Polygenic Link between Blood Lipids and Amyotrophic Lateral Sclerosis

Xu Chen, BMed, MSc[1]

Solmaz Yazdani, MSc[1]

Fredrik Piehl, MD, PhD[2]

Patrik K.E. Magnusson, PhD[1]

Fang Fang, MD, PhD[1]

[1] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 17177 Stockholm, Sweden;
[2] Department of Clinical Neurosciences, Karolinska Institutet, 171 77 Stockholm, Sweden

Corresponding author:
Dr. Fang Fang
Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, PO Box 281
171 77 Stockholm, Sweden.
Phone: +46-8-52486131; Fax: +46-8-314975
E-mail: Fang.Fang@ki.se

Running head: Polygenic link between lipids and ALS

Number of tables: 1

Number of figures: 2

Acknowledgement statement (including conflict of interest and funding sources):

We would like to thank the following consortia for sharing their summary GWAS results: the Global Lipids Genetics Consortium, CARDIoGRAMplusC4D Consortium and Project MinE GWAS Consortium. The authors declare no conflicts of interest. This study was funded by the Swedish Research Council (grant no. 2015-03170), the Karolinska Institutet (Senior Researcher Award and Strategic Research Program in Epidemiology), and the Ulla-Carin Lindquist Foundation.

ABSTRACT

Dyslipidemia is common among patients with amyotrophic lateral sclerosis (ALS), but the genetic evidence for the link between blood lipids and ALS is unclear. We assessed the associations of polygenic risk scores (PRS) for low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC), triglycerides (TG) and high-density lipoprotein cholesterol (HDL-C) with ALS, using summary results from genome-wide association studies. PRS of LDL-C (OR=1.17; 95%CI 1.11-1.23) and TC (OR=1.15; 95%CI 1.09-1.21) were significantly associated with a higher risk of ALS, whereas no association was noted for PRS of TG and HDL-C. Although linkage disequilibrium score regression indicated weak genetic correlations between lipids and ALS, several genetic variants known to up-regulate LDL-C and TC levels were shown to be associated with a higher risk of ALS. This study demonstrates a clear polygenic link between LDL-C, TC and ALS, and suggests that inherited dyslipidemia might be an etiological part of ALS development.

KEY WORDS

Blood lipids, amyotrophic lateral sclerosis, genome-wide association study, polygenic risk score

INTRODUCTION

Dyslipidemia is common among patients with amyotrophic lateral sclerosis (ALS).[1] Higher-than-expected levels of low-density lipoprotein cholesterol (LDL-C), total cholesterol (TC), triglycerides (TG) and low/high-density lipoprotein ratio (LDL-C/HDL-C) have been reported among patients with ALS by some but not all studies.[2-7] In a recent observational study we found that patients with ALS had increased levels of apolipoprotein B and LDL-C up to 20 years before diagnosis.[8] Because muscle atrophy is a contributing factor to metabolic dysregulation and little is known on the earliest onset of muscle atrophy in ALS, it is unclear whether dyslipidemia plays a causal role in ALS or represents purely an early non-motor sign of ALS.

Genetic loci regulating blood LDL-C, TC, TG and HDL-C have been identified in previous genome-wide association studies (GWAS),[9,10] and single nucleotide polymorphisms (SNPs) derived polygenic risk scores (PRS) have been shown to be predictive of circulating lipid trajectories.[11] To this end, we aimed to test the associations of LDL-C, TC, TG and HDL-C, indicated by their corresponding SNPs derived PRS, with the risk of ALS, using the recently published GWAS result of ALS.[12] As a positive comparison, we also conducted a similar analysis for the association of LDL-C with coronary artery disease (CAD).[13,14]

RESULTS

PRS of LDL-C and TC were both significantly associated with ALS risk (**Table**). The associations were similar when using either European-ancestry or trans-ethnic GWAS results of lipids, although the best-fitted P_Ts were lower and the numbers of SNPs included in the best-fitted quantiles were smaller when using European-ancestry GWAS results of lipids. In the analyses using European-ancestry GWAS results of lipids, the best-fitted P_Ts were $5x10^{-5}$ for both LDL-C and TC to predict

ALS, with odds ratio (OR) of 1.17 (95% confidence interval [CI] 1.11-1.23, $P=8.82x10^{-8}$) and 1.15 (95% CI 1.09-1.21, $P=8.17x10^{-7}$) per standard deviation of the PRS for LDL-C and TC, respectively. SNPs in the best-fitted P_T quantiles explained 0.08% (LDL-C) and 0.07% (TC) of the variance of ALS. In contrast, PRS of TG and HDL-C were not associated with ALS.

In the reverse PRS analyses, in which ALS was used as the base whereas LDL-C, TC, TG and HDL-C were used as the targets, no clear association was noted for any of the studied lipids.

Statistically significant associations were noted in all bi-directional analyses of LDL-C and CAD ($P<9x10^{-37}$). In the best-fitted P_T quantiles of LDL-C, less than 500 SNPs explained about 0.25~0.60% variance of CAD, and ORs for PRS of LDL-C on CAD were around 1.45~1.56. The associations for PRS of CAD with LDL-C were weaker but still statistically significant.

Using both European-ancestry and trans-ethnic GWAS results of lipids, the associations between PRS of LDL-C and TC with ALS were similar and statistically significant across quantiles with $P_T \leq 0.005$ (**Figure 1A-D**). Similar results were noted for the quantiles including only genome-wide significance SNPs ($P_T \leq 5x10^{-8}$), the variance of ALS explained by these SNPs was very close to the largest variance explained by the best-fitted quantile (**Figure 1E-F**).

LD score (LDSC) regression was performed to calculate the heritability and genetic correlation between the base and target.[15,16] The genetic correlations (rg) between blood lipids and ALS were weak ($rg \leq 0.06$) and not statistically significant ($P \geq 0.47$) (data not shown). The genetic correlation between trans-ethnic LDL-C and CAD was moderate and statistically significant (rg=0.21, 95%CI 0.13-0.29, $P=1.22x10^{-6}$). The SNP-based univariate heritability of LDL-C, CAD and ALS estimated from LDSC regression was 0.20 (95%CI 0.11-0.29), 0.10 (95%CI 0.08-0.12) and 0.05 (95%CI 0.03-0.07), respectively.

In the Mendelian randomization (MR) analysis, 62 of the 63 SNPs specifically associated with LDL-C or TC were found in the ALS GWAS. The combined effect of these SNPs on ALS was null (OR=1.00, 95%CI 1.00-1.01, P=0.13). Some individual SNPs were however significantly associated with ALS risk (P<0.05, **Figure 2**). Among the alleles up-regulating LDL-C or TC, five were associated with a higher ALS risk, including alleles in the loci close to very-low-density lipoprotein receptor (*VLDLR*), Niemann-Pick C1-like protein 1 (*NPC1L1*), glycerol-3-phosphate acyltransferase 1, Mitochondrial (*GPAM*), LDL receptor (*LDLR*) and topoisomerase I (*TOP1*) genes.

DISCUSSION

Using different polygenic methods on summary GWAS results of blood lipids and ALS, the present study demonstrated a clear polygenic link between LDL-C, TC and ALS. We found that PRS of genome-wide significant SNPs that are known to increase LDL-C and TC levels were associated with a higher risk of ALS. Several genetic variants involved in the regulation of LDL-C and TC were further shown to be associated with the risk of ALS.

The recent finding that ALS patients had consistently elevated levels of apolipoprotein B and LDL-C during the 20 years before diagnosis suggests that dyslipidemia might either be a risk factor for ALS or a prodromal non-motor symptom of ALS.[8] The present findings support the former notion by demonstrating clear genetic evidence that dyslipidemia, in terms of high LDL-C and TC levels specifically, is not purely secondary to ALS but represents an independent risk factor for ALS. Among the 63 genome-wide significant loci specifically associated with LDL-C or TC, five showed clear associations with a higher risk of ALS (P<0.05, **Figure 2**). The fact that these loci are either involved in the cholesterol transportation (*VLDLR, LDLR, NPC1L1*) or lipid modification (*GPAM*)

processes, sheds light on the potential cellular pathways underlying the link between dyslipidemia and ALS. More specifically, mutations in the *NPC1L1* gene are also associated with Niemann-Pick type C disease, which is an inherited metabolic disorder with accumulation of LDL-C and sphingomyelin in the viscera and central nervous system and shares phenotypic features with ALS.[17]

Genetic correlations between blood lipids and ALS were weak and non-significant when estimated by LDSC regression. Similar weak correlations have been reported between ALS and other neurological and psychiatric diseases, including Alzheimer's disease, attention deficit-hyperactivity disorder, bipolar disorder and multiple sclerosis.[18] One potential exception is schizophrenia, demonstrating a genetic correlation of 0.14 with ALS. PRS of schizophrenia was also reported to predict ALS risk, with a best P_T of 0.20. This best P_T quantile included 189,492 SNPs, and explained 0.12% of ALS variance.[18] The SNP-based heritability of ALS from the GWAS data we used was 7%, whereas previous family and twin studies reported a heritability of ALS as high as 60%.[19] Potential explanations for this large difference might include the smaller sample size of ALS GWAS in comparison to GWAS of lipids or CAD, and also potentially due to the special polygenic rare variants architecture of ALS.[12]

Unlike several more common complex diseases (e.g., CAD, schizophrenia), low-frequency or rare alleles have been suggested to explain more heritability of ALS.[12] Because GWAS focus on common SNPs and usually require minor allele frequency to be larger than 1% in the population, they cannot capture potential effects of rare variants. Family and twin studies on the other hand capture all genetic effects from both common and rare variants. Figures 1A-1D illustrate that the explained variance of ALS diminished as more SNPs were included, and the associations of the corresponding PRS with ALS became weaker and not statistically significant. Therefore, deep sequencing for rare variants in future ALS studies might be essential, to understand the genetic

mechanisms of metabolic dysregulation and ALS. In parallel, molecular studies are also needed to shed light on pathogenic mechanisms of the suggested cellular pathways related to lipid processing and ALS risk.

MATERIALS AND METHODS

Data

All data used in this study were summary GWAS results of blood lipids, ALS and CAD, in both European-ancestry and trans-ethnic populations. GWAS results on LDL-C, TC, TG and HDL-C were accessed from the Global Lipids Genetics Consortium, including ~100,000 European-ancestry (95,454 for LDL-C; 100,184 for TC; 96,598 for TG; and 99,900 for HDL-C; 2.69 million SNPs) and 188,577 trans-ethnic (2.44 million SNPs) individuals.[9,10] GWAS results on ALS were accessed from the Project MinE GWAS Consortium, including 36,052 European-ancestry individuals (N_{case}=12,577; 8.71 million SNPs).[12] GWAS results on CAD were accessed from the CARDIoGRAMplusC4D Consortium, including 86,995 European-ancestry (N_{case}=22,233; 2.42 million SNPs) and 124,305 trans-ethnic (N_{case}=60,801; 9.46 million SNPs) individuals.[13,14]

Polygenic analyses

In order to check the directionality of the association between PRS for lipids and ALS, we performed bi-directional PRS analyses using PRSice v1.25,[20] in which ALS caseness was used first as target (outcome) and then as base (exposure). In the base data, we clumped linkage disequilibrium (LD) linked SNPs by using HapMap_ceu_all genotype as the reference (release 22, 60 individuals, 3.96 million SNPs), and used the parameter settings recommended in PRSice; a clumping threshold of p1=p2=0.5, a LD threshold of r^2=0.05 and a distance threshold of 300Kb. We grouped

independent SNPs into different quantiles with gradually increasing P value thresholds (P_T, ranging from 0 to 0.5). We standardized PRS of each quantile before we calculated their effects on the target. The P_T of the quantile that explained most variance of the target was defined as the best-fitted P_T.

We performed LDSC regression by LDSC v1.0.0 to calculate the heritability and genetic correlation between base and target, using information from all available (directly genotyped and imputed) SNPs.[15,16] Finally, we conducted a MR analysis taking advantage of information on 63 loci that are known to be specifically associated with LDL-C or TC (the top SNP in each genome-wide significant locus, $P<5x10^{-8}$), but not other lipids. The 63 SNPs were considered as instrumental variables; their individual and combined effects on ALS were calculated and plotted using "metafor" package in R.3.2.3.

ACKNOWLEDGMENTS

We would like to thank the Global Lipids Genetics Consortium, the CARDIoGRAMplusC4D Consortium , and the Project MinE GWAS Consortium for sharing their summary GWAS results.

COMPETING INTERESTS

The authors declare no conflicts of interest.

FUNDING

This study was funded by the Swedish Research Council (grant no. 2015-03170), the Karolinska Institutet (Senior Researcher Award and Strategic Research Program in Epidemiology), and the Ulla-Carin Lindquist Foundation.

REFERENCES

1. Albert, S.M. Dyslipidemia in ALS: good, bad, or unclear? *Neurology* **70**, 988-9 (2008).
2. Dupuis, L. *et al.* Dyslipidemia is a protective factor in amyotrophic lateral sclerosis. *Neurology* **70**, 1004-9 (2008).
3. Dorst, J. *et al.* Patients with elevated triglyceride and cholesterol serum levels have a prolonged survival in amyotrophic lateral sclerosis. *J Neurol* **258**, 613-7 (2011).
4. Ikeda, K., Hirayama, T., Takazawa, T., Kawabe, K. & Iwasaki, Y. Relationships between disease progression and serum levels of lipid, urate, creatinine and ferritin in Japanese patients with amyotrophic lateral sclerosis: a cross-sectional study. *Intern Med* **51**, 1501-8 (2012).
5. Yang, J.W. *et al.* Hypolipidemia in patients with amyotrophic lateral sclerosis: a possible gender difference? *J Clin Neurol* **9**, 125-9 (2013).
6. Sutedja, N.A. *et al.* Beneficial vascular risk profile is associated with amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry* **82**, 638-42 (2011).
7. Chio, A. *et al.* Lower serum lipid levels are related to respiratory impairment in patients with ALS. *Neurology* **73**, 1681-5 (2009).
8. Mariosa, D. *et al.* Blood biomarkers of carbohydrate, lipid and apolipoprotein metabolisms and risk of amyotrophic lateral sclerosis: a more than 20 year follow-up of the Swedish AMORIS cohort. *Annals of Neurology* (2017).
9. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
10. Global Lipids Genetics, C. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* **45**, 1274-83 (2013).
11. Buscot, M.J. *et al.* The Combined Effect of Common Genetic Risk Variants on Circulating Lipoproteins Is Evident in Childhood: A Longitudinal Analysis of the Cardiovascular Risk in Young Finns Study. *PLoS One* **11**, e0146081 (2016).
12. van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat Genet* **48**, 1043-8 (2016).
13. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**, 333-8 (2011).
14. Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**, 1121-30 (2015).
15. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
16. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
17. Carstea, E.D. *et al.* Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science* **277**, 228-31 (1997).
18. McLaughlin, R.L. *et al.* Genetic correlation between amyotrophic lateral sclerosis and schizophrenia. *Nat Commun* **8**, 14774 (2017).
19. Al-Chalabi, A., van den Berg, L.H. & Veldink, J. Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nat Rev Neurol* **13**, 96-104 (2017).
20. Euesden, J., Lewis, C.M. & O'Reilly, P.F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466-8 (2015).

TABLE

Table. Best-fitted Results from the Bi-directional Polygenic Risk Score Analyses

Base -- Target	Best P_T	OR (SE)	n_{SNP}	r^2(%)	P-value
Blood lipids -- ALS					
LDL-C_EA -- ALS_EA	$5.00x10^{-5}$	1.17 (0.03)	233	0.08	$8.82x10^{-8}$
LDL-C_Trans -- ALS_EA	0.0003	1.19 (0.02)	670	0.15	$1.73x10^{-13}$
TC_EA -- ALS_EA	$5.00x10^{-5}$	1.15 (0.03)	270	0.07	$8.17x10^{-7}$
TC_Trans -- ALS_EA	0.0015	1.11 (0.02)	1476	0.08	$1.37x10^{-7}$
TG_EA -- ALS_EA	$5.00x10^{-7}$	1.07 (0.04)	93	0.00	0.0959
TG_Trans -- ALS_EA	0.0905	0.98 (0.01)	11931	0.00	0.0942
HDL-C_EA -- ALS_EA	0.1604	1.00 (0.00)	32029	0.00	0.2241
HDL-C_Trans -- ALS_EA	0.0020	1.06 (0.02)	1680	0.02	0.0101
ALS -- Blood lipids					
ALS_EA -- LDL-C_EA	0.0009	1.01 (0.00)	617	0.00	0.0859
ALS_EA -- LDL-C_Trans	0.4994	1.00 (0.00)	53875	0.00	0.0007
ALS_EA -- TC_EA	0.0009	1.01 (0.00)	617	0.00	0.0638
ALS_EA -- TC_Trans	0.0057	1.00 (0.00)	2662	0.00	0.0144
ALS_EA -- TG_EA	0.0001	0.99 (0.01)	98	0.00	0.0677
ALS_EA -- TG_Trans	0.0001	0.98 (0.01)	86	0.00	0.0247
ALS_EA -- HDL-C_EA	0.0001	1.01 (0.01)	98	0.00	0.0772
ALS_EA -- HDL-C_Trans	0.0180	1.00 (0.00)	6660	0.00	0.0383
Bi-directional analysis for LDL-C and CAD					
LDL-C_EA --CAD_EA	$5.00x10^{-5}$	1.56 (0.03)	232	0.33	$5.51x10^{-64}$
LDL-C_Trans -- CAD_EA	$5.00x10^{-5}$	1.56 (0.03)	358	0.32	$3.06x10^{-63}$
LDL-C_EA -- CAD_Trans	$5.00x10^{-7}$	1.45 (0.02)	108	0.25	$9.24x10^{-104}$
LDL-C_Trans -- CAD_Trans	$1.00x10^{-4}$	1.55 (0.01)	497	0.60	$1.43x10^{-244}$
CAD_EA -- LDL-C_EA	0.0658	1.01 (0.00)	20136	0.17	$9.46x10^{-37}$
CAD_Trans -- LDL-C_Trans	0.0012	1.13 (0.00)	1148	0.54	$1.02x10^{-225}$

Standardized polygenic risk score (PRS) of the base phenotype was used to predict the target phenotype. P_T=P-value threshold of the association between single nucleotide polymorphisms (SNPs) and the base phenotype, the best P_T was defined as the threshold corresponding to the largest explained variance. OR=odds ratio per standard deviation; SE=standard error; n_{SNP}=number of independent SNPs included in the best P_T quantile; r^2=Nagelkerke r^2, the proportion of target variation explained by PRS of SNPs in the best P_T quantile. Trans=trans-ethnic data; EA=European-ancestry data; LDL-C=low-density lipoprotein cholesterol; ALS= amyotrophic lateral sclerosis; TC=total cholesterol; TG=triglycerides; HDL-C=high-density lipoprotein cholesterol; CAD= coronary artery disease.

FIGURES WITH FIGURE LEGENDS

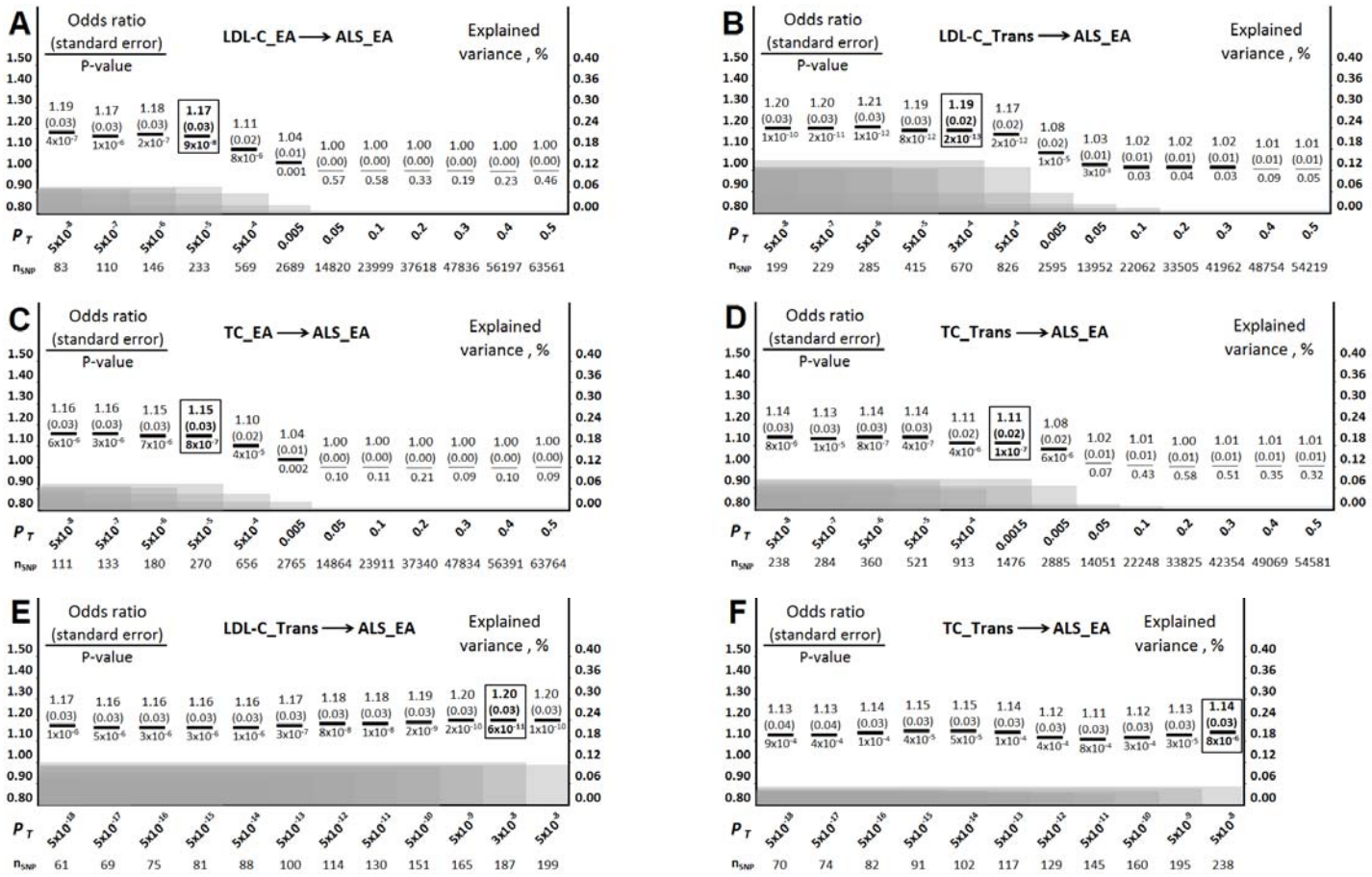

Figure 1. Polygenic Risk Score Analyses of LDL-C and TC on ALS. A, LDL-C_EA. **B,** LDL-C_Trans. **C,** TC_EA. **D,** TC_Trans. **E,** Genome-wide significant SNPs in LDL-C_Trans. **F,** Genome-wide significant SNPs in TC_Trans**.** EA=European-ancestry data; Trans=trans-ethnic data; LDL-C=low-density lipoprotein cholesterol; TC=total cholesterol; ALS=amyotrophic lateral sclerosis; P_T=P-value threshold of the association between single nucleotide polymorphisms (SNPs) and base; n_{SNP}=number of independent SNPs included in the quantile. Bold black bars stand for statistically significant predictions, whereas gray bars stand for predictions that are not statistically significant. The rectangles mean the best-fitted results.

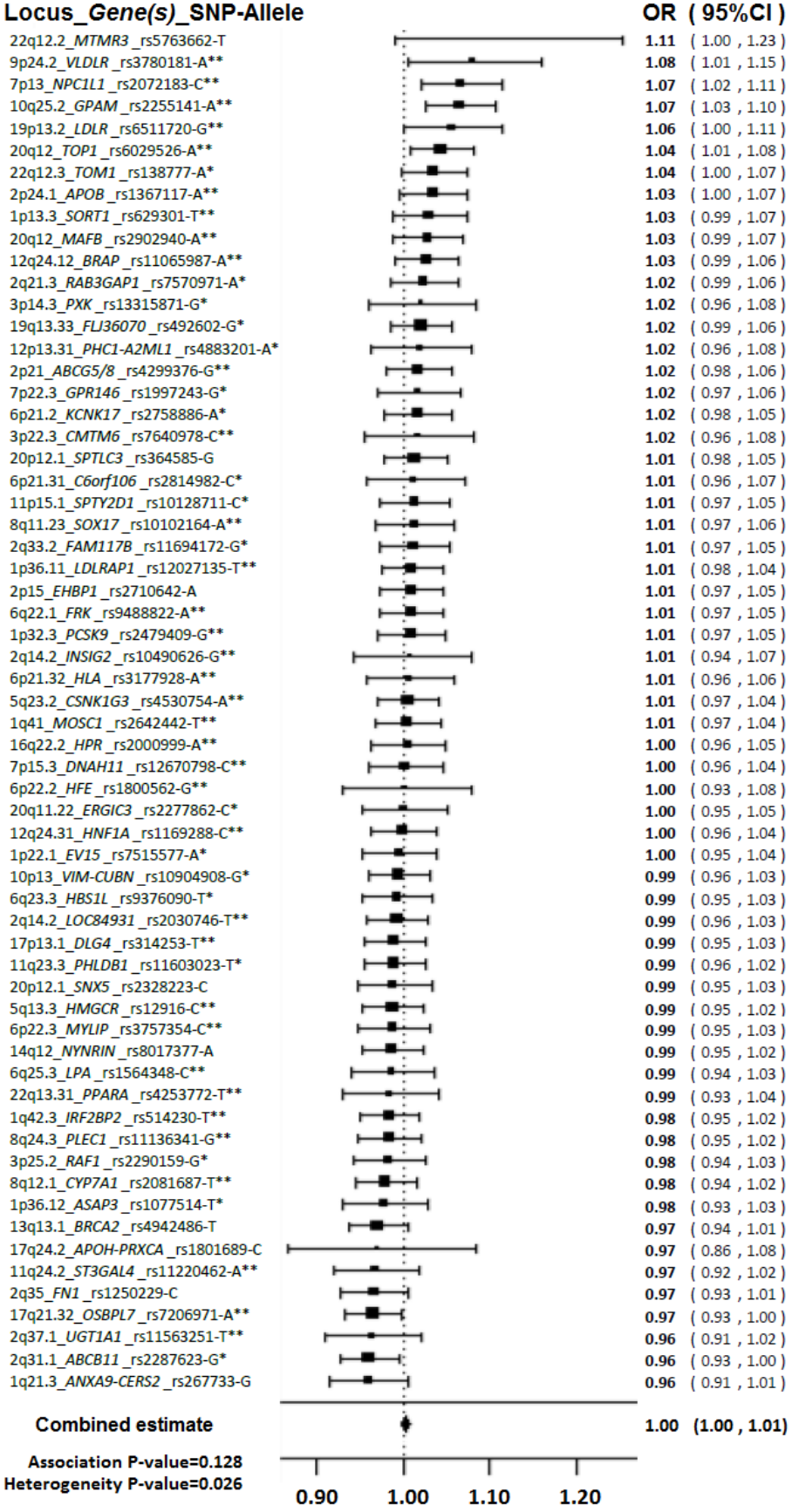

Figure 2. Mendelian Randomization Analyses between LDL-C, TC and ALS. Among the genome-wide significant ($P<5x10^{-8}$) single nucleotide polymorphisms (SNPs)/loci associated with low-density lipoprotein and total cholesterol (LDL-C and TC), 62 independent SNPs that are specifically associated with LDL-C or TC, but not other lipids, were selected as instrumental variables. The individual and combined effects of the LDL-C or TC up-regulating alleles on ALS were plotted. OR=odds ratio per allele; 95%CI=95% confidence interval. SNPs without star are only associated with LDL-C, SNPs with one star (*) are only associated with TC, SNPs with two stars (**) are associated with both LDL-C and TC.