The Pharmacogenomics Journal (2017) 17, 222–229
www.nature.com/tpj

OPEN

ORIGINAL ARTICLE

Statin-induced expression change of *INSIG1* in lymphoblastoid cell lines correlates with plasma triglyceride statin response in a sex-specific manner

E Theusch[1], K Kim[1], K Stevens[1], JD Smith[2], Y-DI Chen[3], JI Rotter[3], DA Nickerson[2] and MW Medina[1]

Statins are widely prescribed to lower plasma low-density lipoprotein (LDL) cholesterol levels. They also modestly reduce plasma triglyceride (TG), an independent cardiovascular disease risk factor, in most people. The mechanism and inter-individual variability of TG statin response is poorly understood. We measured statin-induced gene expression changes in lymphoblastoid cell lines derived from 150 participants of a simvastatin clinical trial and identified 23 genes (false discovery rate, FDR = 15%) with expression changes correlated with plasma TG response. The correlation of insulin-induced gene 1 (*INSIG1*) expression changes with TG response (rho = 0.32, $q = 0.11$) was driven by men (interaction $P = 0.0055$). rs73161338 was associated with *INSIG1* expression changes ($P = 5.4 \times 10^{-5}$) and TG response in two statin clinical trials ($P = 0.0048$), predominantly in men. A combined model including *INSIG1* expression level and splicing changes accounted for 29.5% of plasma TG statin response variance in men ($P = 5.6 \times 10^{-6}$). Our results suggest that *INSIG1* variation may contribute to statin-induced changes in plasma TG in a sex-specific manner.

The Pharmacogenomics Journal (2017) **17,** 222–229; doi:10.1038/tpj.2016.12; published online 1 March 2016

INTRODUCTION

Statins are commonly used to reduce the risk of cardiovascular disease (CVD), the leading cause of death in the United States. They act by inhibiting the activity of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), the rate-limiting cholesterol biosynthesis enzyme. The resulting low intracellular sterol levels release the sterol regulatory element binding transcription factor 2 (SREBF2) from the endoplasmic reticulum, allowing it to be cleaved and activated in the Golgi, travel to the nucleus and bind DNA. This stimulates the upregulation of genes in the cholesterol synthesis and uptake pathways, most notably the low-density lipoprotein (LDL) receptor, which mediates uptake of LDL from the plasma.[1] Statins are most effective at lowering LDL-cholesterol (LDL-C), but they also lower plasma triglyceride (TG) in many people.[2] Elevated plasma TG levels contribute to cause CVD independent of LDL-C or high-density lipoprotein cholesterol levels,[3,4] and are thought to be responsible for some of the residual risk in statin-treated individuals after adjustment for LDL-C and other covariates.[5,6] Although the general mechanism by which statin treatment lowers LDL-C is well understood, little is known about factors underlying statin-induced changes in TG.

Though statins have well-documented clinical efficacy, there is a large degree of inter-individual variability in lipid lowering by statins,[7] and many individuals continue to have CVD events despite statin treatment.[8] There is some debate as to whether statins are as effective at primary prevention of CVD events in women as they are in men,[9–11] so further investigation of potential sex differences in statin response is warranted. A number of pharmacogenetic studies have sought to identify single-nucleotide polymorphisms (SNPs) that influence LDL-C statin response using candidate-gene and genome-wide association approaches, but the findings to date account for less than 10% of the variation in statin-induced LDL-C lowering variability,[12–20] with variants near the apolipoprotein E (*APOE*) and lipoprotein, Lp(a) (*LPA*) genes showing some of the strongest and most replicated signals.[17,18,20] Genetic association studies of TG statin response have been even less fruitful, with two published genome-wide association studies revealing no genome-wide significant hits[15,16] and only a modest number of significant associations reported with genetic variation in candidate genes, such as lipoprotein lipase (*LPL*),[21] cholesteryl ester transfer protein, plasma (*CETP*), ATP-binding cassette, sub-family B (MDR/TAP), member 1 (*ABCB1*)[22] and *APOE*. Interestingly, the *APOE* locus was only associated with TG statin response in men,[23] indicating that there may be sex-specific differences in the genetic basis of TG statin response.

As statins stimulate a large transcriptional response mediated by SREBF2 and other factors, we sought to investigate inter-individual differences in statin-induced gene expression changes to identify genes that may have a role in TG statin response. Toward this objective, we performed transcriptomic analysis on a panel of lymphoblastoid cell lines (LCLs) derived from participants of the Cholesterol and Pharmacogenetics (CAP) simvastatin clinical trial.[7,24] Though LCLs have been transformed by Epstein-Barr virus (EBV) and thus may not perfectly reflect the genetic background of the human donors, we and others have shown that they are a useful model for the study of cholesterol metabolism and statin response *in vitro*.[24–27] Using RNA-seq data collected from 100 Caucasian and 50 African-American control- and statin-exposed CAP LCLs, we identified genes with *in vitro*

[1]Children's Hospital Oakland Research Institute, Oakland, CA, USA; [2]Department of Genome Sciences, University of Washington, Seattle, WA, USA and [3]Departments of Pediatrics and Medicine, Los Angeles Biomedical Research Institute at Harbor–UCLA, Torrance, CA, USA. Correspondence: Dr MW Medina, Children's Hospital Oakland Research Institute, 5700 Martin Luther King Jr. Way, Oakland, CA 94609, USA.
E-mail: mwmedina@chroi.org
Received 16 September 2015; revised 21 January 2016; accepted 27 January 2016; published online 1 March 2016

Table 1. Characteristics of 150 CAP participants used in RNA-seq analysis

	AA men	*AA women*	*Cau men*	*Cau women*	*All*
N	26	24	60	40	150
Ethnicity	Afr. Am	Afr. Am	Caucasian	Caucasian	66.7% Cau
Gender	Male	Female	Male	Female	57.3% male
% smokers	15.4%	12.5%	10.0%	7.5%	10.7%
Age (years)	50.6 ± 12.6	52.6 ± 13.2	53 ± 12.2	55.9 ± 12.5	53.3 ± 12.5
BMI[a]	29.4 ± 5.5	30.9 ± 5.8	27.7 ± 4.2	27.5 ± 6.5	28.5 ± 5.5
Baseline TG	132 ± 65.2	99 ± 42.6	131.5 ± 65.8	122.1 ± 64.1	123.9 ± 62.6
Statin TG	110.9 ± 55.9	78.5 ± 33.4	104.6 ± 70.8	92.8 ± 49.3	98.4 ± 58.6
% Change TG	− 9.3 ± 41.4%	− 15.7 ± 26.5%	− 20.2 ± 25.6%	− 21.8 ± 18.4%	− 18.0 ± 27.6%
Baseline TC	206.4 ± 40.5	202.0 ± 34.6	207.0 ± 33.4	221.1 ± 38.0	209.9 ± 36.4
Statin TC	149.7 ± 35.5	152.9 ± 27.6	146.7 ± 27.8	156.2 ± 33.4	150.8 ± 30.7
% Change TC	− 26.9 ± 13.3%	− 23.9 ± 10.4%	− 28.9 ± 9.9%	− 28.9 ± 11.5%	− 27.8 ± 11.1%
Baseline LDL-C	134.4 ± 36.9	121 ± 35.7	134.3 ± 31.6	135.5 ± 33.9	132.5 ± 33.9
Statin LDL-C	80.6 ± 28.3	70.9 ± 23.3	77.6 ± 23.9	74.7 ± 27.9	76.3 ± 25.6
% Change LDL-C	− 39.4 ± 15.2%	− 40.7 ± 13.1%	− 42.2 ± 13.1%	− 44.7 ± 14.9%	− 42.2 ± 13.9%
Baseline HDL-C[b]	45.6 ± 10.9	61.4 ± 20.5	46.6 ± 12.1	61.4 ± 20.6	52.7 ± 17.6
Statin HDL-C[b]	47.1 ± 10.6	66.5 ± 22.7	48.4 ± 13.0	63.2 ± 20.2	55.0 ± 18.4
% Change HDL-C	4.3 ± 11.0%	8.7 ± 10.8%	4.5 ± 11.4%	3.6 ± 9.4%	4.9 ± 10.8%

Abbreviations: AA, African American; BMI, body mass index; CAP, Cholesterol and Pharmacogenetics; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride. Values are mean ± s.d. Baseline and statin lipid measurements are in mg dl^{-1}. [a]BMI was significantly different between the 100 Caucasians and 50 African Americans ($P = 0.007$, Wilcoxon rank-sum test). [b]Baseline and on-statin HDL-cholesterol levels were significantly different between the 86 men and 64 women ($P = 3 \times 10^{-8}$ for both, Wilcoxon rank-sum test). None of the other phenotypes or covariates listed here were significantly different ($P < 0.05$) between races and sexes in this $N = 150$ CAP subset.

statin-induced expression changes that were correlated with plasma TG response in the cell line donors. We then investigated whether genetic variants near insulin-induced gene 1 (*INSIG1*), the most promising biological candidate from the list of correlated genes, were associated with *in vitro* gene expression changes and plasma TG response, suggesting a role for *INSIG1* in modulating TG statin response.

SUBJECTS AND METHODS

Study population

Of the 609 self-reported Caucasian (white) and 335 self-reported African-American (black) participants of the CAP 40-mg simvastatin/day 6-week clinical trial (ClinicalTrials.gov ID: NCT00451828),[7] LCLs from 100 Caucasian and 50 African-American CAP participants were used in our RNA-seq analyses after quality control filtering. Informed written consent was obtained from the IRBs (Institutional Review Boards) of the sites of recruitment, the University of California Los Angeles and the University of California San Francisco. In addition, all research involving human participants was approved by the Children's Hospital Oakland Research Institute IRB. Clinical covariates and before and on-treatment lipid levels of these 150 participants are summarized in Table 1. TG was measured in fasting plasma samples at two pre-treatment and two on-treatment time points.[7] TG statin response (log(mean on-treatment TG) − log(mean pre-treatment TG)) of this subset was representative of the entire CAP population (Supplementary Figure S1). The pravastatin inflammation/CRP evaluation (PRINCE) 24 week clinical trial of 40 mg per day pravastatin has been previously described.[28] For our genetic analyses, we used 1311 white individuals with genome-wide genotype data as well as before and on statin treatment plasma TG measurements, including 503 individuals from the primary prevention cohort that were allocated statins and 808 individuals from the secondary prevention cohort.

RNA-seq

Using polyA-selected RNA from LCLs exposed to 2 μM simvastatin or control buffer for 24 h, indexed, strand-specific, paired-end Illumina sequencing libraries were prepared using a dUTP second-strand marking method[29] and sequenced to an average depth of 67 million 100 or 101 bp reads (Supplementary Figures S2 and S3; Supplementary Methods). Sequences were aligned to the human (hg19) and EBV (NC_007605) genomes with the Ensembl v67 reference transcriptome allowing four mismatches per aligned read using Tophatv2.0.4.[30] Duplicate fragments were removed from downstream analyses. Identity mismatches, sex mismatches, mixed samples, samples inferred to be from Latinos and 5′ to 3′ bias outliers were excluded. The resulting RNA-seq data are deposited in the database of genotypes and phenotypes (dbGaP) within phs000481.v2.p1. LCL gene fragment counts above our expression threshold were adjusted for library size and variance stabilized in DESeq2.[31] Gene expression changes were calculated and probabilistic estimation of expression residuals (PEER)[32] normalized (Supplementary Figure S4) for correlation analyses.

Statin-responsive genes were identified by first conducting paired two-tailed *t*-tests allowing unequal variance between the variance stabilized control and statin expression levels of each gene and then by calculating the false discovery rate (FDR) adjusted *P*-values for all 14 004 genes using the *P*.adjust function in R. Genes with *q* values (FDR-adjusted *P*-values) under 0.0001 were considered as statin responsive. Spearman's rank order correlations between PEER-adjusted variance stabilized LCL gene expression changes and plasma TG statin response (delta log TG) were calculated using the rcorr function of the Hmisc package in R, and their *P*-values were FDR-adjusted as described above. An FDR threshold of 15% was used to identify candidate genes with statin-induced expression changes significantly correlated with TG response. The sex interaction analyses and other regression analyses were conducted in JMP 11.0.0 statistical software (SAS, Cary, NC, USA). Gene annotation enrichment analyses were conducted in DAVIDv6.7 (ref. 33) using the 13 931 human genes above the expression level threshold in LCLs as the background.

Imputation and genetic association analyses

Participants of the CAP and PRINCE[28] statin clinical trials were genotyped as previously described.[16,34] Using the individuals and SNP genotypes that survived our quality control filters, we imputed to over 35 million SNPs with reference haplotypes from all of the populations in the 1000 Genomes Project Phase I v3 Shapeit2 panel[35] using MaCH[36] and minimac.[37]

We conducted association analyses between genetic variants less than 200 kb[38] upstream or downstream of the *INSIG1* transcription start site and PEER-adjusted variance stabilized changes in gene expression ($N = 99$ Caucasian LCLs) using an additive model in mach2qtl.[36] Owing to our modest sample size, we considered only SNP or indel variants with minor allele frequencies greater than 10% and omitted imputed variants with a MaCH imputation $r^2 < 0.5$.[36] To assess the significance of our results, we

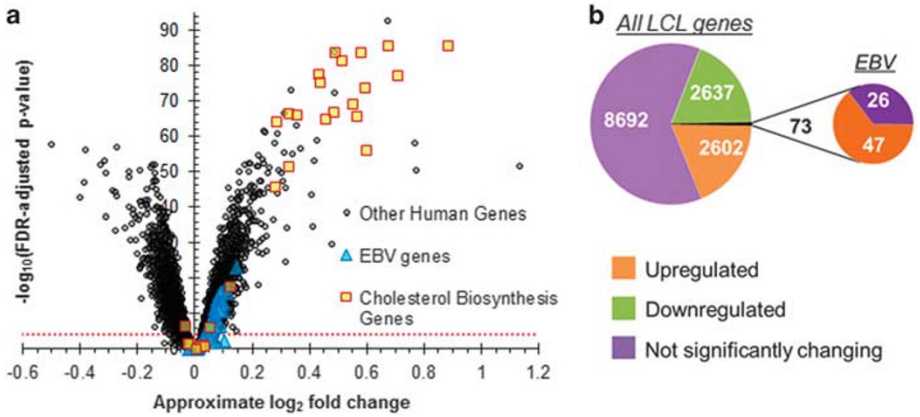

Figure 1. Statin-induced changes in human and Epstein-Barr virus (EBV) gene expression in 150 Cholesterol and Pharmacogenetics (CAP) lymphoblastoid cell lines (LCLs). Gene expression levels were quantified using RNA-seq data from control- and statin-treated LCLs derived from 100 Caucasian and 50 African-American participants of the CAP simvastatin clinical trial. (**a**) Statin-responsive genes were identified using a paired two-tailed *t*-test comparing control- and statin-treated gene expression values. The approximate $\log_2$ fold change was calculated by subtracting the library-size adjusted and variance stabilized estimates of the control from the corresponding statin libraries and averaging across all LCLs. The horizontal red dotted line indicates false discovery rate (FDR)-adjusted $P < 0.0001$. Genes of EBV origin are depicted as blue triangles, and human genes that are annotated with the cholesterol biosynthesis Gene Ontology biological process are shown as yellow squares. (**b**) The numbers of statin-responsive LCL genes in the 150 CAP LCLs (FDR-adjusted $P < 0.0001$) out of a total of 13 931 human and 73 EBV genes were categorized by directionality.

permuted the *INSIG1* gene expression change phenotypes among the 99 individuals 1000 times and used these 1000 randomly generated phenotypes to conduct simulations in mach2qtl, retaining the most significant association from each simulation that survived our minor allele frequencies and r^2 filtering steps and were located in the 200-kb window around *INSIG1*. The empirical *P*-value of our result was calculated by dividing the number of simulations that had a *P*-value at least as significant by the total number of simulations considered (1000). We also conducted association analyses between our candidate SNP and delta log TG in CAP and PRINCE using an additive model and similar methodology as described above.

Splicing analyses

Percent-spliced-in (PSI) values for *INSIG1* alternative splicing events were calculated using JuncBASEv0.6.[39] Events that had less than 25 total reads, had median PSI values less than 1% or greater than 99% or were significantly correlated (with $r^2 > 0.4$) with another event were removed from the data set, leaving 6 *INSIG1* alternative splicing events. Delta PSI values (statin PSI − control PSI) for each event in each cell line were calculated, and the splicing changes were each tested for correlation with delta log TG using Spearman's correlation in JMP. The delta PSI values of two splicing events that exhibited correlations with TG response ($P < 0.05$) were incorporated into a multiple linear regression model that also included statin-induced changes in *INSIG1* expression levels to determine the relationship of both *INSIG1* expression and splicing changes with TG response. This multiple linear regression was conducted on the entire sample set as well as the subset of males separately in JMP.

RESULTS

Statin-responsive LCL genes

Using RNA-seq data from 100 Caucasian and 50 African-American control- and statin-exposed LCLs derived from CAP participants, we found that expression levels of 37.6% of human and 64.4% of EBV genes tested were significantly altered by statin treatment ($N = 150$, two-sided paired *t*-test FDR-adjusted $P < 0.0001$; Figure 1; Supplementary Table S1). The most significant statin-responsive genes were upregulated and were enriched for genes in the cholesterol biosynthesis pathway (Figure 1a). The expression levels of *HMGCR*, an SREBF2 target gene, were increased in all 150 LCLs with statin treatment, confirming the effectiveness of the *in vitro* statin exposure (Supplementary Figure S5). Consistent with our previous findings based on expression array data from 480 CAP LCLs,[26] significantly changing genes were also enriched for other gene ontology biological processes; for example, upregulated genes were enriched for sterol biosynthetic process and downregulated genes were enriched for RNA processing and cell-cycle phase (Supplementary Tables S2 and S3). Though approximately half of all of the statin-responsive human genes went up and the other half went down ($q < 0.0001$), all of the statin-responsive EBV genes were upregulated (Figure 1; Supplementary Table S1).

Gene expression and TG statin response correlations

Plasma TG levels decreased by 15 ± 27% with statin treatment in CAP on average ($N = 944$; minimum −72% change; maximum +206% change), and 22% of individuals actually experienced an increase in TG with statin treatment, indicating a large degree of inter-individual variability in TG statin response (Supplementary Figure S1). Plasma TG changes were not significantly correlated with age, sex, BMI, race or smoking status in the CAP population ($N = 944$, $P > 0.05$ for all), but there was a significantly greater variance in delta log TG in men than in women (−0.21 ± 0.25 in 459 women and −0.21 ± 0.32 in 485 men, Levene's test $P = 7.5 \times 10^{-5}$). To investigate the inter-individual variation in TG response further, we asked which genes had statin-induced changes (statin-control expression levels) in the LCLs that were most correlated with the plasma TG statin responses of the corresponding donors. After adjusting gene expression changes for potential confounders using PEER, 23 genes had statin-induced changes that were significantly correlated with TG response ($N = 150$, FDR = 15%, Table 2), and this list of genes was significantly enriched for genes involved in lipid biosynthesis, sterol metabolism and related Gene Ontology biological processes (Supplementary Table S4). Among the seven genes annotated as being involved in lipid biosynthesis, only fatty acid desaturase 3 (*FADS3*) was negatively correlated with TG response, and this gene was not annotated as being involved in sterol metabolism, unlike the other six lipid biosynthesis genes. Notably, *INSIG1* is involved in sterol metabolism but is an upstream regulator of lipid biosynthesis,[40,41] not a lipid biosynthesis enzyme, unlike the other six sterol metabolism genes in the list (Table 2).

Table 2. Genes with statin-induced LCL expression changes significantly correlated with plasma triglyceride statin response (FDR = 15%)

Gene name	*Ensembl gene ID*	*Location*	*rho*	p	q	*Sterol metab.*	*Lipid biosyn.*
HMGCS1	ENSG00000112972	5p14-p13	0.383	1.30E − 06	0.014	Y	Y
EBP	ENSG00000147155	Xp11.23	0.377	2.05E − 06	0.014	Y	Y
ITGB2-AS1	ENSG00000227039	21q22.3	− 0.347	1.36E − 05	0.063		
IDI1	ENSG00000067064	10p15.3	0.331	3.48E − 05	0.098	Y	Y
UHRF2	ENSG00000147854	9p24.1	− 0.331	3.48E − 05	0.098		
TRIM2	ENSG00000109654	4q31.3	− 0.322	5.83E − 05	0.109		
NSDHL	ENSG00000147383	Xq28	0.322	5.92E − 05	0.109	Y	Y
SC5D	ENSG00000109929	11q23.3	0.320	6.72E − 05	0.109	Y	Y
INSIG1	ENSG00000186480	7q36	0.316	8.00E − 05	0.109	Y	
ARHGAP11B	ENSG00000187951	15q13.2	− 0.316	8.23E − 05	0.109		
FADS3	ENSG00000221968	11q12	− 0.314	9.31E − 05	0.109		Y
NIPSNAP1	ENSG00000184117	22q12	0.314	9.37E − 05	0.109		
FAHD1	ENSG00000180185	16p13.3	0.310	0.000115	0.120		
MARCKSL1	ENSG00000175130	1p35.1	0.306	0.000138	0.120		
YTHDF2	ENSG00000198492	1p35	0.306	0.000140	0.120		
COX10-AS1	ENSG00000223385	17p12	− 0.305	0.000151	0.120		
CYB5B	ENSG00000103018	16q22.1	0.305	0.000151	0.120		
DHCR7	ENSG00000172893	11q13.4	0.303	0.000160	0.120	Y	Y
R3HCC1L	ENSG00000166024	10q24.2	− 0.303	0.000163	0.120		
TRIM38	ENSG00000112343	6p21.3	− 0.301	0.000183	0.128		
ISOC1	ENSG00000066583	5q22.1	0.300	0.000195	0.130		
IDH1	ENSG00000138413	2q34	0.298	0.000209	0.133		
KIF18B	ENSG00000186185	17q21.31	−0.296	0.000240	0.146		

Abbreviations: FDR, false discovery rate; LCL, lymphoblastoid cell line. Rho is the Spearman's rank correlation coefficient, *q* is the FDR-adjusted *P*-value and the sterol metabolism and lipid biosynthesis columns indicate the genes that are annotated with those Gene Ontology biological processes.

INSIG1 associations with TG response

As noted above, one of the top 10 genes with expression changes most correlated with TG response was *INSIG1*, an attractive candidate gene based on its central roles in sterol regulation through its inhibition of SREBF activation[40] and promotion of HMGCR degradation.[41] *INSIG1* expression changes and plasma TG changes were positively correlated, with LCLs derived from individuals who experienced the greatest reductions in plasma TG exhibiting some of the smallest increases in *INSIG1* mRNA expression levels with statin treatment (Spearman's $\rho = 0.32$, $q = 0.11$, $N = 150$, Figure 2a). We noticed that this correlation was stronger in men than in women, with a significant sex interaction term ($P = 0.0055$; Figure 2b). After imputation of the genotypes of CAP Caucasians using a 1000 genomes project reference panel, we looked for common (minor allele frequencies $\geqslant 10\%$) genetic variants within 200 kb of *INSIG1* that were correlated with statin-induced changes in *INSIG1* and found an SNP, rs73161338, that was a differential expression quantitative trait locus less than 100 kb from the transcription start site of *INSIG1* (Figure 2c; Supplementary Figures S6A and S7, unadjusted $P = 5.4 \times 10^{-5}$, $N = 99$). We conducted 1000 simulations using permuted phenotypes and markers within 200 kb of *INSIG1* to further evaluate the significance of this *P*-value, and only 5 of these simulations exhibited more significant *P*-values than that which was observed with the real data, yielding an empirical *P*-value of 0.005. The *INSIG1* differential expression quantitative trait locus SNP (rs73161338) was associated with TG response in Caucasians from the combined CAP and PRINCE statin trial populations ($P = 0.0048$, $N = 1887$), and the association had a similar magnitude and direction of effect in CAP and PRINCE separately (Supplementary Figure S6B). Interestingly, the association was driven by men ($P = 0.0016$ in 1320 men alone) and was essentially absent in women ($P > 0.9$ in 567 women, Figure 2d). The rs73161338 minor (A) allele was associated with smaller statin-induced increases in *INSIG1* in LCLs and with greater reductions in plasma TG *in vivo*, consistent with the positive correlation we observed between change in cellular *INSIG1* and change in plasma TG (Figure 2).

We also calculated PSI values for six common *INSIG1* alternative splicing events in the control- and statin-treated LCLs (Figure 3a; Supplementary Table S5) and found that statin treatment modestly reduced the prevalence of Event #2 from 27.3 to 26.4% on average ($N = 150$, two-tailed Wilcoxon signed rank $P = 0.002$). When statin-induced changes in the six splicing events were tested for correlation with TG response, statin-induced changes in two different *INSIG1* splicing events had modest correlations with TG response (Figures 3b and c). The first was a known event that involved skipping two protein-coding exons that are present in the major isoform (Event #4, Spearman's $\rho = 0.19$, $P = 0.02$, $N = 150$), and the second was intron retention of at least part of the last intron that is normally spliced out of the major isoform (Event #6, Spearman's $\rho = -0.21$, $P = 0.01$, $N = 142$). For both of these events, about half of the cell lines exhibited increases in splicing with statin treatment and the other half exhibited decreases, so the mean change in splicing was close to 0. Similar to the SNP and expression level associations with TG response described above, these correlations were stronger in males than in females (Supplementary Figure S8).

Finally, to assess the combined impact of *INSIG1* gene expression and splicing changes on TG statin response, we incorporated all three measures (*INSIG1* expression and splicing events #4 and #6) into a multiple linear regression model. The overall model explained 14.2% of the variance (r^2) in TG response ($P = 9.2 \times 10^{-5}$, $N = 142$, Supplementary Table S6). Consistent with the gender differences in the individual correlations, statin-induced changes in *INSIG1* expression levels and splicing events #4 and #6 accounted for 29.5% of the variance in TG response in men alone ($P = 5.6 \times 10^{-6}$, $N = 81$, Supplementary Table S6).

DISCUSSION

Though TG lowering is not the primary goal of statin treatment, recent studies demonstrating that high TG levels can cause CVD[3,4] and that TG levels may be responsible for some of the residual CVD risk in statin-treated individuals[5,6] highlight the value of understanding how statin treatment changes plasma TG levels

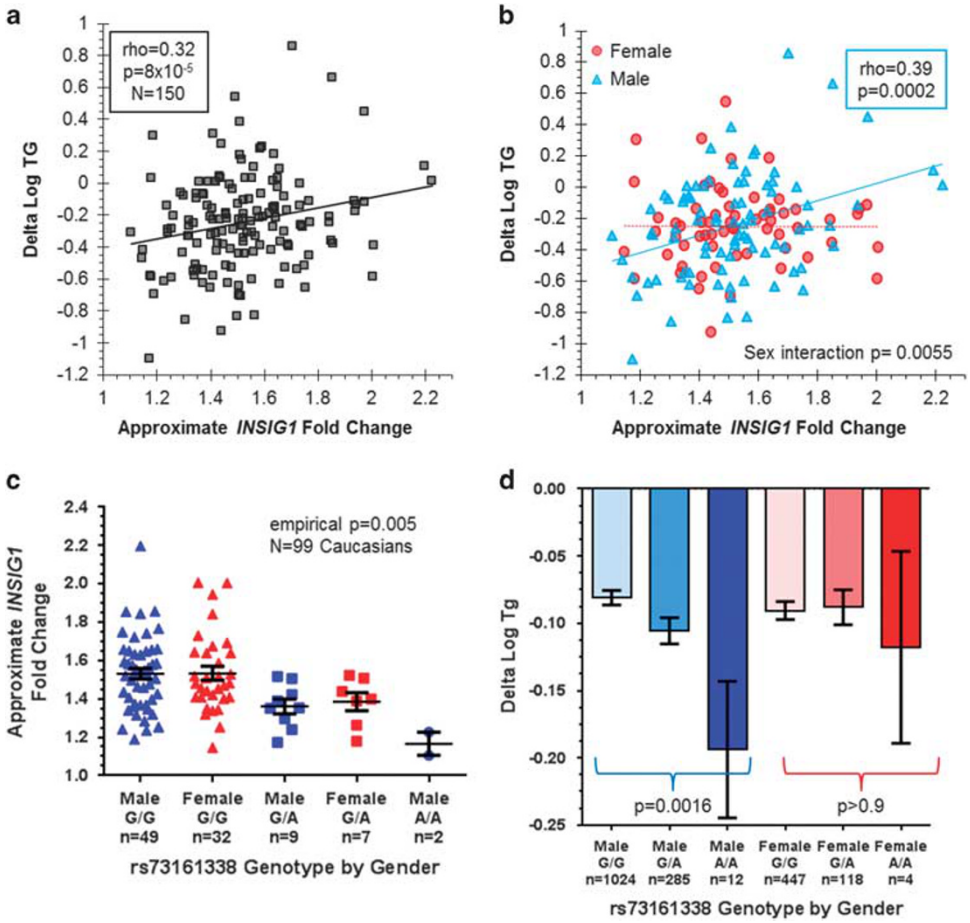

Figure 2. Correlations of insulin-induced gene 1 (*INSIG1*) statin response, plasma triglyceride (TG) statin response and rs73161338 genotype split by sex. (**a**) Correlation of *INSIG1* lymphoblastoid cell line (LCL) gene expression changes with plasma TG response in $N = 150$ Cholesterol and Pharmacogenetics (CAP) participants. The approximate *INSIG1* fold change was calculated by taking 2^(variance stabilized gene expression deltas). Though the trendline is from linear regression and the x axis depicts approximate fold change, the correlation was calculated by taking the Spearman correlation of the *INSIG1* variance stabilized gene expression deltas (adjusted for three probabilistic estimation of expression residuals (PEER) hidden factors) with plasma delta log TG. (**b**) Correlations of *INSIG1* LCL gene expression changes with plasma TG response in $N = 86$ male and $N = 64$ female CAP participants. Correlations in each sex subset were calculated as in (**a**). (**c**) *INSIG1* gene expression changes separated by imputed rs73161338 genotype and sex. The unadjusted *P*-value for the association of delta *INSIG1* versus allelic dosage using an additive allelic model for both sexes combined was 5.4×10^{-5} (empirical $P = 0.005$), with no sex-specific differences identified. (**d**) Plasma TG statin response separated by imputed rs73161338 genotype and sex. *P*-values were calculated using an additive allelic model associating allelic dosage with TG response in $N = 307$ CAP+$N = 1013$ pravastatin inflammation/CRP evaluation (PRINCE) Caucasian men and $N = 269$ CAP+$N = 298$ PRINCE Caucasian women. Means ± s.e. are displayed.

and the importance of identifying genetic factors that contribute to inter-individual variation in TG statin response. To date a few small human kinetic studies have found that statin treatment lowers plasma TG levels by increasing the clearance (fractional catabolic rate) of very low-density lipoprotein (VLDL) particles[42–48] and increasing both preheparin plasma lipase activity and postheparin LPL activity,[43] while other studies have indicated that statins reduce VLDL-TG secretion and VLDL particle size.[49–51] It is possible that the mechanism by which statins reduce plasma TG levels could depend on an individual's baseline fractional catabolic rates and production rates.

The roles of INSIG1 in regulation of cholesterol homeostasis have been studied for over a decade.[40,41] Under conditions of sterol depletion, such as statin treatment, INSIG1 no longer retains SREBF chaperone in the endoplasmic reticulum, releasing it to transport the SREBFs to the Golgi for cleavage and activation. Once activated, the SREBFs can enter the nucleus and promote transcription of their target genes, including *INSIG1*. Though *INSIG1* transcription increases with sterol depletion, INSIG1 proteins are rapidly turned over when they are not bound to SREBF chaperone, thus reducing overall INSIG1 protein levels.[52] The mouse hepatic *Insig1* knockout shows trends toward lower plasma TG levels and increased hepatic TG accumulation. These differences become significant in *Liver-Insig1*$^{-/-}$*Insig2*$^{-/-}$ mice,[53] suggesting that the INSIG proteins promote VLDL secretion and/or inhibit VLDL clearance *in vivo*.

There has been limited evidence for association of genetic variation in *INSIG1* with endogenous plasma TG levels,[54,55] but here we identify an SNP associated with both *INSIG1* expression changes and TG statin response. In our study, the A (minor) allele of rs73161338 appears to blunt the normal statin-induced increase in *INSIG1* mRNA levels. As it is located about 95 kb from the transcription start site of *INSIG1*, rs73161338 could affect *INSIG1* expression changes by modifying binding to an enhancer or it could be in LD with other genetic variation that more directly affects *INSIG1* gene expression changes but was not genotyped or well imputed in our data set. With smaller increases in *INSIG1* mRNA with statin treatment, we hypothesize that less new

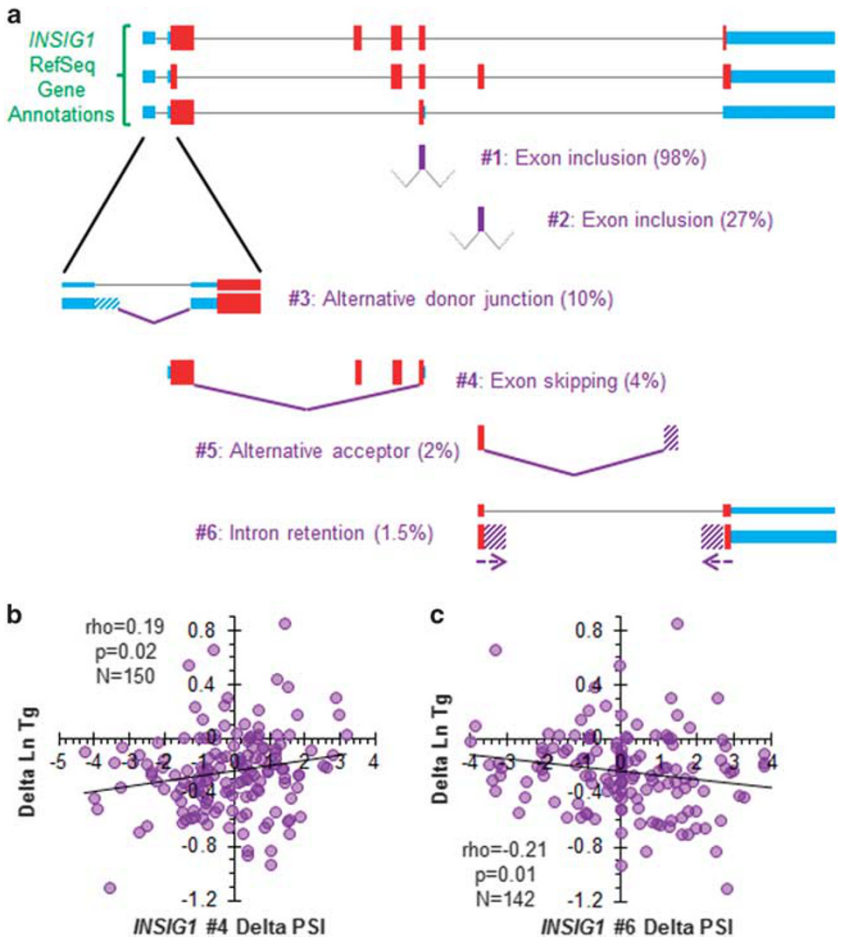

Figure 3. Insulin-induced gene 1 (*INSIG1*) splicing events and correlations of *INSIG1* splicing statin response with plasma triglyceride (TG) statin response. (**a**) Percent-spliced-in (PSI) values for six common *INSIG1* splice variants were quantified in control- and statin-treated lymphoblastoid cell lines (LCLs) using JuncBASE. (**b, c**) Statin-induced changes in (**b**) Event #4 and (**c**) Event #6 PSI values were tested for correlation with delta log TG using Spearman's correlation. Event #6 measures the proportion of sequence reads that cross an exon–intron boundary on either end of the intron, and thus may not reflect complete intron retention.

INSIG1 protein is made to replenish the INSIG1 proteins that are degraded as a result of statin-induced sterol depletion. Consistent with the phenotypes observed in the *Insig* knockout mice, greater statin-induced reductions in INSIG1 protein could cause reduced VLDL secretion and/or increased VLDL clearance, thus causing greater reductions in plasma TG with statin treatment (Supplementary Figure S9). Though this is a plausible model for how statins and rs73161338 influence plasma TG changes through *INSIG1*, further studies would be necessary to rule out the possibility that the *INSIG1* and TG change relationship is purely correlative.

Despite the abundance of alternatively spliced isoforms of *INSIG1* that produce different protein products and our finding that at least one *INSIG1* splicing event is statin responsive, the role of these alternate isoforms in the regulation of cellular cholesterol metabolism has not been studied to date. Since we observe that statin-induced changes in two different *INSIG1* transcripts that may code for altered protein products are correlated with TG response even after accounting for overall changes in *INSIG1* expression levels, the alternate INSIG1 isoforms may be regulated or function differently from the major isoform, but this requires further investigation. Given that the alternative splicing of *HMGCR* also accounts for some of the inter-individual variation in plasma LDL-C statin response,[24] differences in alternative splicing are an important regulatory mechanism of cellular sterol and lipid metabolism and statin response.

Another intriguing aspect of our findings is the sex specificity of the TG response correlations. Since there are many known differences in VLDL metabolism between the sexes[56,57] that may be influenced by plasma sex steroid levels[58–60] or adipose tissue distribution,[61] it is possible that there are also mechanistic differences in TG statin response between women and men. Further studies are needed to fully explore this possibility.

In summary, we leveraged RNA-seq data from control- and statin-treated LCLs derived from 150 statin clinical trial participants to identify novel statin-responsive genes and genes whose statin-induced expression changes were correlated with plasma TG statin response *in vivo*. *INSIG1*, a central regulator of cholesterol metabolism, emerged as a promising candidate modulator of TG statin response. This study is just one step toward a better understanding of factors that influence TG statin response, which could help to predict the efficacy of statins in lowering CVD risk and potentially inform the development of new therapeutics.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We would like to thank Dr Ronald Krauss for his expert advice and support throughout this project. We would also like to thank Drs Daniel Chasman and Paul Ridker for providing the PRINCE data. Lastly, we thank Arnie Acosta and Devesh Naidoo for their assistance in cell culturing and processing. This study would not have been possible without the contributions of the CAP participants.

REFERENCES

1 Shimano H. Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. *Prog Lipid Res* 2001; **40**: 439–452.
2 Branchi A, Fiorenza AM, Rovellini A, Torri A, Muzio F, Macor S *et al.* Lowering effects of four different statins on serum triglyceride level. *Eur J Clin Pharmacol* 1999; **55**: 499–502.
3 Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM *et al.* Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet* 2013; **45**: 1345–1352.
4 Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP *et al.* Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* 2014; **36**: 539–550.
5 Miller M, Cannon CP, Murphy SA, Qin J, Ray KK, Braunwald E. Impact of triglyceride levels beyond low-density lipoprotein cholesterol after acute coronary syndrome in the PROVE IT-TIMI 22 trial. *J Am Coll Cardiol* 2008; **51**: 724–730.
6 Schwartz GG, Abt M, Bao W, DeMicco D, Kallend D, Miller M *et al.* Fasting triglycerides predict recurrent ischemic events in patients with acute coronary syndrome treated with statins. *J Am Coll Cardiol* 2015; **65**: 2267–2275.
7 Simon JA, Lin F, Hulley SB, Blanche PJ, Waters D, Shiboski S *et al.* Phenotypic predictors of response to simvastatin therapy among African-Americans and Caucasians: the Cholesterol and Pharmacogenetics (CAP) Study. *Am J Cardiol* 2006; **97**: 843–850.
8 Medina MW, Krauss RM. The role of HMGCR alternative splicing in statin efficacy. *Trends Cardiovasc Med* 2009; **19**: 173–177.
9 Petretta M, Costanzo P, Perrone-Filardi P, Chiariello M. Impact of gender in primary prevention of coronary heart disease with statin therapy: a meta-analysis. *Int J Cardiol* 2010; **138**: 25–31.
10 Mora S, Glynn RJ, Hsia J, MacFadyen JG, Genest J, Ridker PM. Statins for the primary prevention of cardiovascular events in women with elevated high-sensitivity C-reactive protein or dyslipidemia: results from the Justification for the Use of Statins in Prevention: An Intervention Trial Evaluating Rosuvastatin (JUPITER) and meta-analysis of women from primary prevention trials. *Circulation* 2010; **121**: 1069–1077.
11 Kostis WJ, Cheng JQ, Dobrzynski JM, Cabrera J, Kostis JB. Meta-analysis of statin effects in women versus men. *J Am Coll Cardiol* 2012; **59**: 572–582.
12 Chasman DI, Posada D, Subrahmanyan L, Cook NR, Stanton VP Jr, Ridker PM. Pharmacogenetic study of statin therapy and cholesterol reduction. *JAMA* 2004; **291**: 2821–2827.
13 Donnelly LA, Doney AS, Dannfald J, Whitley AL, Lang CC, Morris AD *et al.* A paucimorphic variant in the HMG-CoA reductase gene is associated with lipid-lowering response to statin treatment in diabetes: a GoDARTS study. *Pharmacogenet Genomics* 2008; **18**: 1021–1026.
14 Krauss RM, Mangravite LM, Smith JD, Medina MW, Wang D, Guo X *et al.* Variation in the 3-hydroxyl-3-methylglutaryl coenzyme a reductase gene is associated with racial differences in low-density lipoprotein cholesterol response to simvastatin treatment. *Circulation* 2008; **117**: 1537–1544.
15 Thompson JF, Hyde CL, Wood LS, Paciga SA, Hinds DA, Cox DR *et al.* Comprehensive whole-genome and candidate gene analysis for response to statin therapy in the Treating to New Targets (TNT) cohort. *Circ Cardiovasc Genet* 2009; **2**: 173–181.
16 Barber MJ, Mangravite LM, Hyde CL, Chasman DI, Smith JD, McCarty CA *et al.* Genome-wide association of lipid-lowering response to statins in combined study populations. *PLoS One* 2010; **5**: e9763.
17 Chasman DI, Giulianini F, Macfadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin induced LDL-C reduction: The JUPITER Trial. *Circ Cardiovasc Genet* 2012; **5**: 257–264.
18 Deshmukh HA, Colhoun HM, Johnson T, McKeigue PM, Betteridge DJ, Durrington PN *et al.* Genome-wide association study of genetic determinants of LDL-c response to atorvastatin therapy: importance of Lp(a). *J Lipid Res* 2012; **53**: 1000–1011.
19 Hopewell JC, Parish S, Offer A, Link E, Clarke R, Lathrop M *et al.* Impact of common genetic variation on response to simvastatin therapy among 18 705 participants in the Heart Protection Study. *Eur Heart J* 2013; **34**: 982–992.
20 Postmus I, Trompet S, Deshmukh HA, Barnes MR, Li X, Warren HR *et al.* Pharmacogenetic meta-analysis of genome-wide association studies of LDL cholesterol response to statins. *Nat Commun* 2014; **5**: 5068.
21 Goodarzi MO, Taylor KD, Scheuner MT, Antoine HJ, Guo X, Shah PK *et al.* Haplotypes in the lipoprotein lipase gene influence high-density lipoprotein cholesterol response to statin therapy and progression of atherosclerosis in coronary artery bypass grafts. *Pharmacogenomics J* 2007; **7**: 66–73.
22 Bercovich D, Friedlander Y, Korem S, Houminer A, Hoffman A, Kleinberg L *et al.* The association of common SNPs and haplotypes in the CETP and MDR1 genes with lipids response to fluvastatin in familial hypercholesterolemia. *Atherosclerosis* 2006; **185**: 97–107.
23 Pedro-Botet J, Schaefer EJ, Bakker-Arkema RG, Black DM, Stein EM, Corella D *et al.* Apolipoprotein E genotype affects plasma lipid response to atorvastatin in a gender specific manner. *Atherosclerosis* 2001; **158**: 183–193.
24 Medina MW, Gao F, Ruan W, Rotter JI, Krauss RM. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. *Circulation* 2008; **118**: 355–362.
25 Medina MW, Gao F, Naidoo D, Rudel LL, Temel RE, McDaniel AL *et al.* Coordinately regulated alternative splicing of genes involved in cholesterol biosynthesis and uptake. *PLoS One* 2011; **6**: e19420.
26 Mangravite LM, Engelhardt BE, Medina MW, Smith JD, Brown CD, Chasman DI *et al.* A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. *Nature* 2013; **502**: 377–380.
27 Bolotin E, Armendariz A, Kim K, Heo SJ, Boffelli D, Tantisira K *et al.* Statin-induced changes in gene expression in EBV-transformed and native B-cells. *Hum Mol Genet* 2014; **23**: 1202–1210.
28 Albert MA, Danielson E, Rifai N, Ridker PM. Effect of statin therapy on C-reactive protein levels: the pravastatin inflammation/CRP evaluation (PRINCE): a randomized trial and cohort study. *JAMA* 2001; **286**: 64–70.
29 Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res* 2009; **37**: e123.
30 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013; **14**: R36.
31 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**: 550.
32 Stegle O, Parts L, Durbin R, Winn J. A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLoS Comput Biol* 2010; **6**: e1000770.
33 Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**: 44–57.
34 Theusch E, Medina MW, Rotter JI, Krauss RM. Ancestry and other genetic associations with plasma PCSK9 response to simvastatin. *Pharmacogenet Genomics* 2014; **24**: 492–500.
35 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.
36 Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010; **34**: 816–834.
37 Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012; **44**: 955–959.
38 Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE *et al.* Patterns of cis regulatory variation in diverse human populations. *PLoS Genet* 2012; **8**: e1002639.
39 Brooks AN, Yang L, Duff MO, Hansen KD, Park JW, Dudoit S *et al.* Conservation of an RNA regulatory map between Drosophila and mammals. *Genome Res* 2011; **21**: 193–202.
40 Yang T, Espenshade PJ, Wright ME, Yabe D, Gong Y, Aebersold R *et al.* Crucial step in cholesterol homeostasis: sterols promote binding of SCAP to INSIG-1, a membrane protein that facilitates retention of SREBPs in ER. *Cell* 2002; **110**: 489–500.
41 Sever N, Yang T, Brown MS, Goldstein JL, DeBose-Boyd RA. Accelerated degradation of HMG CoA reductase mediated by binding of insig-1 to its sterol-sensing domain. *Mol Cell* 2003; **11**: 25–33.
42 Watts GF, Barrett PH, Ji J, Serone AP, Chan DC, Croft KD *et al.* Differential regulation of lipoprotein kinetics by atorvastatin and fenofibrate in subjects with the metabolic syndrome. *Diabetes* 2003; **52**: 803–811.
43 Isley WL, Miles JM, Patterson BW, Harris WS. The effect of high-dose simvastatin on triglyceride-rich lipoprotein metabolism in patients with type 2 diabetes mellitus. *J Lipid Res* 2006; **47**: 193–200.
44 Chan DC, Watts GF, Ooi EM, Ji J, Johnson AG, Barrett PH. Atorvastatin and fenofibrate have comparable effects on VLDL-apolipoprotein C-III kinetics in men with the metabolic syndrome. *Arterioscler Thromb Vasc Biol* 2008; **28**: 1831–1837.

45 Chan DC, Nguyen MN, Watts GF, Ooi EM, Barrett PH. Effects of atorvastatin and n-3 fatty acid supplementation on VLDL apolipoprotein C-III kinetics in men with abdominal obesity. *Am J Clin Nutr* 2010; **91**: 900–906.
46 Ooi EM, Ng TW, Watts GF, Chan DC, Barrett PH. Effect of fenofibrate and atorvastatin on VLDL apoE metabolism in men with the metabolic syndrome. *J Lipid Res* 2012; **53**: 2443–2449.
47 Ng TW, Ooi EM, Watts GF, Chan DC, Barrett PH. Atorvastatin plus omega-3 fatty acid ethyl ester decreases very-low-density lipoprotein triglyceride production in insulin resistant obese men. *Diabetes Obes Metab* 2014; **16**: 519–526.
48 Forster LF, Stewart G, Bedford D, Stewart JP, Rogers E, Shepherd J *et al.* Influence of atorvastatin and simvastatin on apolipoprotein B metabolism in moderate combined hyperlipidemic subjects with low VLDL and LDL fractional clearance rates. *Atherosclerosis* 2002; **164**: 129–145.
49 Arad Y, Ramakrishnan R, Ginsberg HN. Lovastatin therapy reduces low density lipoprotein apoB levels in subjects with combined hyperlipidemia by reducing the production of apoB-containing lipoproteins: implications for the pathophysiology of apoB production. *J Lipid Res* 1990; **31**: 567–582.
50 Watts GF, Naoumova RP, Kelly JM, Riches FM, Croft KD, Thompson GR. Inhibition of cholesterogenesis decreases hepatic secretion of apoB-100 in normolipidemic subjects. *Am J Physiol* 1997; **273**: E462–E470.
51 Myerson M, Ngai C, Jones J, Holleran S, Ramakrishnan R, Berglund L *et al.* Treatment with high-dose simvastatin reduces secretion of apolipoprotein B-lipoproteins in patients with diabetic dyslipidemia. *J Lipid Res* 2005; **46**: 2735–2744.
52 Gong Y, Lee JN, Lee PC, Goldstein JL, Brown MS, Ye J. Sterol-regulated ubiquitination and degradation of Insig-1 creates a convergent mechanism for feedback control of cholesterol synthesis and uptake. *Cell Metab* 2006; **3**: 15–24.
53 Engelking LJ, Liang G, Hammer RE, Takaishi K, Kuriyama H, Evers BM *et al.* Schoenheimer effect explained--feedback regulation of cholesterol synthesis in mice mediated by Insig proteins. *J Clin Invest* 2005; **115**: 2489–2498.
54 Martin LJ, Kissebah AH, Olivier M. Accounting for a quantitative trait locus for plasma triglyceride levels: utilization of variants in multiple genes. *PLoS One* 2012; **7**: e34614.
55 Smith EM, Zhang Y, Baye TM, Gawrieh S, Cole R, Blangero J *et al.* INSIG1 influences obesity-related hypertriglyceridemia in humans. *J Lipid Res* 2010; **51**: 701–708.
56 Magkos F, Patterson BW, Mohammed BS, Klein S, Mittendorfer B. Women produce fewer but triglyceride-richer very low-density lipoproteins than men. *J Clin Endocrinol Metab* 2007; **92**: 1311–1318.
57 Wang X, Magkos F, Mittendorfer B. Sex differences in lipid and lipoprotein metabolism: it's not just about sex hormones. *J Clin Endocrinol Metab* 2011; **96**: 885–893.
58 Iverius PH, Brunzell JD. Relationship between lipoprotein lipase activity and plasma sex steroid level in obese women. *J Clin Invest* 1988; **82**: 1106–1112.
59 Desmeules A, Couillard C, Tchernof A, Bergeron J, Rankinen T, Leon AS *et al.* Post-heparin lipolytic enzyme activities, sex hormones and sex hormone-binding globulin (SHBG) in men and women: the HERITAGE Family Study. *Atherosclerosis* 2003; **171**: 343–350.
60 Smith GI, Reeds DN, Okunade AL, Patterson BW, Mittendorfer B. Systemic delivery of estradiol, but not testosterone or progesterone, alters very low density lipoprotein-triglyceride kinetics in postmenopausal women. *J Clin Endocrinol Metab* 2014; **99**: E1306–E1310.
61 Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues - the biology of pear shape. *Biol Sex Differ* 2012; **3**: 13.

Supplementary Information accompanies the paper on the The Pharmacogenomics Journal website (http://www.nature.com/tpj)