Abstract Plasma triglyceride (TG) concentration is re-emerging as an important cardiovascular disease risk factor. More complete understanding of the genes and variants that modulate plasma TG should enable development of markers for risk prediction, diagnosis, prognosis, and response to therapies and might help specify new directions for therapeutic interventions. Recent genome-wide association studies (GWAS) have identified both known and novel loci associated with plasma TG concentration. However, genetic variation at these loci explains only ~10% of overall TG variation within the population. As the GWAS approach may be reaching its limit for discovering genetic determinants of TG, alternative genetic strategies, such as rare variant sequencing studies and evaluation of animal models, may provide complementary information to flesh out knowledge of clinically and biologically important pathways in TG metabolism. Herein, we review genes recently implicated in TG metabolism and describe how some of these genes likely modulate plasma TG concentration. We also discuss lessons regarding plasma TG metabolism learned from various genomic and genetic experimental approaches. Treatment of patients with moderate to severe hypertriglyceridemia with existing therapies is often challenging; thus, gene products and pathways found in recent genetic research studies provide hope for development of more effective clinical strategies.—Johansen, C. T.; Kathiresan, S.; and Hegele, R. A.


Supplementary key words hypertriglyceridemia • hypotriglyceridemia • dyslipidemias • genetic variation • genomics • lipoproteins • resequencing

Plasma triglyceride (TG) concentration is a complex polygenic trait that follows a rightward-skewed distribution in the population (Fig. 1). As a clinical measurement, it integrates multiple TG-rich lipoprotein (TRL) species that circulate in plasma, predominantly intestinally synthesized chylomicrons (CMs) in the postprandial state and hepatically synthesized very low density lipoproteins (VLDL) in the fasted state. Epidemiological evidence indicates that plasma TG concentration is a strong independent risk factor for cardiovascular disease (CVD), suggesting that prolonged residence of plasma TRLs, especially in the postprandial state, may contribute to CVD susceptibility (1–7). Together with environmental influences, common and rare variants in multiple genes may collectively determine a patient’s plasma TG concentration. Identifying genes and genetic variants associated with plasma TG concentration will enrich our understanding of biochemical pathways involved in TRL metabolism, enabling identification of

Abbreviations: ABL, abetalipoproteinemia; ANGPTL3, angiopoietin-like 3 protein; CAD, coronary artery disease; CHREBP, carbohydrate response element binding protein; Cld, combined lipase deficiency; CM, chylomicrons; CRE, carbohydrate response element; CRP, C-reactive protein; CVD, cardiovascular disease; cQTL, expression quantitative trait locus; FCH, familial combined hypolipidemia; FCHL, familial combined hyperlipidemia; GCK, glucokinase; GCKR, glucokinase regulatory protein; GLGC, Global Lipids Genetics Consortium; GPI-HBP1, glycosylphosphatidylinositol-anchored HDL-binding protein 1; GWAS, genome-wide association studies; HHBL, homozygous hypobetalipoproteinemia; HL, hepatic lipase; HLP, hyperlipoproteinemia; HTG, hypertriglyceridemia; LD, linkage disequilibrium; LMF1, lipase maturation factor 1; LXR, liver X receptor; Ly6, lymphocyte antigen 6; MTP, microsomal TG transfer protein; PLTP, phospholipid transfer protein; PPARα, peroxisome proliferator-activator receptor α; SNP, single nucleotide polymorphism; TG, triglyceride; TRL, TG-rich lipoprotein; USF1, upstream transcription factor 1.

1To whom correspondence should be addressed.

e-mail: hegele@roberts.ca
subjects with increased susceptibility to disordered metabolism, and development of therapeutic interventions to improve plasma TG concentration and ameliorate CVD risk.

Our understanding of the genetic architecture, that is, the integrated contribution of all genetic variants to interindividual variations of plasma TG concentrations, has increased substantially following reports of genome-wide association studies (GWAS) in the literature. Briefly, GWAS test for associations between common genetic variants with frequencies >1% (called single nucleotide polymorphisms [SNPs]) and either quantitative or discrete traits (8). Common variants used in GWAS are chosen to "tag" additional variants (tagSNP) by virtue of blocks of linkage disequilibrium (LD), which cover large regions of the genome and often contain multiple genomic elements including variants, genes, and regulatory elements. GWAS cannot impute causation to a particular gene or variant underlying an association, although they can implicate new genomic regions for further study. Herein, we define a GWAS locus as a genomic region marked by a common variant statistically associated with TG. Methodological features of GWAS will not be discussed here, as comprehensive reviews of study design, execution, and interpretation can be found elsewhere (9–11).

GWAS have identified signals associated with regions containing both classically established genes and previously unknown genomic regions as determinants of plasma TG concentration. However, complementary experimental approaches, including rare variant sequencing, study of animal models, linkage and family-based studies, and functional cellular and biochemical experiments, will be required to more fully elucidate the genetic architecture of plasma TG concentration. Herein, we review recent advances in genetics of plasma TG, summarizing recent GWAS findings supplemented with results from complementary lines of genetic investigation.

TRIGLYCERIDE-ASSOCIATED LOCI IDENTIFIED FROM GENOME-WIDE ASSOCIATION STUDIES

Genome-wide association studies of population-based TG concentration

Because fasting plasma TG is a stable biochemical analyte, GWAS can combine findings from multiple epidemiologic studies to identify associated genetic loci (12–18). These earlier findings have been confirmed and expanded by the recent definitive study from the Global Lipids Genetics Consortium (GLGC) (19), which reported a meta-analysis of >100,000 subjects encompassing multiple ethnic groups and multiple lipid and CVD phenotypes. The GLGC analysis has provided the most comprehensive list of TG-associated loci to date, identifying 32 loci harboring common variants that contribute to variations in plasma TG concentrations (Table 1). Many genes at GWAS-identified

![Fig. 1. Frequency distribution of fasting plasma triglyceride (TG) concentrations. White bars represent male subjects, and black bars represent female subjects. Subjects with plasma TG concentration >3.37 mmol·L⁻¹ are in the 95th percentile, considered the threshold for hypertriglyceridemia (HTG). To convert mmol·L⁻¹ to mg·dl⁻¹, multiply by 88.6. The maximum plasma TG concentration in this sample was 45 mmol·L⁻¹. Data were obtained from the Canadian Heart Health Survey, a cross-sectional population-based study including >26,000 participants of multiple ancestries and ages from 18 to 74 years old, from metropolitan, urban, and rural areas of Canada (170).](image-url)
looser significance threshold of only. Astonishingly, virtually every TG-associated locus is HDL-C; while 11 loci had specific associations with TG, 7 loci were associated with total cholesterol, LDL-C, and (LDL-C) only; 8 loci were associated with HDL-C only; and 1) e.g., 6 loci were associated with total or LDL-cholesterol metabolism. However, while small in effect, the associated loci are statistically robust and replicable, given the unprecedented statistical power derived from GLGC. The lead GWAS-discovered SNPs at each of 32 TG-associated loci cumulatively explained 9.6% of the total variation in plasma TG concentrations, corresponding to 25–30% of the total genetic contribution to TG variability (19). The 21 loci newly discovered by GLGC, most of which contain novel TG-associated genes, adds only ~2.2% to the total variation that was accounted for in previous GWAS meta-analyses (17). The remaining sources of “missing variability” are the subject of intense speculation (20).

GLGC also provided in vitro and in vivo evidence supporting the involvement of specific genes in plasma TG metabolism (19). First, each locus was analyzed as a putative expression quantitative trait locus (eQTL) by testing for the association of common variants with mRNA transcript abundance in hepatic or adipose tissue of nearby genes. Common variants at 9/32 TG-associated loci were associated with transcript abundance at one or more genes (Table 1). Identification of an eQTL at a locus implicates promoter or regulatory elements in the modulation of plasma TG and provides strong evidence favoring involvement of genes whose expression is altered. Second, functional analysis of the HDL-cholesterol (HDL-C) and TG-associated GALNT2 locus in mouse models demonstrated that Galnt2 overexpression decreased HDL-C concentration, while RNA interference-mediated knockdown of Galnt2 increased HDL-C concentration. These results provide strong evidence for a causal association between GALNT2 and plasma lipids.

Pleiotropy or simply correlation. One lesson learned from GLGC GWAS was that pleiotropy, referring to the involvement of a single gene or genetic variant in multiple biological phenotypes, is a relatively common feature of the genetic determinants of plasma lipids. For instance, 21 of 32 TG-associated loci were associated with at least one additional lipid trait at the genome-wide significance threshold for association, according to a P-value < 5 × 10^-8 (Table 1): e.g., 6 loci were associated with total or LDL-cholesterol (LDL-C) only; 8 loci were associated with HDL-C only; and 7 loci were associated with total cholesterol, LDL-C, and HDL-C; while 11 loci had specific associations with TG only. Astonishingly, virtually every TG-associated locus is associated with another lipid phenotype, when using a looser significance threshold of P-value < 0.05. However, some of these findings may represent false-positive associations rather than true biological effects. Only GCKR was exclusively associated with plasma TG, although this locus is pleiotropic for nonlipid traits involved in carbohydrate metabolism.

The complexity of lipoprotein metabolism means that genes and biochemical pathways can be involved in metabolism of several lipoprotein classes. Thus, it is not surprising that 15/32 TG-associated loci in GLGC studies are also associated with HDL-C at genome-wide significance levels (Table 2), without further adjustment for related lipid phenotypes (19). Among jointly associated loci, one-third were associated with TG and HDL-C with similar strengths of association. Most loci, however, seemed preferentially associated with one phenotype, notably, APOA5 and MTHFR with TG, and CETP and LIPC with HDL-C, with P values differing by 233, 49, 368, and 83 orders of magnitude, respectively. These findings cannot distinguish between a mechanistic model in which genomic variations primarily modulate one lipoprotein variable, with subsequent secondary effects on the others, or in which a common single mechanism primarily affects multiple lipoprotein species concurrently. Postgenomic biochemical interactions among circulatingTRLs and HDL may ultimately underlie such associations. Interestingly, the directions of genetic associations with TG and HDL-C were opposite at most loci, except at LIPC and APOE. Furthermore, common variants at LIPC and APOE are also eQTLs, suggesting that single regulatory element loci at these loci may control multiple gene products.

Given the limitations of the GWAS approach, the mechanism(s) that underlies joint associations between TG and HDL-C cannot be simply explained. Furthermore, the strength and effective size of significantly associated loci only minimally reflect the potential biologic importance of the true causal gene; they only suggest that genetic variation in a genomic region appears to modulate the trait of interest. Therefore, loci jointly associated with >1 trait may prove to have biological effects that are quite different from what has been suggested by GWAS P values. Ultimately, biochemical experiments are needed to explain the (dys)function underlying GWAS-discovered signals.

Studies in multiethnic populations. A general limitation of GWAS is that they have been conducted predominantly with European subjects to date, although the extended patterns of LD in European populations have facilitated GWAS identification of TG-associated loci. A few replication studies have confirmed some TG associations in small multiethnic cohorts (21–23), although this has begun to change with GLGC studies in which 22/32 TG associations were replicated based on concordant direction of effect in European (n = 7,000), East Asian (n = 15,000), South Asian (n = 9,700), and African (n = 8,000) populations (19). However, substantial GWAS have still not been conducted with multiethnic populations. Differences in allele frequencies and more complex patterns of LD among populations of varying ancestry could also affect, either positively or negatively, the statistical power necessary to discover novel TG-associated loci (24), providing additional opportunities to identify loci and variants specific to ethnic groups.

The focus on European subjects also fails to identify functional variants that are restricted to populations of non-European ancestry. For example, the APOA5 variant
causing a glycine-to-cysteine substitution at position 185 (G185C), which causes a 23% reduction in LPL activity, is found in nearly 27% of Taiwanese subjects with high TG, which causes a 23% reduction in LPL activity, causing a glycine-to-cysteine substitution at position 185.

**Genome-wide association study of hypertriglyceridemia**

Do the same genes and variants that modulate plasma TG concentration in GWAS performed in the general population also contribute to pathological TG levels in patients ascertained through specialty clinics? Hypertriglyceridemia (HTG) is defined as a fasting plasma TG concentration $>$95th percentile (Fig. 1). Clinical and biochemical features in HTG subjects include eruptive, tuberos or palmar crease xanthomas, lipemia retinalis, hepatosplenomegaly, and pancreatitis (27, 28). HTG is thought to result from the accumulation of genetic determinants of plasma TG associated with defective TRL lipolysis (29) combined with environmental and metabolic stressors such as alcohol consumption, medications, renal disease, nonalcoholic fatty-liver disease, pregnancy, obesity, metabolic syndrome, and type 2 diabetes (T2D) (27), which culminate in a clinical phenotype. Together, HTG and accompanying comorbidities also significantly increase CVD risk (30). However, while both common and rare genetic determinants of HTG have been identified, the majority of genetic susceptibility in people with an HTG diagnosis remains unattributed (31). The study of HTG patients could improve understanding of the mechanisms by which genetic burden contributes to elevated TG.

**TABLE 1. GLGC-identified loci harboring common genetic variants associated with fasting plasma TG concentration**

<table>
<thead>
<tr>
<th>Locus</th>
<th>CHR</th>
<th>SNP</th>
<th>Associated traits</th>
<th>Risk allele</th>
<th>Allele frequency</th>
<th>TG effect (mg/dl)</th>
<th>$P$</th>
<th>$\text{eQTL}^b$</th>
<th>New GWAS-identified loci $^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOA5</td>
<td>11</td>
<td>rs964184</td>
<td>TC, LDL, HDL</td>
<td>G</td>
<td>0.13</td>
<td>16.95</td>
<td>$7.0 \times 10^{-50}$</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>GCKR</td>
<td>2</td>
<td>rs1260326</td>
<td>TC</td>
<td>T</td>
<td>0.41</td>
<td>8.76</td>
<td>$6.0 \times 10^{-135}$</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>LPL</td>
<td>11</td>
<td>rs12678919</td>
<td>HDL</td>
<td>A</td>
<td>0.88</td>
<td>13.64</td>
<td>$2.0 \times 10^{-11}$</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>MLXIPL</td>
<td>7</td>
<td>rs7811295</td>
<td>HDL</td>
<td>A</td>
<td>0.81</td>
<td>7.91</td>
<td>$9.0 \times 10^{-50}$</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>TRIB1</td>
<td>8</td>
<td>rs2954029</td>
<td>TC, LDL, HDL</td>
<td>A</td>
<td>0.53</td>
<td>5.64</td>
<td>$3.0 \times 10^{-55}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>APOB</td>
<td>1</td>
<td>rs1042034</td>
<td>TC, LDL, HDL</td>
<td>T</td>
<td>0.78</td>
<td>5.99</td>
<td>$1.0 \times 10^{-45}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>ANGPTL3</td>
<td>1</td>
<td>rs2131925</td>
<td>TC, LDL</td>
<td>T</td>
<td>0.68</td>
<td>4.94</td>
<td>$9.0 \times 10^{-45}$</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>APOE</td>
<td>19</td>
<td>rs139401</td>
<td>TC, LDL, HDL</td>
<td>C</td>
<td>0.64</td>
<td>5.50</td>
<td>$1.0 \times 10^{-36}$</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>CTP1</td>
<td>15</td>
<td>rs2019960</td>
<td>TC, LDL, HDL</td>
<td>T</td>
<td>0.93</td>
<td>7.83</td>
<td>$2.0 \times 10^{-26}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>FADS1-2-3</td>
<td>11</td>
<td>rs174546</td>
<td>TC, LDL, HDL</td>
<td>T</td>
<td>0.34</td>
<td>3.82</td>
<td>$5.0 \times 10^{-24}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>PLTP</td>
<td>20</td>
<td>rs4810479</td>
<td>HDL</td>
<td>T</td>
<td>0.24</td>
<td>3.32</td>
<td>$5.0 \times 10^{-19}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>HLA</td>
<td>6</td>
<td>rs2247056</td>
<td>TC, LDL</td>
<td>C</td>
<td>0.75</td>
<td>2.99</td>
<td>$2.0 \times 10^{-15}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>NAT2</td>
<td>8</td>
<td>rs1495743</td>
<td>TC</td>
<td>G</td>
<td>0.22</td>
<td>2.97</td>
<td>$4.0 \times 10^{-14}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>GALNT1</td>
<td>2</td>
<td>rs1291257</td>
<td>HDL</td>
<td>C</td>
<td>0.39</td>
<td>2.76</td>
<td>$2.0 \times 10^{-14}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>LIPC</td>
<td>15</td>
<td>rs261342</td>
<td>TC, LDL, HDL</td>
<td>G</td>
<td>0.22</td>
<td>2.99</td>
<td>$2.0 \times 10^{-13}$</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>CETP</td>
<td>3</td>
<td>rs7268084</td>
<td>TC, LDL, HDL</td>
<td>T</td>
<td>0.55</td>
<td>2.88</td>
<td>$1.0 \times 10^{-12}$</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>JMJ1C</td>
<td>10</td>
<td>rs1076713</td>
<td>T</td>
<td>0.57</td>
<td>2.38</td>
<td>$3.0 \times 10^{-12}$</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>TIMD4</td>
<td>5</td>
<td>rs1553318</td>
<td>TC, LDL</td>
<td>C</td>
<td>0.64</td>
<td>2.63</td>
<td>$4.0 \times 10^{-12}$</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>KLHL8</td>
<td>4</td>
<td>rs42177</td>
<td>T</td>
<td>0.59</td>
<td>2.25</td>
<td>$9.0 \times 10^{-12}$</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>FRMD5</td>
<td>15</td>
<td>rs2929282</td>
<td>T</td>
<td>0.05</td>
<td>5.13</td>
<td>$2.0 \times 10^{-11}$</td>
<td>Y</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>MAPK31</td>
<td>5</td>
<td>rs9668661</td>
<td>T</td>
<td>0.20</td>
<td>2.57</td>
<td>$1.0 \times 10^{-10}$</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>COBL1</td>
<td>11</td>
<td>rs10195252</td>
<td>HDL</td>
<td>T</td>
<td>0.60</td>
<td>2.01</td>
<td>$2.0 \times 10^{-10}$</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>LRPL1</td>
<td>12</td>
<td>rs1161352</td>
<td>HDL</td>
<td>T</td>
<td>0.77</td>
<td>2.70</td>
<td>$4.0 \times 10^{-10}$</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>TYW1B</td>
<td>7</td>
<td>rs13283803</td>
<td>C</td>
<td>0.96</td>
<td>7.91</td>
<td>$1.0 \times 10^{-9}$</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>PINX1</td>
<td>8</td>
<td>rs11776767</td>
<td>C</td>
<td>0.37</td>
<td>2.01</td>
<td>$1.0 \times 10^{-8}$</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>ZNF664</td>
<td>12</td>
<td>rs12310567</td>
<td>HDL</td>
<td>A</td>
<td>0.66</td>
<td>2.42</td>
<td>$1.0 \times 10^{-8}$</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>CAPN3</td>
<td>15</td>
<td>rs1412710</td>
<td>T</td>
<td>0.02</td>
<td>7.00</td>
<td>$2.0 \times 10^{-8}$</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>CYP26A1</td>
<td>10</td>
<td>rs2068888</td>
<td>G</td>
<td>0.54</td>
<td>2.28</td>
<td>$2.0 \times 10^{-7}$</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>IRS1</td>
<td>2</td>
<td>rs2484344</td>
<td>HDL</td>
<td>T</td>
<td>0.63</td>
<td>1.89</td>
<td>$2.0 \times 10^{-8}$</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>CT1F1</td>
<td>16</td>
<td>rs11649653</td>
<td>T</td>
<td>0.60</td>
<td>2.13</td>
<td>$3.0 \times 10^{-8}$</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>MSL2L1</td>
<td>13</td>
<td>rs645040</td>
<td>T</td>
<td>0.78</td>
<td>2.22</td>
<td>$3.0 \times 10^{-8}$</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>PLA2G6</td>
<td>22</td>
<td>rs5769531</td>
<td>T</td>
<td>0.60</td>
<td>1.54</td>
<td>$4.0 \times 10^{-8}$</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Risk allele refers to the minor allele at each locus, but not necessarily at the same SNP, at genome-wide significance levels.

$^b$ $\text{eQTL}$ are SNPs that are associated with $\geq 1$ transcript within 500-kb associated at $P < 0.001$ in liver and omental or subcutaneous fat.

$^c$ “Y” indicates locus was first implicated in TG metabolism identified by GWAS; “N” indicates locus was identified by other means prior to GWAS.

Note: The table lists common genetic variants associated with fasting plasma triglyceride (TG) levels. Variants are identified by their genetic marker (SNP), chromosome (CHR), and associated traits. The table includes information on the risk allele, allele frequency, TG effect in milligrams per deciliter (mg/dl), $P$-value, and whether the SNP is associated with eQTLs (expression quantitative trait locus) in liver and omental or subcutaneous fat. The table also notes whether the locus was previously identified by GWAS or other means. The study of HTG patients could improve understanding of the mechanisms by which genetic burden contributes to elevated TG. Only one GWAS study evaluating subjects at the extremes of lipid phenotypes has been conducted to date. Using a case-control design comparing $\sim$500 HTG patients with $\sim$1,200 normotriglyceridemic population-based controls (32), we showed genome-wide significant associations among the APOA5, GCKR, LPL, and APOB loci, and replicated the MLXIPL, TRIB1, ANGPTL3, and NCAN loci, with trends toward significance as determined by $P$-values near significance at the FADS1-FADS2-FADS3 (FADS1-2-3) cluster, XKR6/PINX1, and PLTP. Thus, loci that determine...
plasma TG concentration in the general population also underlie HTG susceptibility in patients. Furthermore, we demonstrated that 29/32 TG-raising alleles associated with TG in the GLGC analyses were overrepresented in HTG patients compared with controls, leading to an increased genetic burden of risk (Fig. 2) (19). However, replication of all 32 newly identified TG-associated loci will be required in a larger HTG cohort, because larger sample sizes will more likely detect variants with clinically relevant effects. Although our study identified and replicated several TG-associated loci in HTG pathophysiology, it was not exhaustive and does not exclude additional loci that potentially modulate plasma TG concentration or determine HTG risk. A complete list of HTG-associated loci is shown in Table 3.

We next searched for rare variations in genes under GWAS peaks by resequencing the protein coding exons of APOA5, GCKR, LPL, and APOB (32). We demonstrated a significant accumulation of rare variants in HTG patients, defined by a minor allele frequency of <1% in controls. Several variants found in HTG patients were either truncation mutations or were predicted to have deleterious effects in silico or were previously demonstrated in vitro to be dysfunctional, particularly in LPL, a long-established functional candidate for TG metabolism. In aggregate, 154 rare missense and nonsense variants were seen in 28.1% of 438 HTG patients compared with 53 rare variants found in 15.3% of 327 controls. A more stringent analysis restricted to variants found exclusively in HTG patients or controls confirmed the mutation skew: 47 rare variants were found in 10.3% of HTG patients, while only 9 rare variants were found in 2.8% of controls. Carrying rare variants thus decreased the unattributed variation in the estimation of variation attributable to rare variants in HTG patients.

An emerging mosaic genetic model of hypertriglyceridemia. The overall HTG phenotype is heterogeneous due to distinct differences in populations of TRLs (33), classically defined using the Fredrickson HLP phenotype system based on lipoprotein fractions quantified by ultracentrifugation and electrophoresis (28). However, genetic studies suggest that HLP phenotypes are essentially indistinguishable at the level of common SNP genotypes (34). We suggest that genetic susceptibility requires a minimal burden (a "quorum") of common HTG risk alleles, with rare loss-of-function variants that also predispose to HTG (Table 3). This minimal number of common variants is associated with susceptibility to develop moderate HTG (HLP type 4). The "stacking" of additional HTG risk alleles on top of the quorum leads to severe HTG (HLP type 5). The presence of the APOE e2/e2 genotype in the background of the quorum of HTG risk alleles is associated with dysbetalipoproteinemia (HLP type 3). The presence of LDL-associated alleles drives the phenotype toward HLP type 2B. Thus,

### TABLE 2. GLGC meta-analysis of genetic loci associated with concentrations of both plasma TG and HDL-C

<table>
<thead>
<tr>
<th>Locus</th>
<th>Lead trait</th>
<th>Role in lipoprotein metabolism</th>
<th>TG P</th>
<th>HDL P</th>
<th>SNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOA5</td>
<td>TG</td>
<td>Activator of lipoprotein lipase</td>
<td>7×10⁻²⁰</td>
<td>5×10⁻⁴⁷</td>
<td>rs964184</td>
</tr>
<tr>
<td>LPL</td>
<td>TG</td>
<td>Hydrolysis of TG-rich lipoproteins</td>
<td>2×10⁻¹⁵</td>
<td>9×10⁻⁸⁸</td>
<td>rs3278919</td>
</tr>
<tr>
<td>MLX1P</td>
<td>TG</td>
<td>Activation of glycolytic and lipogenic enzymes</td>
<td>9×10⁻⁹⁹</td>
<td>1×10⁻⁹⁹</td>
<td>TGrs7811265</td>
</tr>
<tr>
<td>TRIB1</td>
<td>TG</td>
<td>Unknown</td>
<td>3×10⁻⁵⁵</td>
<td>6×10⁻¹⁵</td>
<td>HDL rs1745758</td>
</tr>
<tr>
<td>APOB</td>
<td>TG</td>
<td>Backbone of atherogenic lipoproteins</td>
<td>1×10⁻⁴⁵</td>
<td>1×10⁻³⁰</td>
<td>TG rs9349947</td>
</tr>
<tr>
<td>APOE</td>
<td>TG</td>
<td>TG-rich lipoprotein receptor ligand for the LDLR and LDLR-related protein (LRP1)</td>
<td>1×10⁻⁵⁰</td>
<td>1×10⁻²¹</td>
<td>HDL rs10988546</td>
</tr>
<tr>
<td>FADS</td>
<td>TG</td>
<td>Fatty acid desaturation</td>
<td>5×10⁻²⁴</td>
<td>2×10⁻²²</td>
<td>TG rs174546</td>
</tr>
<tr>
<td>PLTP</td>
<td>HDL-C</td>
<td>Transfer phospholipid between TG-rich lipoprotein and HDL</td>
<td>5×10⁻¹⁸</td>
<td>2×10⁻²²</td>
<td>HDL rs6089096</td>
</tr>
<tr>
<td>GALNT2</td>
<td>HDL-C</td>
<td>0-linked glycosylation of proteins</td>
<td>2×10⁻¹⁴</td>
<td>4×10⁻²¹</td>
<td>TG rs4810479</td>
</tr>
<tr>
<td>LIPC</td>
<td>HDL-C</td>
<td>TG lipase</td>
<td>2×10⁻¹³</td>
<td>3×10⁻⁶⁶</td>
<td>HDL rs4846914</td>
</tr>
<tr>
<td>CETP</td>
<td>HDL-C</td>
<td>Exchanges cholesterol ester for TG between HDL and TG-rich lipoproteins</td>
<td>1×10⁻¹²</td>
<td>7×10⁻⁰⁸⁰</td>
<td>TG rs1321257</td>
</tr>
<tr>
<td>COBLL1</td>
<td>TG</td>
<td>Unknown</td>
<td>2×10⁻¹⁰</td>
<td>3×10⁻¹⁰</td>
<td>HDL rs1532085</td>
</tr>
<tr>
<td>LRP1</td>
<td>TG</td>
<td>TG-rich lipoprotein receptor via apoE</td>
<td>4×10⁻¹⁰</td>
<td>2×10⁻⁸⁸</td>
<td>TG rs10195252</td>
</tr>
<tr>
<td>IRS1</td>
<td>HDL-C</td>
<td>Involved in insulin signaling</td>
<td>2×10⁻⁸</td>
<td>2×10⁻⁹⁹</td>
<td>HDL rs3741414</td>
</tr>
<tr>
<td>ZNF664</td>
<td>HDL-C</td>
<td>Unknown</td>
<td>1×10⁻⁸</td>
<td>3×10⁻¹⁰</td>
<td>HDL rs2943645</td>
</tr>
</tbody>
</table>

HDL-C, high-density lipoprotein cholesterol; LDLR, low-density lipoprotein receptor; SNP, single nucleotide polymorphism; TG, triglyceride.

*Lead trait* is the strongest associated trait.

*SNP column is the strongest trait-associated SNP for TG and HDL-C; a single SNP is the strongest associated variant for both traits.*
The multisystem clinical phenotypes for ABL and HHBL include retinopathy, disturbed bone metabolism, neuropathy, myopathy, and coagulopathy and stem from deficiency in fat-soluble vitamins. Atherosclerosis risk in ABL and HHBL patients is low, probably due to reduced LDL-C rather than to reduced plasma TG levels. Additional features of ABL and HHBL, and also heterozygous hypobetalipoproteinemia, is hepatosteatosis, suggesting that the pathogenic accumulation of TG within hepatocytes is related to a deficiency of circulating TRLs. The biochemical phenotype of FCH resembles that of HHBL (35), although mutations causing FCH likely do not interfere with TRL assembly and export, because hepatosteatosis and other clinical phenotypes characteristic of ABL and HHBL are not observed.

New treatments directed against these molecular targets, for example, pharmacological inhibition of MTP in the case of lomitapide (38) and RNA interference of APOB in the case of mipomersen (39), should lower both LDL-C and TG levels, with a potential increased risk of hepatosteatosis. Molecular inhibition of ANGPTL3 would be predicted to reduce circulating lipoproteins, potentially without off-target effects including hepatosteatosis, given the absence of additional pathologies in FCH patients with loss-of-function mutations in ANGPTL3 (35).

Functional variants in classical triglyceride-associated genes

Lipoprotein lipase. The gene encoding LPL is located on chromosome 8, and its product has long been recognized as the fulcrum for hydrolysis of plasma TRLs (40). LPL is highly synthesized in tissues requiring free fatty acids for energy metabolism, predominantly heart, adipose tissue, and skeletal muscle (41). It is secreted into the vasculature supplying these tissues, where it binds to glycosylphosphatidylinositol-anchored HDL-binding protein 1 (GPIHBP1) on endothelial cell surfaces (42). LPL hydrolyzes TRLs to release free fatty acids, which may be used for TG resynthesis in adipose tissue or lipid oxidation in muscle. LPL is under strong dietary and hormonal regulation (43), and its activity depends upon additional cofactors, most notably apoC-II which is absolutely required for hydrolysis (44). Murine models lacking Lpl have markedly increased plasma TG concentration, with VLDL accumulation after birth and chylomicronemia after suckling (45). TRL accumulation is lethal within 18 h of onset of feeding. Conversely, overexpression of Lpl in mice confers a protective phenotype, improving clearance of both CM and VLDL and reducing plasma TG concentration nearly 75% (46).

The lead TG-associated Lpl variant (rs12678919) is an intergenic SNP located downstream of Lpl, the only gene within a large block of LD spanning the GWAS-described
locus. The minor allele frequencies range from 9% to 14%, lowest in African and highest in European and Japanese populations. In population-based GWAS (19), this variant increases plasma TG concentration by 0.16 mmol·l⁻¹ (13.95 mg·dl⁻¹). A second HTG-associated LPL variant (rs7016880) downstream of LPL is strongly associated with the population-based associated SNP (r² = 0.86) conferring nearly 3-fold increase in risk of HTG (32). Some tagging SNPs associated with the LPL locus have further been associated with plasma TG in multiple ethnicities, with allele frequencies differing by up to 10% in various populations (21, 22). The association of lead LPL-associated variants with plasma TG concentration indicates that common variants at this locus affect plasma TG; however, the precise functional variants have not been identified.

Several LPL variants have systematically been identified as determinants of plasma TG concentration and CVD risk (47). Notably, functional coding sequence changes include aspartic acid-to-asparagine substitution at position 9 (D9N), which decreases both LPL activity and mass, resulting in a 25% increase in plasma TG concentration among carriers (48); the asparagine-to-serine substitution at position 291 (N291S), which decreases LPL activity and prolongs TRL clearance (52). The LPL D9N variant is significantly overrepresented in patients with HTG, while the N291S variant is not (53). Conversely, the LPL S447X variant has been repeatedly associated with lower TG and higher HDL-C concentrations (54) and is less prevalent in HTG patients (34, 55). Interestingly, the S447X variant is now the basis of a gene therapy strategy, alipogene tiparvovec, for treatment of severe chylomicronemia (56).

Rare loss-of-function LPL variants are also associated with HTG. Both homozygous and compound heterozygous loss-of-function LPL mutations cause childhood-onset HTG, with prevalence of ~1 case in 1,000,000 population (27, 28). Numerous functional variants reduce LPL activity when in the homozygous state, preventing hydrolysis and resulting in TRL accumulation (57–61). Interestingly, homozygous APOC2 mutations also cause HTG (31, 62). The APOC2 mutations were first definitively shown to cause HTG after they were characterized at the amino acid level (40). Rare heterozygous LPL mutations also accumulate in patients with late-onset HTG (32) and were observed in ~6.4% of these patients (53). Both common and rare LPL variants are thus strongly associated with both population-based plasma TG concentration and diagnosis of HTG.

**Apolipoprotein A-V.** The APOA5 gene was bioinformatically identified within the APOA5–APOA4–APOC3–APOA1 gene cluster on chromosome 11 (63, 64). It encodes the liver-expressed apoA-V, which enhances LPL activity (65–67). About 80% of apoA-V associates with CMs, VLDL, and HDL at very low concentrations (68), while ~20% is retained in an intrahepatic pool associated with lipid droplets (69). Plasma apoA-V likely affects the distribution of apoC-III molecules on VLDL and thus promotes lipolysis (70); however, the role of intracellular apoA-V is less clear. In mice, overexpression of the human APOA5 gene markedly decreases plasma TG concentration, whereas mice lacking the Apoa5 gene become severely hypertriglyceridemic (63, 71). APOA5 is regulated by several transcription factors that are involved in plasma lipoprotein and glucose homeostasis, suggesting it responds to multiple environmental cues (72); however, such responses may differ between mice and humans. Peroxisome proliferator-activator receptor α (PPARα) agonists, such as fibrates, that induce APOA5 expression in humans do not equivalently induce expression of Apoa5 or transgenic human APOA5 in mice, the result of a nonfunctional PPARα response element in the mouse Apoa5 promoter sequence (73). Regardless, the APOA5 gene is clearly a crucial determinant of TRL metabolism in both mice and humans, although its regulation and molecular function are incompletely understood.

The lead TG-associated APOA5 variant (rs964184) is an intergenic SNP located downstream of the gene cluster containing APOA5. Its minor allele frequency ranges from 12% to 33%, lowest in European and highest in Japanese populations. In population-based GWAS (19), this locus was identified.
individuals. In GWAS of population-based samples (19), each APOA5 risk allele increases plasma TG by 0.19 mmol·l⁻¹ (16.95 mg/dl). Interestingly, the same variant was also associated with HTG: the minor allele frequency was 33% in HTG patients of European descent compared with 14% in controls, increasing the HTG odds ratio by 3.3-fold (32). The APOA5 variants are also associated with plasma TG in multiple ethnicities (21, 22). As with LPL, the precise functional variant(s) that leads to the GWAS-determined association by virtue of LD, and the nature of their dysfunction, are not known.

Several common functional APOA5 variants have been identified (74). However, because these variants are largely not genotyped by the International HapMap Consortium, it is impossible to determine whether they underlie the associations with the APOA5 lead SNP (rs964184). For instance, a coding sequence variant causing a serine-to-tryptophan substitution at amino acid residue 19 (S19W) in the signal peptide of apoA-V (75) impairs protein translocation and secretion (76). Several promoter variants have also been identified (63, 75), including −3A>G in the APOA5 Kozak sequence and −1131T>C. Functional analyses in vitro suggest that it is primarily the −3A>G variant that decreases APOA5 expression, whereas −1131T>C has minimal effect on expression (76, 77). Interestingly, the −1131T>C variant is located on a haplotype distinct from S19W and is in strong LD with functional APOC3 variants (75, 78), suggesting that could underlie the APOC3 association with plasma TG (78, 79). Both the S19W and the −1131T>C variants are also strong predictors of HTG (34, 55, 80) and increased CVD risk (81, 82). The APOA5 S19W carriers are >6 times more likely to be diagnosed with severe HTG than controls (80). Rare loss-of-function APOA5 variants are also associated with HTG (32, 53, 83). Functional analysis of such variants has generally revealed loss of LPL activity, particularly with C-terminal truncation mutations that interfere with apoA-V ability to interact with lipid and lipoproteins (61). Finally, 41.8% of patients with severe HTG were carriers for either the APOA5 S19W mutation or at least one of the above-mentioned rare APOA5, LPL, or APOC2 variants compared with only 8.9% of controls who carried these variants (53).

Apolipoprotein C-III. The APOC3 gene is also part of the APOA5-APOA4-APOC3-APOA1 gene cluster on chromosome 11. ApoC-III inhibits LPL-mediated TG hydrolysis by opposing the effect of apoC-II, and the apoC-II/apoC-III ratio determines net activation or inhibition of LPL (84, 85). Intracellular apoC-III also promotes assembly and secretion of VLDL particles (86). Both putative functions of apoC-III are consistent with in vivo studies showing HTG in mice overexpressing APOC3 (87) and hypotriglyceridemia and enhanced postprandial TG clearance in APOC3-deficient mice (88). In humans, linkage studies have implicated APOC3 as a candidate gene for familial combined hyperlipidemia (FCHL); however, specific common variants have not been associated with disease (89).

APOC3 is not strictly speaking a GWAS-identified locus. The APOA5-APOA4-APOC3-APOA1 locus produces the strongest association signal in GWAS of plasma TG concentration, although there is ambiguity regarding those causal genes and variants that underlie this signal. Strong yet complex patterns of LD encompass APOA5 and APOC3; therefore, common functional APOC3 variants may also contribute to the GWAS-identified signal (75, 78). Regardless, APOC3 is a crucial determinant of plasma TG concentration. Two functional variants, namely -482C>T and -455T>C, in the APOC3 regulatory sequences have been shown to attenuate APOC3 insulin responsiveness in vitro (90) and to increase fasting plasma TG concentration by ~24-fold in Asian Indian men (91). These APOC3 variants are further associated with increased risk of metabolic syndrome in multiethnic studies (92, 93).

Because APOC3 inhibits TRL lipolysis, rare APOC3 variants increase lipolysis and reduce plasma TG. For example, GWAS of fasting and postprandial TG in 806 Old Order Amish subjects implicated APOC3, and resequencing identified a heterozygous nonsense mutation at residue 19 (R19X) in 2.8% of subjects; plasma apoC-III concentrations were reduced 50% compared with that in noncarriers (94). This variant was associated with a favorable plasma lipoprotein profile: decreased plasma TG and LDL and increased plasma HDL-C. Subclinical atherosclerosis, as measured by electron beam computed tomography, was also significantly reduced in carriers. Another rare APOC3 variant, the alanine-to-threonine missense mutation at residue 23 (A23T), was identified in three YuCatan Indian subjects with apoC-III deficiency (95). Functional characterization of A23T in vitro demonstrated attenuated VLDL assembly and secretion from hepatocytes, with TG accumulation within the microsomal lumen that was independent of MTP activity (96). Perhaps pharmacologic inhibition of APOC3 could improve elevated plasma TG phenotypes, but additional mechanistic evidence is required to definitively establish the benefit of inhibiting APOC3 expression or function, given the complexity of phenotypes and mechanisms of the genes and gene products at this locus.

Angiopoietin-like 3 protein. Angiopoietin-like 3 protein (ANGPTL3) was initially identified as a regulator of lipid metabolism in mice (97, 98). A mutant hypolipidemia phenotype was mapped to this locus, revealing that disruption of Angt3 decreased plasma TG concentrations in addition to improving other lipid parameters including total cholesterol and nonesterified fatty acids, while ANGPTL3 overexpression significantly worsened the lipoprotein profile (98). ANGPTL3 appears to reversibly inhibit LPL catalytic activity (99, 100), which is inhibited in turn by LPL-stabilizing proteins such as GPIHBP1 (101). ANGPTL3 is abundant in plasma (102), which may reflect its low affinity and reversible LPL binding, and larger concentrations may be required to effectively inhibit LPL activity (100). ANGPTL3 expression is regulated by the promoter liver X receptor (LXR) response element, increasing plasma concentration in response to LXR agonists (103).

The lead SNP identifying ANGPTL3 as the TG-associated locus (rs2131925) is an intronic SNP in the DOCK7 gene,
corresponding to a location upstream of ANGPTL3. The frequency of this allele varies among different populations, ranging from 88% in Japanese to 25% in African individuals; thus, in addition to variable LD patterns, variable allele frequencies might explain nonreplication of ANGPTL3 associations in multiethnic studies (21, 22). Interestingly, the ANGPTL3 GWAS-described signal is an eQTL for ANGPTL3 expression (17, 19), and the risk allele associated with increased expression is also associated with increased plasma TG concentration. This suggests that common ANGPTL3 variants likely lie within regulatory elements; whether they affect multiple genes in TRL metabolism remains to be determined.

Population-based resequencing in the Dallas Heart Study (n = 3,551 subjects) has revealed several heterozygous nonsynonymous sequence variants that compromised protein synthesis or secretion of ANGPTL3, in turn decreasing plasma TG concentration (104). In ANGPTL3, 15 nonsynonymous variants were identified in subjects at the lower 25th percentile of population-based TG concentration versus 5 variants identified in subjects at the upper 25th percentile (104). Furthermore, ANGPTL3 belongs to a family of seven angiopoietin-like proteins involved in angiogenesis and lipoprotein metabolism, including ANGPTL4, which also carries nonsynonymous sequence variants with synthesis or secretion defects that modulate plasma concentrations of TG and HDL (104, 105). Similar extreme sampling within this cohort revealed 13 nonsynonymous variants in subjects with low TG versus 2 variants in subjects with higher TG concentrations (105). Such variants likely extend to subjects with very low circulating plasma TG in other populations, such as in subjects with FCH (35).

**Novel TG-associated loci**

Identification and replication of established genes in lipoprotein metabolism, serving as “positive controls” for the GWAS approach, have supported GWAS methodology as the tools with which to identify novel TG-associated loci of likely biological relevance. However, unlike classically established loci, the precise genes and variants responsible for novel GWAS peaks are unclear. The genes described below are likely biological candidates underlying their respective TG-associated signals, either because of prior implication in lipoprotein metabolism or because they are the sole gene at the GWAS-identified locus. Future studies using larger samples, fine mapping, deep resequencing, and bioinformatic analyses will be required to elucidate the function of novel genes in TG metabolism.

**Glucokinase regulatory protein.** The GCKR locus, encoding the glucokinase (GCK) regulatory protein (GCKR), is the strongest completely novel locus implicated in TG metabolism (106). It has been robustly replicated in GWAS of plasma TG concentration (19) and HTG (32) and implicated in HTG by an excess of rare variants in patients (32). GCKR is expressed in the liver and acts as an allosteric regulator of GCK, allowing it to be rapidly mobilized in response to increased cellular glucose concentrations (107).

A functional GCKR variant, a proline-to-leucine substitution at position 446 (P446L) that is commonly genotyped using high-density microarrays, may be the basis of this association (108). Studies in vitro suggest that allosteric regulation of GCK is attenuated by this GCKR variant, disrupting its sensitivity to physiological concentrations of fructose-6-phosphate (109). It mediates a dampened response to GCK inhibition in response to endogenous agonists, indirectly promoting GCK activity (109), which might increase glycolytic flux, increasing cellular glucose uptake, de novo TG synthesis and inhibition fatty acid oxidation.

Interestingly, GCKR is also associated with diverse traits that could modulate CVD risk and link carbohydrate with lipoprotein metabolism. The same risk allele for increased plasma TG concentration is associated with improved glycemia indices, including reduced plasma glucose, improved insulin sensitivity, and decreased T2D risk (108, 110–113). The functional GCKR variant is associated with other phenotypes, such as plasma C-reactive protein (CRP) and serum uric acid (108, 114). How the functional GCKR variant mediates these other interactions is unknown.

**Carbohydrate response element binding protein.** The MLXIPL locus encodes the carbohydrate response element (CRE) binding protein (CHREBP). The lead TG-associated SNP (rs7811265) is located downstream of MLXIPL, in a block of LD encompassing several genes; however, MLXIPL is a strong candidate gene. The lead SNP is associated with MLXIPL transcript abundance in subcutaneous fat, although with an atypical U-shaped direction of effect across genotypes, and will require replication and functional evaluation (19). In any event, the minor allele is robustly associated with lower TG concentration (19) and protection from HTG (32). CHREBP is a basic helix-loop-helix transcription factor expressed in lipogenic tissues, responsible for the coordinate activation of glycolytic enzymes such as GCK and lipogenic enzymes such as fatty acid synthase, required to convert dietary carbohydrate to TG (115).

CHREBP activates target genes in TG synthesis by binding to a CRE in the promoter (116, 117). The phenotype of Mlxipl knockout mice is consistent with the protective effect of the minor allele, as seen in humans, showing delaying activation of CHREBP and dampened cellular glycolytic and lipogenic programs (118). These mice have decreased hepatic TG content, decreased plasma free fatty acids, and suppression of both glycolysis and lipogenesis. The association between CHREBP and TRL metabolism should invigorate interest in metabolic pathways downstream from MLXIPL, as responsive genes may suggest potential targets for intervention to reduce plasma TG concentration.

**Drosophila Tribbles homolog 1.** The lead TG-associated SNP (rs2954029) is located downstream of the gene, in a block of LD spanning only the TRIB1 locus, suggesting it is probably the gene underlying the GWAS association signal. TRIB1 is expressed in most tissues but is most abun-
dant in skeletal muscle, thyroid, pancreas, blood leukocytes, and bone marrow (119). TRIB1 modulates mitogen-activated protein kinase activity in vascular smooth muscle cell proliferation and also chemotaxis (120). Trib1 knockout mice have impaired macrophage activation of signal transduction pathways involved in toll-like receptor activation in response to lipopolysaccharides, but lipoprotein phenotypes were not investigated in model systems (121).

Apolipoprotein B. The APOB gene is another familiar player in lipoprotein metabolism. Prior to GWAS, variation in APOB was most often associated with LDL-C (122). Only recently has APOB been implicated as a TG-associated locus. GWAS have identified two signals on APOB: 1) a missense variant (rs1042034) causing a serine-to-isoleucine substitution at position 4338 (S4338I) was associated with population-based TG levels (19); and 2) a variant ∼120 kb upstream of APOB (rs4635554) was associated with increased HTG risk (32). Neither variant has been functionally implicated in the modulation of plasma TG; however, we anticipate that they are merely tagSNPs with sufficient power to detect associations at this locus. Additional direct evidence is provided by the observation that rare variants in APOB accumulate in HTG patients (32), suggesting that by an unknown mechanism, mutations are capable of increasing plasma TG concentration in addition to causing disorders such as familial defective apoB-100, where LDL-C is increased but TRLs are typically not affected (123), and hypobetalipoproteinemia, where TRLs and LDL-C are decreased (124).

Phospholipid transfer protein. The phospholipid transfer protein (PLTP) locus is considered to be involved primarily in HDL metabolism but is similarly associated with plasma TG concentration (Table 2). The lead SNP in PLTP (rs4810479) is located upstream of the gene; however, it was not among eQTLs identified by the GLGC (19), suggesting that the underlying functional variant likely affects protein function. PLTP is secreted into the plasma by multiple tissues, where it shuttles phospholipids from TRLs to HDL and may remodel HDL producing pre-β HDL particles (125). Its expression is modulated by bound farnesoid X receptor (FXR) and LXRs (126). Several common variants carried by PLTP are associated with decreased PLTP expression, decreased activity, increased number of small HDL molecules, and a protective cardiovascular effect, including the lead SNP identified by GLGC; however, measures of TRL metabolism were not investigated in depth (127, 128). In mice, overexpression of Pltp modestly increased HDL-C, while Pltp knockout mice had reduced HDL-C, phospholipids, and apolipoprotein A-I (129, 130). Other studies in mice have suggested that Pltp contributes to TG-rich particle assembly by increasing VLDL secretion (131, 132). The emergence of PLTP as a TG-associated GWAS-discovered locus solidifies the rationale to better understand its role in lipoprotein metabolism.

Fatty acid desaturases. The FADS1-2-3 gene cluster is associated with plasma TG and inversely with plasma HDL-C concentrations (Table 2) (17, 19). These three genes each encode fatty acid desaturases responsible for the metabolism of polyunsaturated fatty acids involved in cell signaling (133); however, little is known about the role of desaturases in TG or HDL metabolism. Interestingly, the lead TG-associated SNP (rs174546) is found within the 3’ untranslated region of FADS1, and the lead HDL-C-associated SNP (rs174601) is within an intron of FADS2. However, both SNPs are located within a large LD block and are in strong LD with each other (r² = 0.86). The FADS1-associated SNP is also an eQTL in liver and omental fat (19). Common variants at this locus have also been associated with serum phospholipid concentrations including arachidonic acid (134, 135), desaturase activity (increased arachidonic-to-linoleic acid ratio), increased CRP concentration and CVD risk (136), phospholipid subfractions including sphingomyelin species (137), and increased plasma glucose and T2D risk (112). These findings suggest a proinflammatory response mediated by increased desaturase activity, although the mechanistic link between these pathways and lipoprotein metabolism is unclear. Similarly, it is uncertain whether one or all of these genes are involved in TRL metabolism.

GENETIC DETERMINANTS OF TRIGLYCERIDE FOUND USING OTHER APPROACHES

Important genes are not always identified by GWAS. There are several possible reasons for this: 1) a locus encodes a crucial protein, but there is no structural or functionally relevant genetic variation in humans; 2) only very rare functional variants exist at a locus, and these have no consistent LD relationship with the common genetic variation that is genotyped on microarrays; 3) common functional genetic variation at a locus affects TG metabolism but is not included on the genotyping microarray nor is it in strong LD with surrogate markers on microarrays; 4) common functional variation has too small an effect size to be significantly associated; 5) GWAS studies that evaluate an additive mode of inheritance may miss associations with variants that act through simple dominant or recessive models, or nonadditive gene–gene or gene–environment interactions, or non-Mendelian mechanisms including inheritance of mitochondrial DNA; or 6) common variation is present at a biologically important locus, detectable and included on genotyping microarrays, but it has no functional consequence. Thus, non-GWAS methodologies such as linkage and family based studies or analysis of human homologs of animal models may help with gene discovery. In the following sections we describe TG-associated genes identified using such complementary methodologies.

Linkage and family studies

Upstream transcription factor 1. Linkage analysis is a classical genetic approach to statistically implicate chromosomal loci (and ultimately genes) that cosegregate with
disease phenotypes which show vertical transmission in multigenerational families or kindreds. The approach is most powerful for disorders in which a single gene completely explains the phenotype.

Among familial HTG disorders in which linkage analysis has been attempted is FCHL. FCHL is a relatively common familial dyslipidemia defined by increased plasma concentrations of total and LDL-C, TG, and apoB and by decreased plasma concentrations of HDL-C (89). The mechanism underlying FCHL has been considered the overproduction of apoB-containing lipoproteins together with delayed clearance of TRLs. Although it does not demonstrate pure autosomal dominant inheritance in most affected families, it is nonetheless strongly heritable; the phenotypic variability in family members has sometimes been attributed to variable penetrance or to gene-diet or gene-hormone interactions. Classical genetic linkage approaches have been applied in specific FCHL families (89). Notably, linkage mapping was used to identify the 1q21-q23 locus as associated with the FCHL-associated locus reported in multiple independent cohorts (138–140), shown to be mediated by upstream transcription factor 1 (USF1) (141).

USF1 encodes a transcription factor responsible for the regulation of many proteins involved in glucose and lipid metabolism, including apoA-V and apoC-III and metabolic enzymes such as fatty acid synthase and GCK (142, 143). Recently, the risk allele of a strongly associated intronic variant in USF1 was shown to attenuate the insulin induction of USF1, thus preventing activation of responsive genes (144). Coding sequence variants in USF1 are rare, suggesting that this intrinsic variant is the functional variant underlying dyslipidemia. The potential functional variant is found in a highly conserved nucleotide in a putative transcriptional regulatory element, which may alter expression of USF1 target genes including APOE (145).

Human homologs of mouse models of hypertriglyceridemia

Combined lipase deficiency (lipase maturation factor 1). Combined lipase deficiency (cld) was first described in a spontaneous mouse mutant that was deficient in both LPL and hepatic lipase (HL) activity (146). Mice carrying autosomal recessive mutations in Cld develop postpartum chylomicronemia and death from ischemia and cyanosis from increased blood viscosity after nursing (146, 147). Using linkage analysis, the cld phenotype was mapped to murine chromosome 17, and further fine mapping identified the Tmem112 gene, renamed lipase maturation factor 1 (Lmf1), as the cause (147). Subsequent studies showed that Lmf1 is coexpressed in tissues that express LPL or HL; its gene product localizes to the membrane of endoplasmic reticulum, stimulating maturation of both LPL and HL (147, 148).

The cld mutation is a premature truncation of Lmf1 affecting a conserved domain of uncertain function (DUF1222) that cripples maturation of the target lipases, apparently independent of its subcellular localization and without requiring any direct interaction (148). Naturally occurring murine mutations in Lmf1 have also been demonstrated to compromise Lmf1 function in vitro (149). For instance, the Lmf1 glycine-to-glutamic acid missense mutation at position 181 (G181E), located next to a transmembrane domain, has severely compromised function (149).

The HTG phenotype in cld-carrying mice suggested that the human ortholog LMF1 on chromosome 16 was a candidate gene for TG metabolism (Table 3). In one study, 11 patients with chylomicronemia and defective lipase activity were resequenced for mutations in LMF1. A tyrosine-to-stop codon (Y439X) mutation in LMF1 exon 9 was identified in an HTG patient with recurrent pancreatitis resulting from a plasma TG concentration of ~30 mmol·l⁻¹, caused by a 93% reduction in post-heparin plasma activity of LPL (147). In a second study, another homozygous mutation was identified in a patient with recurrent pancreatitis and a plasma TG concentration of ~27 mmol·l⁻¹ (150). The tryptophan-to-stop codon (W464X) mutation was found to have reduced LPL and HL activities by 76% and 27%, respectively (150). Thus, LMF1 appears to be a bona fide gene that regulates TG metabolism, as indicated by phenotypic and functional consequences of rare large-effect mutations in humans and mice, while its common variants interrogated using GWAS failed to produce an association signal.

Homologous protein sequences that share significant sequences or even protein domain structures may point toward genes with similar function or common interactions in TG metabolism. Thus, evaluation of protein families may help identify additional genes involved in lipoprotein metabolism. For instance, lipase maturation factor 2 (LMF2) is a potential TG metabolism candidate gene (147). Lmf2 is homologous to Lmf1, with 42% protein sequence identity, including the conserved C-terminal DUF1222 (147). Both Lmf1 and Lmf2 are conserved across vertebrates, although Lmf2 cannot complement or rescue Lmf1-deficient cells, suggesting it has a physiologically relevant and nonredundant role (147).

Glycosylphosphatidylinositol-anchored HDL-binding protein 1 deficiency. Gpihbp1 was identified through a large scale murine cDNA library screen (151). Gpihbp1 knockout mice have chylomicronemia (152). Gpihbp1 deficiency results in progressive onset of HTG from decreased ability of LPL to efficiently metabolize CMs (152). Gpihbp1 is expressed primarily in capillaries of heart, adipose tissue, and skeletal muscle, where it localizes to the luminal surface of endothelial cells (151, 152). Gpihbp1 probably directs fat toward energy sinks, such as the heart and adipose tissue (153). Gpihbp1 is regulated by PPARY, with increased expression during fasting and return to baseline after refeeding (154). Gpihbp1 contains an N-terminal acidic domain and a lymphocyte antigen 6 (Ly6) domain. The 25-residue acidic domain appears to bind for heparin binding substrates, including LPL, apoA-V and CMs (155). The ~80-residue Ly6 domain is required for N-linked glycosylation, forming a three-fingered structural motif (156), essential for translocation to the endothelial cell surface (157, 158).
Defining the role of the human ortholog GPIHBP1 in TG metabolism was inspired by its role in murine TG metabolism. For instance, the coding region of GPIHBP1 from 60 HTG patients without known disease-causing mutations was sequenced to find potential GPIHBP1 mutations (159). In a patient whose plasma TG was \( \sim 38 \) mmol·l\(^{-1} \), a homozygous glutamine-to-proline substitution at amino acid residue 115 (Q115P) was found; this mutation was within the Ly6 domain of GPIHBP1. Follow-up in vitro analyses demonstrated that the GPIHBP1 mutant was properly shuttled to the endothelial cell surface but was deficient in binding both LPL and CMs. Additional GPIHBP1 mutations were found in three Swedish siblings whose plasma TG concentrations were between 18 and 27 mmol·l\(^{-1} \); each sibling was a compound heterozygote for a cysteine-to-serine mutation at amino acid residue 65 (C65S) and a cysteine-to-glycine mutation at amino acid residue 68 (C68G) (160). Another patient was homozygous for a cysteine-to-tyrosine mutation also at residue 65 (C65Y) (161). These mutations all replace cysteine residues required for disulfide bonding in the GPIHBP1 Ly6 domain. Functional analyses in vitro showed that each mutant protein could reach the cell surface but that each was defective in LPL and CM binding (160, 161). Finally, screening of 160 patients with severe HTG identified a homozygous glycine-to-arginine mutation at GPIHBP1 residue 56 (G56R) in two siblings, localized to a linker domain between the N-terminal acidic domain and the Ly6 domain (162). This mutation cosegregated with the clinical and biochemical phenotype (162): the proband had an average lifetime plasma TG concentration of \( \sim 40 \) mmol·l\(^{-1} \), with recurrent pancreatitis beginning at 22 years of age, while her older brother had a similar medical history and lipid profile. Furthermore, several heterozygote relatives had less severe HTG, while relatives who were homozygous for the wild-type gene had normal lipid profiles. Analysis of this variant in vitro demonstrated that GPIHBP1-56R reached the cell surface and bound all ligands normally (163), suggesting that some other function of GPIHBP1 was affected. Cumulatively, evidence from mouse models and rare human mutations supports a role for GPIHBP1 in TG metabolism (Table 3).

**FUTURE DIRECTIONS**

The legacy of GWAS will be gauged by the number of genetic loci that help to elucidate and functionally annotate disease-associated metabolic pathways (Fig. 3). However, the greatly expanded number of TG-associated genes may challenge our ability to rapidly study their functions. Taking full advantage of the new opportunities that have arisen through GWAS discoveries will require analogous technological advances allowing fine mapping of association peaks, deep resequencing for functional variants, and finally high-throughput, robust functional validation at all stages: in vitro, in vivo, and ultimately in clinical trials. In the interim, TG-associated loci may prove useful for personalized health management strategies.

**Genetic risk prediction of HTG**

Early identification of subjects at risk for developing HTG could prompt early lifestyle modification or evidence-based pharmacological intervention to reduce risk of clinical end points. We have shown that there is a significant difference in mean score of risk alleles in HTG patients compared with those in population-based controls (19) (Fig. 2). TG risk scoring could help identify subjects at risk for developing HTG and who might benefit from early intervention. However, while few HTG patients have very low risk allele scores and few normal individuals have very high risk allele scores, there is substantial overlap toward the centers of the distributions of risk allele scores. Thus, the risk allele scores tend to discriminate primarily between normal and HTG for those with extreme values. Emerging methodologies such as whole-genome sequencing, including known common functional variants and rare deleterious or protective alleles, may help produce more discriminating TG risk scores.

**Genetic risk prediction of coronary artery disease**

The GLGC studied coronary artery disease (CAD) risk in \( \sim 25,000 \) cases and \( \sim 66,000 \) controls (19). Six TG-associated loci from GLGC studies were associated with CAD, including APOA5, LPL, TRIB1, APOE, CILP2, NAT2, and IRS1, compared with 11 LDL-C-associated loci and 8 HDL-associated loci. Interestingly, only two TG-associated loci that were independent of TG or LDL-C were associated with CAD, specifically LPL and IRS1, both of which were also HDL-C-associated loci. IRS1 is similarly associated with increased risk of T2D, insulin resistance, and hyperinsulinemia (164). Thus, it has been hypothesized that some mechanisms which increase plasma TG concentration may also directly increase CAD risk. In support of this observation, a recent Mendelian randomization experiment with the APOA5 functional variant \( −1131T>C \) has suggested a causal relationship between increased plasma TG concentration and CAD (81).

Genetic variables effectively determine a subject’s “genetic burden” for developing an unfavorable TRL profile; this factor may improve existing risk prediction algorithms for CAD, such as the Framingham risk score (165). Genetic risk scores could be composed of risk alleles exclusively for a single trait such as plasma TG or risk alleles from multiple independent traits, such as a general lipoprotein risk score. Cumulative risk scores constructed as a weighted sum of allelic effect estimates have been associated with increased CVD risk, significantly improving risk stratification of patients by 6.1% overall and by 26% for subjects in the intermediate risk category (166). Increased lipoprotein risk scores are associated with increasing plasma lipids, including plasma TG (17, 21, 167). The modest improvement of risk prediction provided by genetics, while biologically important, may be due to small effect sizes. Many additional genetic variables, or improved composite genetic risk scores, might improve the performance of risk prediction algorithms (168).
Pharmacogenomics

Fibrates and niacin are the primary pharmacotherapies for HTG. Interindividual variations in plasma TG response have been reported to differ according to genotype of common APOA5 SNPs (169) and to the presence of rare loss-of-function variants of candidate genes in TG metabolism (53). A complete genetic profile related to TG metabolism and possibly plasma TG responsiveness will include common SNP genotypes and eventually rare variants and genomic copy number variation, from GWAS of clinical trials of TG-lowering strategies. As genetic information is included as covariates in clinical trials, an individual’s genetic score might predict both biochemical TG and CVD outcome response to drugs, diet, exercise, or other interventions.

CONCLUSION

Plasma TG concentration appears to be an independent predictor of CVD risk. A combination of genetic approaches, including study of genes that encode proteins identified from classical biochemistry, genes identified through GWAS of normal and HTG samples, and genes identified by studies of families and animal models, together accounts for a relatively large proportion of variation in plasma TG. Clinical HTG has a complex genetic basis that includes a burden of both common variants, such that each has a relatively small effect on TG levels, and rare variants, such that each has a relatively large effect on TG levels. A more complete understanding of the genes and variants that modulate plasma TG will enable development of markers for risk prediction, diagnosis, prognosis, and response to therapies and might help specify new directions for therapeutic interventions.

REFERENCES


