intraspecific positive selection. Specifically, negative Tajima's *D* values were found for the 5′ untranslated region, intron 4, and 3′ untranslated region (*D*=−2.381, −1.947, and −1.821 respectively; Figure 2D), indicating intraspecific positive selection in these regions. In contrast, the other regions of the gene did not show evidence of positive selection, with *D* values ranging from −1.800 to 0.000 (Table 2). Taken together, the data from humans and chimpanzees indicate different types of selective pressure (balancing versus positive) acting on a segment that includes intron 4 in both humans and chimpanzees.

Evolutionary History of the *OLR1* Locus Is Consistent With Species Phylogeny
The action of selection on a locus can cause its phylogeny to deviate significantly from that of the species.[10,11] To test whether this was the case for the *OLR1* locus, we reconstructed the evolutionary history of the locus using both coding and noncoding data (Figure I in the online-only Data Supplement). Although the locus history is largely in accordance with species phylogeny, sequences from species with multiple individuals often fail to form monophyletic groupings, with some of the alleles representing trans-species polymorphisms (eg, alleles in humans that are more similar to alleles in chimpanzees than they are to alleles derived from other humans).[12] Although only balancing selection can maintain trans-species polymorphisms for the long-term, transient trans-species polymorphisms could also be consistent with directional selection or neutrality.[10] To test this possibility, we compared whether the maximum likelihood estimated topologies, that is, the branching patterns of phylogenetic trees, from the entire set of coding and noncoding sequences were significantly different from topologies constrained to obey the primate phylogeny and monophyly of alleles from each species using the Shimodaira–Hasegawa test.[13] The maximum likelihood topologies were not significantly different

from the species phylogeny (both tests were not significant at *P*=0.05), indicating that trans-specific *OLR1* polymorphisms do not exist in primates.

Discussion
Analyses of primate sequences of *OLR1* indicate complex patterns of selection based on species and the part of the gene being assessed. Our results support the hypothesis that intron 4 and its surrounding region have been shaped by differential selection in human versus nonhuman primate lineages. To summarize, although analyses of the coding region indicate selective neutrality, we found strong evidence of ancient positive selection in the intron 4 region in primates (Table 1). Comparable analyses of 18 chimpanzee sequences also indicated intraspecific positive selection (Figure 2D; Table 2), whereas in humans (especially Europeans) there was strong evidence for balancing selection in intron 4 (Figure 2A–2C; Table 2). Hence, the mode of recent selection varies between these very closely related species.

Differential patterns of selection in *OLR1* can be interpreted in light of how the LOX-1 and LOXIN proteins function and their impact on phenotypic variation. For LOX-1 to bind its ligands, it must contain the sequence encoded in exon 5. The ability to bind ligands, however, has been significantly associated with several predominantly human diseases (CVD, hypertension, Alzheimer's; Table I in the online-only Data Supplement). In contrast to its presumed negative effect on cardiovascular disease, the function of LOX-1 is necessary for binding to bacteria and to trigger the subsequent response to infection.[3]

Antagonistic pleiotropy acting on *OLR1* is consistent with the pattern of balancing selection in certain human populations. Balancing selection, which maintains alleles in a polymorphic state, may in the case of *OLR1* be because of both the direction and strength of selection varying with immediate environments. For example, for a late-onset disease such as coronary artery disease, the selection pressure may be weaker than that from infection in certain populations (eg, humans in tropical regions), whereas it may be stronger in humans in nontropical climates, such as Europe. In contrast, selection pressure for late-onset noncommunicable disease among nonhuman primates is likely low, and therefore locus variation in these

Figure 2. Representation of linkage disequilibrium (LD) of single-nucleotide polymorphisms (SNPs; r2) in HapMap (**A**) Caucasian European, (**B**) Han Chinese from Beijing+Japanese from Tokyo, (**C**) Yoruba from Ibadan, and (**D**) chimps from the present study. Tajima's *D* values are shown by the colored bars above the exons (if $D \geq 3$ then balancing selection, $P<0.001$; if $2<D>3$ then balancing selection, $P<0.01$; if $D=2$ then balancing selection, $P<0.05$; if $D<-2$ then positive selection, $P<0.001$; if $D=2$ then positive selection, $P<0.05$; if $-2<D>1.8$ then neutral). Level of significance for each region is also indicated above the exons by *$P<0.05$, **$P<0.01$, and ***$P<0.001$, and darker shades represent stronger selection signals. The region that demarks the C-type lectin domain also demarks

Figure 2. (continued) what we refer to as the LD block (between SNPs rs12822177 and rs11615002). Markers: 1=rs2742110, 2=rs11053654, 3=rs2634162, 4=rs2742113, 5=rs2742114, 6=rs2742115, 7=rs16910917, 8=rs3741860, 9=rs3912640, 10=rs11611453, 11=rs11611438, 12=rs2010655, 13=rs11053649, 14=rs6488265, 15=rs11053648, 16=rs12822177, 17=rs11053646, 18=rs3736232, 19=rs3736233, 20=rs3736234, 21=3736235, 22=rs3816844, 23=rs2634156, 24=rs12309394, 25=rs1050283, 26=rs10505755, 27=rs1050286, and 28=rs1050289 (**A–C**). Polymorphisms in chimpanzees have no rs numbers.

Table 2.   Intraspecific Selection Analyses in Humans and Chimpanzees by Region of the Gene *OLR1*

Species	Population/Region	Tajima's D	Fu and Li D	Fu and Li F	Population Differentiation (F_{st})
Humans					
Homo sapiens	CEU–overall	3.54***	1.84**	3.00**	0.117***
Homo sapiens	YRI–overall	1.77	1.40	1.84*	
Homo sapiens	CHB+JPT–overall	3.50***	1.83**	3.01**	
Homo sapiens	CEU–LD block	3.49***	1.41	1.84*	0.270***
Homo sapiens	YRI–LD block	0.89	0.81	1.01	
Homo sapiens	CHB+JPT–LD block	2.37*	1.34	2.05**	
Homo sapiens	CEU–non-LD block	2.88**	1.58*	1.58*	0.080***
Homo sapiens	YRI–non-LD block	2.30*	1.41	2.06*	
Homo sapiens	CHB+JPT–non-LD block	3.70***	1.45	2.76**	
Chimpanzees					
Pan troglodytes	5′UTR+exon 1	−2.381**	−3.60**	−3.77**	…
Pan troglodytes	Intron 4	−1.947*	−2.19	−2.45	…
Pan troglodytes	Exon 5	−1.012	1.60	1.32	…
Pan troglodytes	Exon 6′+3′UTR	−1.821*	−2.58*	−2.74*	…

For human populations, population stratification indexes are also reported.

CEU indicates Caucasian European; LD, linkage disequilibrium; SNP, single-nucleotide polymorphisms; and UTR, untranslated region.

*$P<0.05$, **$P<0.02$, and ***$P<0.001$. LD block refers to the SNPs between rs12822177 and rs11615002, and non-LD block refers to all the SNPs in *OLR1* that are out of that region.

has more likely been shaped by positive selection in response to early onset diseases such as those because of infectious agents. Balancing selection could also be caused by heterozygote advantage. However, our data are insufficient to directly address this possibility.

This and previous physiological and immunological studies support a likely role for balancing selection in humans, but the case in nonhuman primates is less obvious. Clearly, our interspecific analyses provide support for ancestral positive selection, but functional data in nonhuman primates are lacking. Therefore, it is impossible to infer exactly what the biological role of these variants is in nonhuman primates. Nonetheless, the analyses clearly provide support to the idea that the polymorphisms in intron 4 confer differences based on variable selective effects between species. However, the exact variants cannot be determined because they are not directly functionally comparable between humans and primates.

Our data provide a unique perspective on the role of selection in a gene that has multiple and putative antagonistic roles, indicating that future studies examining the phenotypic role of *OLR1* might best be focused on the region where we found evidence of selection. In addition, we argue that association studies with *OLR1* will need to adjust for ethnicity because the patterns of selection vary by population. Pleiotropy leading to a complex signature of selection may play out differently depending on the environmental context and may provide a paradigm for genes, such as *OLR1*, that confer risk for multiple diseases.

In summary, the results suggest that antagonistic pleiotropy may obscure genetic associations between coronary artery disease and *OLR1*, a gene whose multiple functions influence diverse, antagonistic, or competing biological processes. For example, atherogenesis is affected by both immunity and oxidized low-density lipoprotein internalization. In addition, the 2 variants, LOX-1 and LOXIN, affect at least one of these processes antagonistically. Nonetheless, our data are consistent with the concept that evolution shapes genetic structure differently in relatively short-lived human populations with a high pathogen load (eg, developing world populations), where there is likely to be strong selection for the variant that increases the proportion of LOX-1 product compared with populations in the developed world with access to quality care and nutrition. In the latter, environment infection may not provide a large selective pressure, allowing LOXIN to increase in frequency. In both cases, changing patterns of allelic variation may over the long-term reduce power to detect phenotype association in some, but not all populations. Antagonistic roles may also drive balancing selection in some populations, further masking effects on disease, especially when separate phenotypes are being assessed, such as myocardial infarction, stroke, or carotid intima-media thickness (Table I in the online-only Data Supplement). These complex biological roles may explain why functional studies on LOX-1 have clearly explained its role in the pathogenesis of coronary artery disease, whereas epidemiological studies have produced inconsistent and at times opposite outcomes.

Our data showing complex patterns of selection in *OLR1* support the argument that environmentally dependent antagonistic pleiotropy is likely to affect the pattern of genetic variation in this gene, and hence association results as such studies are impacted by allelic distributions and patterns of LD. Therefore, assessing patterns of selection can identify genes, such as *OLR1*, for which knowledge of function is key to defining disease risk because they may be more sensitive to gene–environment interactions than those with simpler patterns of selection.

Experimental Procedures

Samples

A total of 48 samples were collected and sequenced from human and nonhuman primates, covering 12 species. The 3 human samples used were deidentified. An additional 6-reference sequence, 1 each from *Homo sapiens*, *P. troglodytes*, *Pongo pygmaeus*, *Nomascus leucogenys*, *Macaca mulatta*, and *Callithrix jacchus*, was downloaded from the Ensembl database. Details regarding the samples are presented in Table IV in the online-only Data Supplement. DNA was extracted from blood samples using the phenol–chloroform method.[14]

Amplification and Sequencing

Primers were designed using conserved regions of the Ensembl 6 primate alignment (http://www.ensembl.org/Homo_sapiens/Gene/Compara_Alignments?align=511&db=core&g=ENSG00000173391&r=12%3A10310902-10324737). At least 50 bp of flanking intronic sequence were included in all exon amplimers. The entirety of intron 4 was amplified because previous data suggested that it included a variant that affected alternative splicing. Primer sequences and reaction conditions for all the primers used are shown in Table V in the online-only Data Supplement, and the amplified regions are shown in Figure II in the online-only Data Supplement. Amplification reactions were performed with Taq Platinum according to the manufacturer's protocols (Invitrogen). Each fragment was sequenced in both the forward and reverse directions. Sanger sequencing was performed in the Vanderbilt University Genome Resources Core.

Tests for Ancient Selection on Coding and Noncoding Regions

We tested for ancient selection in both coding and noncoding regions of 2 different taxon sets. The coding region contained exons 1, 2, 3, 4, 5, and 6. The noncoding region sequence contained intron 4. The first taxon set (primate data set) included a single sequence from each of 12 different primate species, whereas the second taxon set (human-chimp data set) included 4 human sequences and 17 de novo chimpanzee sequences.

Sequence Alignment

All coding sequence alignments were performed in protein space using the MAFFT software, version 6.811[15,16] and back-translated to codons using the PAL2NAL software,[17] whereas all noncoding sequence alignments were performed in nucleotide space using the MAFFT software, version 6.811.[15,16]

Test for Ancient Selection in Coding Regions

We tested the coding region of the primate data set for evidence of positive selection by estimating the ω ratio of the nonsynonymous substitution rate (d_N) to the synonymous substitution rate (d_S) using the CODEML module from the PAML software package, version 4.4.[18] To do so, we first evaluated the log likelihood of the null M1a model, which allows codon sites to exhibit variable selective pressure but no positive selection. Under M1a, ω values at different codon sites can belong to 1 of 2 categories: sites in the first category have ω values that range >0 but <1 (and are assumed to be under purifying selection), whereas sites in the second category have ω values equal to 1 (and are assumed to be neutral). We then compared the log likelihood of the M1a model with that estimated by the alternative model M1b, which allows codon sites to exhibit variable selective pressure with positive selection. Under M1b, ω values at different codon sites can belong to 1 of 3 categories: sites in the first category have ω values that range between 0 and 1 (and are assumed to be under purifying selection), sites in the second category have ω values equal to 1 (and are assumed to be neutral), and sites in the third category have ω values >1 (and are assumed to be products of positive selection).

Test for Ancient Selection in Noncoding Regions

We tested the noncoding primate data set for evidence of positive selection by estimating the ζ ratio using the batch file developed by Olivier Fedrigo[19] for use with the HyPhy software package, version 2.0.[20] The ζ ratio is similar in concept to the ω ratio. To assess whether a noncoding region is evolving under purifying selection, neutrally, or under positive selection, it compares the ratio of the substitution rate of a noncoding region ($d_{NONCODING}$) relative to the substitution rate of the synonymous sites from the adjacent coding region (d_S), which are assumed to be evolving neutrally.[21,22] Similar to the coding region tests, we first evaluated the log likelihood of the null model (noncoding M1a or ncM1a), which allows noncoding sites to exhibit variable selective pressure but no positive selection. Under ncM1a, ζ values at different sites can belong to 1 of 2 categories: sites in the first category have ζ values that are >0 but range between 0 and <1 (and are assumed to be under purifying selection), whereas sites in the second category have ζ values equal to 1 (and are assumed to be neutral). We then compared the log likelihood of the ncM1a model relative to that estimated by the alternative model (noncoding M1b or ncM1b), which allows sites to exhibit variable selective pressure with positive selection. Under ncM1b, ζ values at different sites can belong to 1 of 3 categories: sites in the first category have ζ values that range from 0 to 1 (and are assumed to be under purifying selection), sites in the second category have ζ values equal to 1 (and are assumed to be neutral), and sites in the third category have ζ values >1 (and are assumed to be undergoing positive selection).

Tests for Intraspecific Selection

Genotype data for SNPs spanning the whole gene (chr12:10 202 166–10 216 057) were downloaded from the HapMap database available at http://hapmap.ncbi.nlm.nih.gov/cgi-perl/gbrowse/hapmap24_B36/#search server. The considered populations were the Chinese (Han Chinese from Beijing, 45 unrelated Han Chinese from Beijing, China), Japanese (Japanese from Tokyo, 45 unrelated Japanese from Tokyo, Japan), Nigerian (Yoruba from Ibadan, 30 Yoruba mother–father–child trios from Ibadan, Nigeria), and European (Caucasian European, 30 mother–father–child trios from the CEPH collection Utah residents with ancestry from Northern and Western Europe) from the second release of the project. Offspring genotypes were removed to analyze only unrelated subjects. Formatted files were submitted to the Phase v2.1.1 software for haplotype reconstruction.[23–25] Analysis of

molecular variance and Tajima's *D* was then computed using Arlequin 2.0 version.[26,27] Through analysis of molecular variance, we calculated F_{st}, which is a measure of how genetic variation is partitioned within versus among populations, and therefore serves as a means to discern genetic differences among populations. SNPs used in these analyses are listed in Table VI in the online-only Data Supplement.

In a second phase of the analysis, the haplotypes were split into 2 regions: the LD block that includes intron 4 and the rest of the gene. Analyses were then repeated in these regions separately and compared (Figure 1 and Table VI in the online-only Data Supplement).

Analyses on humans were also replicated in the samples from the 1000 genomes project (Table III in the online-only Data Supplement). We used data from 55 Tuscans from Italy; 86 Great Britains from both England and Scotland; 92 Finnish from Finland; 86 Yoruba from Ibadan, Nigeria; 58 African Americans from the Southwest; 99 Southern Han Chinese; 97 Han Chinese from Beijing, China; and 88 Japanese from Tokyo, Japan. Unfortunately, only 4 sequences from the Caucasian European populations were available for this gene region and were excluded from our analyses.

Tajima's *D* was also calculated on 18 *P. troglodytes* sequences using the DNAsp software, version 5.10.01.[28] Tajima's *D* was calculated separately for different regions (overall coding region including all exons, 5′ untranslated region and exon 1, exon 2 and surrounding intronic region, exon 3 and surrounding intronic region, exon 4 and upstream region, exon 4, exon 6, and surrounding region, and 3′ untranslated region). In all cases, significance of selection was determined using a *P* value cutoff of 0.05.

Phylogenetic Analysis

We estimated the evolutionary history of both coding and noncoding regions of all sequences from all taxa using the maximum likelihood optimality criterion, as implemented in the RAxML software, version 7.2.6.[29] The maximum likelihood estimates of the phylogenies of both sets of sequences were obtained, assuming a general time reversible model of nucleotide evolution, empirically measured nucleotide frequencies, and allowing for rate heterogeneity across sites. The appropriate number of bootstrap replicates was assessed through the frequency-based stopping criterion implemented in RAxML.[30] Because the maximum likelihood estimates of the evolutionary history of both coding and noncoding regions may differ from the standard species phylogeny,[31,32] we evaluated whether each of the maximum likelihood trees was significantly different from the species phylogeny using the Shimodaira–Hasegawa test,[13] as implemented in RAxML.

Disclosures

None.

References

1. Wagner GP, Zhang J. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet*. 2011;12:204–213.
2. Sawamura T, Kume N, Aoyama T, Moriwaki H, Hoshikawa H, Aiba Y, et al. An endothelial receptor for oxidized low-density lipoprotein. *Nature*. 1997;386:73–77.
3. Jeannin P, Bottazzi B, Sironi M, Doni A, Rusnati M, Presta M, et al. Complexity and complementarity of outer membrane protein A recognition by cellular and humoral innate immunity receptors. *Immunity*. 2005;22:551–560.
4. Mango R, Clementi F, Borgiani P, Forleo GB, Federici M, Contino G, et al. Association of single nucleotide polymorphisms in the oxidised LDL receptor 1 (OLR1) gene in patients with acute myocardial infarction. *J Med Genet*. 2003;40:933–936.
5. Mango R, Biocca S, del Vecchio F, Clementi F, Sangiuolo F, Amati F, et al. *In vivo* and *in vitro* studies support that a new splicing isoform of OLR1 gene is protective against acute myocardial infarction. *Circ Res*. 2005;97:152–158.
6. Biocca S, Falconi M, Filesi I, Baldini F, Vecchione L, Mango R, et al. Functional analysis and molecular dynamics simulation of LOX-1 K167N polymorphism reveal alteration of receptor activity. *PLoS One*. 2009;4:e4648.
7. Biocca S, Filesi I, Mango R, Maggiore L, Baldini F, Vecchione L, et al. The splice variant LOXIN inhibits LOX-1 receptor function through hetero-oligomerization. *J Mol Cell Cardiol*. 2008;44:561–570.
8. Cheng Y, Wei Y, Li W, Chen J, Zhang W, Hui R, et al. Associations between oxidized-lipoprotein receptor 1 G501C and 3′-UTR-C188T polymorphisms and coronary artery disease: a meta-analysis. *Cardiology*. 2011;119:90–95.
9. Yang Z, Wong WS, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection. *Mol Biol Evol*. 2005;22:1107–1118.
10. Klein J, Sato A, Nagl S, O'hUigin C. Molecular trans-species polymorphism. *Annu Rev Ecol Syst*. 1998;29:1–21.
11. Rokas A, Carroll SB. Bushes in the tree of life. *PLoS Biol*. 2006;4:e352.
12. Muirhead CA, Glass NL, Slatkin M. Multilocus self-recognition systems in fungi as a cause of trans-species polymorphism. *Genetics*. 2002;161:633–641.
13. Shimodaira H, Hasegawa M. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol*. 1999;16:1114–1116.
14. Carlson CS, Thomas DJ, Eberle MA, Swanson JE, Livingston RJ, Rieder MJ, et al. Genomic regions exhibiting positive selection identified from dense genotype data. *Genome Res*. 2005;15:1553–1565.
15. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–3066.
16. Katoh K, Toh H. Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform*. 2008;9:286–298.
17. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 2006;34:W609–W612.
18. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24:1586–1591.
19. Haygood R, Fedrigo O, Hanson B, Yokoyama KD, Wray GA. Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. *Nat Genet*. 2007;39:1140–1144.
20. Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*. 2005;21:676–679.
21. Wong WS, Nielsen R. Detecting selection in noncoding regions of nucleotide sequences. *Genetics*. 2004;167:949–958.
22. Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol*. 2005;22:2472–2479.
23. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*. 2001;68:978–989.
24. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet*. 2003;73:1162–1169.
25. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet*. 2005;76:449–462.
26. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol*. 1995;12:921–927.

27. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992;131:479–491.

28. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25:1451–1452.

29. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22:2688–2690.

30. Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A. How many bootstrap replicates are necessary? *J Comput Biol*. 2010;17:337–354.

31. Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. Molecular phylogenetics and the origins of placental mammals. *Nature*. 2001;409:614–618.

32. Tosi AJ, Morales JC, Melnick DJ. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution*. 2003;57:1419–1435.

CLINICAL PERSPECTIVE

The region spanning intron 4 to the 3′ untranslated region of the lectin-like oxidized lipoprotein receptor 1 gene (*OLR1*) is under divergent patterns of selection in different primates, with positive selection in nonhuman primates, especially chimpanzees, and balancing selection in humans, with different patterns in different populations. Intron 4 sequences regulate alternative splicing, especially loss of exon 5 and production of a truncated isoform, called LOXIN. The proteins encoded by the 2 transcripts have opposing roles in disease because lectin-like oxidized lipoprotein receptor 1 binds both oxidized low-density lipoprotein and bacterial cell wall proteins, the first favoring cardiovascular disease and the second protecting from infection, whereas LOXIN binds neither. A lower LOXIN/lectin-like oxidized lipoprotein receptor 1 ratio has been reported in subjects with coronary artery disease compared with controls, but no clear correlation between *OLR1* polymorphisms and disease was found in >20 genetic epidemiological studies, leading to uncertainty about the role of *OLR1* in human atherosclerosis. Our work provides evidence that genes related to multiple complex traits can be subject to different and even antagonistic selection pressures depending on environmental variation and that this phenomenon underlies our inability to detect or validate strong genotype–disease associations. Therefore, biological function may be shrouded by complicated evolutionary histories that will need to be considered before dismissing the role of genes affecting disease because of lack of validation in different studies.

SUPPLEMENTAL MATERIAL

Supplemental Figure 1: Phylogenetic trees of coding (a) region and intron 4 (b)
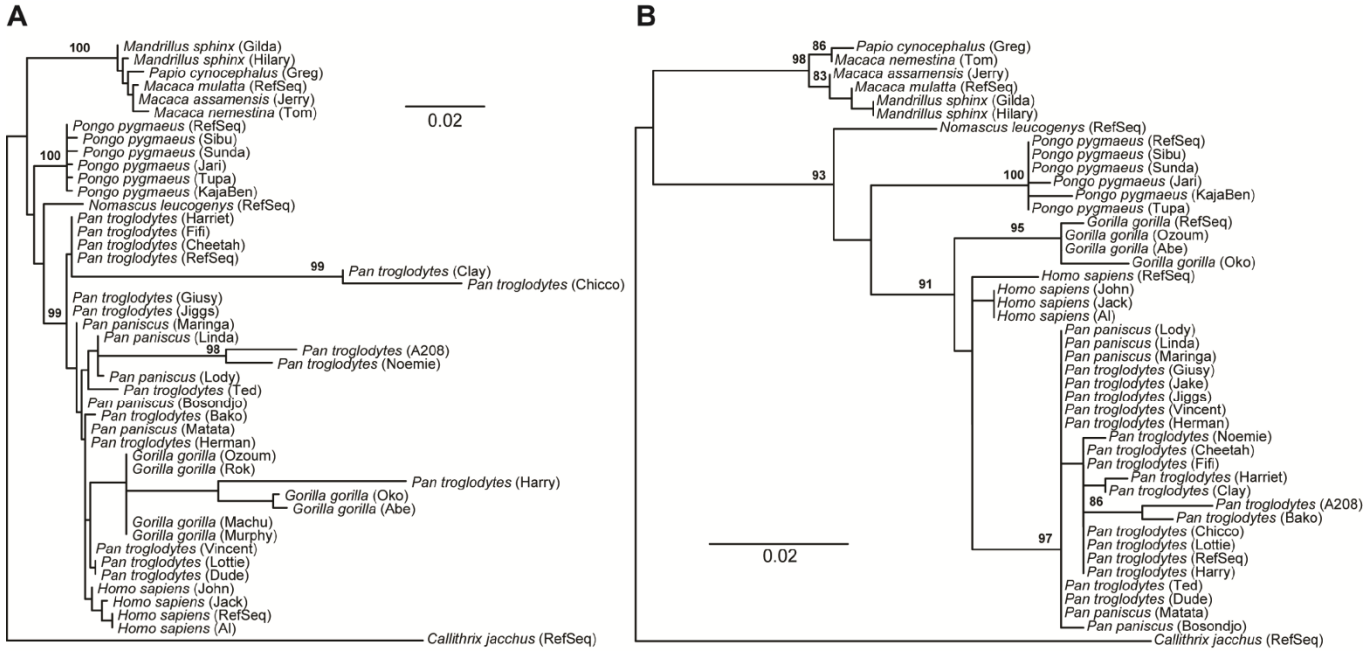
Panel A: coding sequences
Panel B: non-coding sequences
Bootstrap values < 80% not shown

Supplemental Figure 2: schematic of the analyzed regions of *OLR1* gene

Supplemental Table S1. List of association studies on *OLR1* SNPs and disease (CAD: coronary artery disease, AMI: acute myocardial infarction, ACS: acute coronary syndrome)

Reference	Phenotype	Study design	Sample number	GCAAGC haplotype	N (Asn) variant
Luedecking-Zimmer et al 2002 [1]	Alzheimer	Case control	800 vs 700	Dependent on APOE*4[1]	---
Lambert et al 2003 [2]	Alzheimer	Case control	589 vs 663	Risk	---
Bertram et al 2004 [3]	Alzheimer	Family based	437 families (1439 subjects)	No association	---
Colacicco et al 2009 [4]	Alzheimer	Case control	169 vs 264	Risk	---
Mango et al 2003 [5]	CAD, AMI	Case control	103 vs 150	Protective	Protective
Tatsuguchi et al 2003 [6]	AMI	Case control	102 vs 102	---	Risk
Ohmori et al 2004 [7]	CAD severity	Case control	167 vs 39 vs 419	---	Protective
Sentinelli et al 2006 [8]	CAD	Case control	351 vs 251	No association	---
Knowles et al 2008 – ADVANCE [9]	AMI	Case control	1200 vs 1045	Protective	Protective
Knowles et al 2008 – ARIC [9]	AMI	Case control	1156 vs 8702	Risk	---
Trabetti et al 2006 [10]	CAD, AMI	Case control	190 vs 160 vs 327	No association	Protective
Puccetti et al 2007 [11]	Statin effectiveness	Cohort	751	Protective	---
Puccetti et al 2011 [12]	Statin effectiveness	Cohort	1039	Protective	---
Morgan et al 2007 [13]	AMI mortality	Cohort	811	---	No association
Chen et al 2003 [14]	Stenosis gravity	Case-control	227 vs 150 vs 206	Protective	---
Wang et al 2010 [15]	ACS	Case-control	198 vs 204	Risk	No association
Hou et al 2008 [16]	Hypertension	Case-control	280 vs 284	---	Risk
Hattori et al 2006 [17]	Ischemic stroke	Case-control	274 vs 274	---	No association
Wang et al 2011 [18]	Carotid plaque phenotypes	Case-control	134 vs 153	---	Risk
Predazzi et al 2012 [19]	IMT	Cohort	2200	---	Risk
Sciacqua et al 2012[20]	Hypertension	Case-control	178 vs 36	Protective	---
Gu et al 2012[21]	Conventional factors for CAD	Cohort	1075	No association	---

Supplemental Table S2. Ancient inter-specific selection on intron 4 (Neg = Negative selection, Neu = Neutrality, Pos = Positive selection)

	Bayes Empirical Algorithm[1]			Naïve Empirical Bayes Algorithm[1]			Intron 4 sequence											
Site	Neg	Neu	Pos	Neg	Neu	Pos	*H. sapiens*	*P. troglodytes*	*P. paniscus*	*G. gorilla*	*P. pygmaeus*	*N. leucogenys*	*M. sphinx*	*P. cynocephalus*	*M. mulatta*	*M. nemestina*	*M. assamensis*	*C. jacchus*
46	0.047	0.428	0.526*	0	0.438	0.562*	T	T	T	T	T	C	C	C	C	C	C	T
55	0.047	0.427	0.526*	0	0.433	0.567*	T	T	T	T	T	T	G	G	G	G	G	A
112	0.004	0.068	0.927*	0	0.054	0.946*	T	T	T	A	G	G	T	T	T	T	T	G
180	0.036	0.327	0.636*	0	0.321	0.680*	C	C	C	C	G	C	G	G	G	G	G	-
207	0.003	0.056	0.941*	0	0.046	0.954*	G	A	A	G	A	G	A	A	A	A	A	-
219	0.046	0.402	0.553*	0	0.404	0.596*	T	T	T	T	T	T	T	C	T	C	T	T
224	0.001	0.014	0.986*	0	0.010	0.991*	G	G	G	A	G	G	A	G	A	G	A	G
312	0.054	0.471	0.474*	0	0.485	0.515*	A	A	A	A	A	A	C	A	C	A	A	A

[1]posterior probability of type of selection

Supplemental Table S3. Results of intra-species analyses in human samples from the 1000 genomes project (TSI, Toscani in Italia; GBR, British in England and Scotland; FIN, Finnish in Finland; YRI, Yoruba in Ibadan, Nigeria; PUR, Puerto Ricans from Puerto Rico; ASW, African American Southwest; CHS, Southern Han Chinese; CHB, Han Chinese in Bejing, China; JPT, Japanese in Tokyo, Japan). Significant reusults are in RED.

| Variant | Region | Tajima's D | | | | | | | | |
| | | European Ancestry | | | Admixed Ancestry | | Western African Ancestry | Asian Ancestry | | |
		TSI	GBR	FIN	PUR	ASW	YRI	CHS	CHB	JPT
rs2171602	LD block	**2.080***	1.117	1.601	0.997	0.709	0.15	1.494	1.443	1.34
rs7964641	LD block	0.008	0.195	-0.136	-0.137	-0.442	-0.126	1.054	1.017	-0.132
rs10505755	LD block	1.546	0.71	1.401	1.317	-0.429	-0.68	0.4724	0.435	0.345
rs2634156	LD block	**2.283***	**2.158***	**2.097***	0.933	0.123	-0.325	1.557	0.914	1.311
rs3816844	LD block	**2.685***	**2.614***	1.893	**2.414***	0.863	-0.081	0.539	1.499	1.311
rs34807991	LD block	1.447	0.889	0.903	1.274	-0.351	-0.608	1.27	1.245	1.199
rs35146920	LD block	**2.525***	1.604	1.651	1.042	0.435	-0.273	1.464	1.41	1.235
rs2634159	LD block	0.279	1.163	1.093	0.039	-0.549	1.092	0.833	0.846	0.18
rs34970770	LD block	1.093	1.635	0.661	0.09	0.058	-0.6251	1.464	1.411	1.235
chr12:10317633[1]	LD block	-0.794	-0.402	-1.254	-1.201	-1.386	-1.212	0	0	-1.109
rs35827587	Non LD block	0	-1.215	-1.215	-0.96	-0.9	-0.745	-1.029	-1.031	-1.042
rs34163097	Non LD block	0.035	0.434	0.388	-1.011	-0.63	-0.291	1.362	0.821	1.56
rs35567600	Non LD block	-1.092	-0.977	-0.974	-1.135	-1.336	-1.07	0	-1.092	-1.042
rs35492478	Non LD block	-0.21	0.499	0.569	0.298	-0.006	0.502	**2.406***	0.92	0.895
chr12:10321197[1]	Non LD block	0.179	-0.385	-0.179	-0.59	-0.186	0.627	1.781	1.774	-0.568
rs16910917	Non LD block	-0.736	-0.832	-0.787	-0.603	-0.729	0.19	-1.485	-1.431	-1.497
chr12:10322484[1]	Non LD block	-0.51	1.238	1.558	-0.592	-1.081	-0.757	-0.504	-0.513	1.773
chr12:10323156[1]	Non LD block	0.231	1.059	0.587	0.276	-0.31	0.781	1.195	**2.302***	**2.246***
chr12:10324057[1]	Non LD block	-0.005	0.085	-0.131	-0.738	-1.231	-1.185	-0.157	0.468	-0.126
rs35553961	Non LD block	0.212	0.126	-0.416	-0.464	-1.051	-0.694	0.299	-0.199	0.26
chr12:10325077[1]	Non LD block	0	0	-1.037	-1.393	-1.633	-1.701	0	0	0
rs2742112	Non LD block	1.284	1.61	-0.225	0.255	0.559	-0.148	0.472	0.469	0.432

*Sites that show evidence of balancing selection

[1](Human genome Reference Sequence NCBI36)[22].

Supplemental Table S4. List of species and samples used for sequence analyses

Sample Number (n)		*Species*	Time since last common ancestor
5 (5 de novo[1])	*Bonobo*	*Pan paniscus*	2 MYA[2]
18 (17 de novo + Ensembl Refseq)	*Chimpanzee*	*Pan troglodytes*	
4 (3 de novo + Ensembl Refseq)	*Human*	*Homo sapiens*	6 MYA
3 (3 de novo)	*Gorilla*	*Gorilla gorilla*	7 MYA
7 (6 de novo + Ensembl Refseq)	*Orang*	*Pongo pygmaeus*	14 MYA
1 (Ensembl Refseq)	*Gibbon*	*Nomascus leucogenys*	18 MYA
1 (Ensembl Refseq)	*Macaque*	*Macaca mulatta*	25 MYA
1 (de novo)	*Macaque*	*Macaca nemestina*	
1 (de novo)	*Macaque*	*Macaca assamensis*	
2 (2 de novo)	*Mandrill*	*Mandrillus sphinx*	
1 (1 de novo)	*Baboon*	*Papio cynocephalus*	
1 (Ensembl Refseq)	*Marmoset*	*Callithrix jacchus*	40 MYA

[1]Sequenced for this project as denoted in Supplemental Figure 3
[2]MYA equals millions of years ago

Supplemental Table S5. List of primers used for amplification and sequencing reactions and amplicon characteristics

Primer Name	Region Amplified	Sequence	Amplicon size	Annealing T
F1	Exon 1	ATTCTAAAATCACCCAGGAC	776	56
R1		TCCCATACTTGGGTGTTTAG		
F2		CAGGCTCTCTGCACTGTGGA	1213	57
R2	Exon 2	GCCTCTCAGATGATCTTGGA		
R2 - Macaques		GCCTCTCAGATGATCTTGGA		
F3		CCTGGAGGGATTTTTATTTGG	998	57
F3 - Macaques	Exon 3	CCTGGAGGGATTTTTATCTGG		
R3		GCCTGTTCAATATGGCAAAAC		
R3 - Macaques		GCCTGTTCAATATGGTAAAAC		
F4 - Humans		TTCTCCCAACCCTCACTTTC	1042	56
F4 - Chimps	Exon 4, Intron 4 and	TTCTCCCAGCCCTCACTTTC		
F4 - Orangs	Exon 5	TTCCTCCAACCCTCACTTTC		
F4 - Macaques		TTCCCCCAACCCTCACTTTC		
R4		GTGGGGCATCAAAGGAGAA		
F5	Exon 4, Intron 4 and	AGGAATTCTTGTTTGAGAG		
R5	Exon 5 (nested)	CTCTCAAACAAGAATTCCT		
F6	Exon 6 Forward	CCCATTCTTTTTGATGCCTC	1005	55
R6	Exon 6 Reverse	AGTGGTAGAGTCTGGAGATG		

Supplemental Table S6. List of Hapmap SNPs used for the analysis

SNP	REGION	CEU		CHB		JPT		YRI		ANALYSIS
rs11615002	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block[1]
rs11609310	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs12231578	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs12231585	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs3994135	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs11053645	3'UTR	0.1	0.9	0.2	0.8	0.2	0.8	0.2	0.8	LD Block
rs7964641	3'UTR	0.9	0.1	1	0	0.9	0.1	0.4	0.6	LD Block
rs1050286	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs1050283	3'UTR	0.5	0.5	0.8	0.2	0.8	0.2	0.8	0.2	LD Block
rs1050289	3'UTR	0.1	0.9	0.1	0.9	0	1	0	1	LD Block
rs12309394	Int 5	1	0	1	0	1	0	0.8	0.2	LD Block
rs2634156	Int 5	0	1	0	1	0	1	0.1	0.9	LD Block
rs3816844	Int 5	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs3736235	Int 4	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs3736234	Int 4	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs3736233	Int 4	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs3736232	Int 4	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs11053646	ex4 (K167N)	0.9	0.1	0.8	0.2	0.8	0.2	0.8	0.2	LD Block
rs12822177	Int 3	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	LD Block
rs11053648	Int 3	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	Not in LD Block
rs2742109	Int 3	1	0	1	0	1	0	1	0	Not in LD Block
rs6488265	Int 3	0.4	0.6	0.8	0.2	0.8	0.2	0.5	0.5	Not in LD Block
rs11053649	Int 3	0.5	0.5	0.8	0.2	0.8	0.2	0.9	0.1	Not in LD Block
rs2010655	Int 3	0.6	0.4	0.5	0.5	0.4	0.6	0.4	0.6	Not in LD Block
rs11611438	Int 3	0.1	0.9	0.2	0.8	0.2	0.8	0	1	Not in LD Block
rs11611453	Int 3	0.1	0.9	0.2	0.8	0.2	0.8	0	1	Not in LD Block
rs3912640	Int 2	0.3	0.7	0.5	0.5	0.6	0.4	0.6	0.4	Not in LD Block
rs3741860	Int 2	1	0	1	0	1	0	1	0	Not in LD Block
rs2742118	Int 2	0	1	0	1	0	1	0	1	Not in LD Block
rs2742117	Int 2	1	0	1	0	1	0	1	0	Not in LD Block
rs2634161	Int 2	1	0	1	0	1	0	1	0	Not in LD Block
rs12316417	Int 2	1	0	1	0	1	0	1	0	Not in LD Block
rs2742116	Int 2	1	0	1	0	1	0	1	0	Not in LD Block
rs2742115	Int 1	0.7	0.3	0.5	0.5	0.4	0.6	0.8	0.2	Not in LD Block
rs2742114	Int 1	0.7	0.3	0.5	0.5	0.4	0.6	0.4	0.6	Not in LD Block
rs2742113	Int 1	0.3	0.7	0.5	0.5	0.6	0.4	0.6	0.4	Not in LD Block
rs2634162	Int 1	0.7	0.3	0.5	0.5	0.4	0.6	0.8	0.2	Not in LD Block
rs11053654	Int 1	0.5	0.5	0.2	0.8	0.2	0.8	0.1	0.9	Not in LD Block
rs2634163	5'UTR	1	0	1	0	1	0	1	0	Not in LD Block
rs12308385	5'UTR	0	1	0	1	0	1	0	1	Not in LD Block
rs2742111	5'UTR	1	0	1	0	1	0	1	0	Not in LD Block

[1] LD block as shown in Figure 1 (main text).

1. Luedecking-Zimmer E, DeKosky ST, Chen Q, Barmada MM, Kamboh MI. Investigation of oxidized ldl-receptor 1 (olr1) as the candidate gene for alzheimer's disease on chromosome 12. *Hum Genet*. 2002;111:443-451

2. Lambert JC, Luedecking-Zimmer E, Merrot S, Hayes A, Thaker U, Desai P, et al. Association of 3'-utr polymorphisms of the oxidised ldl receptor 1 (olr1) gene with alzheimer's disease. *J Med Genet*. 2003;40:424-430

3. Bertram L, Parkinson M, Mullin K, Menon R, Blacker D, Tanzi RE. No association between a previously reported olr1 3' utr polymorphism and alzheimer's disease in a large family sample. *J Med Genet*. 2004;41:286-288

4. Colacicco AM, Solfrizzi V, D'Introno A, Capurso C, Kehoe PG, Seripa D, et al. Alpha-2-macroglobulin gene, oxidized low-density lipoprotein receptor-1 locus, and sporadic alzheimer's disease. *Neurobiol Aging*. 2009;30:1518-1520

5. Mango R, Clementi F, Borgiani P, Forleo GB, Federici M, Contino G, et al. Association of single nucleotide polymorphisms in the oxidised ldl receptor 1 (olr1) gene in patients with acute myocardial infarction. *J Med Genet*. 2003;40:933-936

6. Tatsuguchi M, Furutani M, Hinagata J, Tanaka T, Furutani Y, Imamura S, et al. Oxidized ldl receptor gene (olr1) is associated with the risk of myocardial infarction. *Biochem Biophys Res Commun*. 2003;303:247-250

7. Ohmori R, Momiyama Y, Nagano M, Taniguchi H, Egashira T, Yonemura A, et al. An oxidized low-density lipoprotein receptor gene variant is inversely associated with the severity of coronary artery disease. *Clin Cardiol*. 2004;27:641-+

8. Sentinelli F, Filippi E, Fallarino M, Romeo S, Fanelli M, Buzzetti R, et al. The 3 '-utr c > t polymorphism of the oxidized ldl-receptor 1 (olr 1) gene does not associate with coronary artery disease in italian cad patients or with the severity of coronary disease. *Nutr Metab Cardiovas*. 2006;16:345-352

9. Knowles JW, Assimes TL, Boerwinkle E, Fortmann SP, Go A, Grove ML, et al. Failure to replicate an association of snps in the oxidized ldl receptor gene (olr1) with cad. *BMC Med Genet*. 2008;9:23

10. Trabetti E, Biscuola M, Cavallari U, Malerba G, Girelli D, Olivieri O, et al. On the association of the oxidised ldl receptor 1 (olr1) gene in patients with acute myocardial infarction or coronary artery disease. *Eur J Hum Genet*. 2006;14:127-130

11. Puccetti L, Pasqui AL, Bruni F, Pastorelli M, Ciani F, Palazzuoli A, et al. Lectin-like oxidized-ldl receptor-1 (lox-1) polymorphisms influence cardiovascular events rate during statin treatment. *Int J Cardiol*. 2007;119:41-47

12. Puccetti L, Scarpini F, Cappellone R, Auteri A. Genetic influence in statin intolerance. *Clin Pharmacol Ther*. 2011;90:365

13. Morgan TM, Krumholz HM, Lifton RP, Spertus JA. Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. *Jama*. 2007;297:1551-1561

14. Chen Q, Reis SE, Kammerer C, Craig WY, LaPierre SE, Zimmer EL, et al. Genetic variation in lectin-like oxidized low-density lipoprotein receptor 1 (lox1) gene and the risk of coronary artery disease. *Circulation*. 2003;107:3146-3151

15. Wang Y, Rao L, Zhou B, Chen Y, Peng Y, Song Y, et al. The g501c polymorphism of the oxidized low-density lipoprotein-receptor 1 gene is associated with acute coronary syndrome in the han chinese population. *DNA Cell Biol*. 2010;29:201-205

16. Hou XW, Wang LF, Wang NF, Pang DT, Hui B, Zhou YI, et al. The g501c polymorphism of oxidized ldl receptor gene [olr-1] is associated with susceptibility and serum c-reactive protein concentration in chinese essential hypertensives. *Clin Chim Acta*. 2008;388:200-203

17. Hattori H, Sonoda A, Sato H, Ito D, Tanahashi N, Murata M, et al. G501c polymorphism of oxidized ldl receptor gene (olr1) and ischemic stroke. *Brain Res*. 2006;1121:246-249

18. Wang LY, Yanuck D, Beecham A, Gardener H, Slifer S, Blanton SH, et al. A candidate gene study revealed sex-specific association between the olr1 gene and carotid plaque. *Stroke*. 2011;42:588-592

19. Predazzi IM, Norata GD, Vecchione L, Garlaschelli K, Amati F, Grigore L, et al. Association between *OLR1* k167n snp and intima media thickness of the common carotid artery in the general population. *Plos One*. 2012;7:e31086

20. Sciacqua A, Presta I, Perticone M, Tassone EJ, Andreozzi F, Quitadamo MC, et al. 3'-utr olr1/lox-1 gene polymorphism and endothelial dysfunction: Molecular and vascular data in never-treated hypertensive patients. *Intern Emerg Med*. 2012

21. Gu Y, Liu Z, Li L, Guo CY, Li CJ, Wang LS, et al. Olr1 , pon1 and mthfr gene polymorphisms, conventional risk factors and the severity of coronary atherosclerosis in a chinese han population. *Cell Physiol Biochem*. 2013;31:143-152

22. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56-65