Expression Quantitative Trait Loci Acting Across Multiple Tissues Are Enriched in Inherited Risk for Coronary Artery Disease

Hassan Foroughi Asl, Msc; Husain A. Talukdar, Msc; Alida S.D. Kindt, PhD; Rajeev K. Jain, PhD; Raili Ermel, MD; Arno Ruusalepp, MD, PhD; Khanh-Dung H. Nguyen, PhD; Radu Dobrin, PhD; Dermot F. Reilly, PhD; Heribert Schunkert, MD, PhD; Nilesh J. Samani, MD, PhD; Ingrid Braenne, PhD; Jeanette Erdmann, MD, PhD; Olle Melander, MD, PhD; Jianlong Qi, PhD; Torbjörn Ivert, MD, PhD; Josefin Skogsberg, PhD; Eric E. Schadt, PhD; Tom Michoel, PhD*; Johan L.M. Björkegren, MD, PhD*; CARDIoGRAM Consortium

Background—Despite recent discoveries of new genetic risk factors, the majority of risk for coronary artery disease (CAD) remains elusive. As the most proximal sensor of DNA variation, RNA abundance can help identify subpopulations of genetic variants active in and across tissues mediating CAD risk through gene expression.

Methods and Results—By generating new genomic data on DNA and RNA samples from the Stockholm Atherosclerosis Gene Expression (STAGE) study, 8156 *cis*-acting expression quantitative trait loci (eQTLs) for 6450 genes across 7 CAD-relevant tissues were detected. The inherited risk enrichments of tissue-defined sets of these eQTLs were assessed using 2 independent genome-wide association data sets. eQTLs acting across increasing numbers of tissues were found increasingly enriched for CAD risk and resided at regulatory hot spots. The risk enrichment of 42 eQTLs acting across 5 to 6 tissues was particularly high (≤7.3-fold) and confirmed in the combined genome-wide association data from Coronary Artery Disease Genome Wide Replication And Meta-Analysis Consortium. Sixteen of the 42 eQTLs associated with 19 master regulatory genes and 29 downstream gene sets (n>30) were further risk enriched comparable to that of the 153 genome-wide association risk single-nucleotide polymorphisms established for CAD (8.4-fold versus 10-fold). Three gene sets, governed by the master regulators *FLYWCH1*, *PSORSIC3*, and *G3BP1*, segregated the STAGE patients according to extent of CAD, and small interfering RNA targeting of these master regulators affected cholesterol-ester accumulation in foam cells of the THP1 monocytic cell line.

Conclusions—eQTLs acting across multiple tissues are significant carriers of inherited risk for CAD. *FLYWCH1*, *PSORSIC3*, and *G3BP1* are novel master regulatory genes in CAD that may be suitable targets. **(*Circ Cardiovasc Genet*. 2015;8:305-315. DOI: 10.1161/CIRCGENETICS.114.000640.)**

Key Words: coronary artery disease ■ expression quantitative trait locus ■ genetics

Genomic data sets are changing how scientists approach research of common complex diseases, such as coronary artery disease (CAD).[1] Up to 153 common single-nucleotide polymorphisms (SNPs) associated ($P<10^{-5}$) with CAD and myocardial infarction (MI) have been identified in genome-wide association (GWA) studies.[2,3] Forty-six of these SNPs were found genome-wide significant ($P=5\times10^{-8}$) also in meta-analyses of GWA studies.[2,4,5] An important task is to reveal the biological mechanisms by which these SNPs exert risk for CAD and MI. An even larger task is to discover additional CAD risk variants because those found to date explain only ≈10% of the inherited risk variation in the population, whereas the remaining ≈90% of CAD risk heritability reside remains elusive.

Clinical Perspective on p 315

Some of the missing heritability will be explained by rare disease variants and epigenetic inheritance (possibly not mediated by DNA). Another possibility is that parts of the missing heritability are explained by many disease-associated DNA variants acting together in and across tissues through common biological processes, networks, or individual pathways.[6] In traditional analysis of GWA data sets, where DNA variants are considered one by one, such variants may not reach genome-wide significance ($<10^{-5}$) and are therefore typically ignored. However, if several of these subsignificant variants ($P=10^{-2}$ to 10^{-5}) are considered together, they may constitute a substantial contribution to CAD heritability.[1,7]

Received April 28, 2014; accepted December 16, 2014.

*Drs Michoel and Björkegren contributed equally to this work and share last authorship.

The members of the CARDIoGRAM Consortium are listed in the data supplement for this article.

The Data Supplement is available at http://circgenetics.ahajournals.org/lookup/suppl/doi:10.1161/CIRCGENETICS.114.000640/-/DC1.

Correspondence to Johan L.M. Björkegren, MD, PhD, Cardiovascular Genomics Group, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden. E-mail johan.bjorkegren@ki.se

Disease-associated DNA variants may modify CAD risk directly by acting on the arterial wall during coronary artery lesion formation (primarily involving leukocytes and endothelial and smooth muscle cells). However, because CAD is a systemic disease, risk loci may modify risk indirectly in metabolic tissues.[8] As the most proximal sensor of DNA variation, RNA abundance at any stage of disease reflects interactions between the genotype and the microenvironment within one or several tissues. Accordingly, the risk contributions of expression quantitative trait loci (eQTLs) proximal to disease phenotypes can be identified through parallel assessments of allelic distributions of SNPs, tissue expression profiles across disease-relevant tissues, and GWA data sets.[1,9]

In this study, we generated new genotype and global gene expression data on DNA and RNA samples from CAD patients in the Stockholm Atherosclerosis Gene Expression (STAGE) study[10] to identify eQTLs in and across 7 tissues. To evaluate the importance of different subsets of the eQTLs as carriers of CAD risk, we primarily used 2 independent GWA data sets of CAD, one from the Myocardial Infarction Genetics Consortium (MIGen)[11] and one from the Wellcome Trust Case Control Cohort (WTCCC).[12]

Methods

The STAGE Study

In the STAGE study, 7 vascular and metabolic tissues of well-characterized CAD patients were sampled during coronary artery bypass grafting as described.[10] The Ethical committee of the Karolinska Hospital approved the study, and patients gave written consent (Dnr 004-02). Patients were included if they were eligible for coronary artery bypass grafting and had no other severe systemic diseases (eg, widespread cancer and active systemic inflammatory disease).

Genotyping and Global Gene Expression Analyses in the STAGE Study

STAGE tissue samples were obtained from atherosclerotic arterial wall (AAW), internal mammary artery (IMA), liver, skeletal muscle (SM), subcutaneous fat (SF), and visceral fat (VF)[10] and stored at –80°C in RNAlater (Ambion). Preoperatively, patients were examined and fasting whole blood (WB) was obtained for DNA and RNA isolations and biochemistry. DNA and RNA were isolated with Qiagen Mini/Midi-Kits (Ambion). DNA from 109 patients (Table 1) with sufficient quantities and qualities (≥1 μg, 1.7>260/280>1.9 with Nanodrop, Agilent) was genotyped with GenomeWideSNP_6 arrays (Affymetrix). Allele frequencies for 909 622 SNPs were determined with the Birdseed algorithm in Affymetrix Power Tools (v 1.14.2); 530 222 autosomal SNPs with call rates of 100% and minor allele frequency >5% and in Hardy-Weinberg equilibrium ($P>10^{-6}$; quality controlled [QC] SNP set) were used for downstream analysis. Custom-made HuRSTA-2a520709 arrays (Affymetrix) were used for gene expression profiling of total RNA samples (≥5 μg, 1.95>260/280>2.05 with Nanodrop, Agilent) from the 109 genotyped patients (AAW, n=68; IMA, n=79; liver, n=77; SM, n=78; SF, n=63; VF, n=88, and WB, n=102). Robust multiarray average was used for background subtraction, normalization, and summarizing of raw microarray data (Affymetrix Power Tools, v 1.14.2). A custom-made chip description file was used to match 381 707 probes on the array to 19 610 probe sets for unique genes (to avoid cross-hybridization between alternative transcripts) according to the hg19 human genome assembly.

eSNP Discovery

Each expression trait (probe set signal) was tested for association with the allele frequency of each genotyped SNP using the Kruskal–Wallis test and false-discovery rates (FDRs) to control family-wise error rates.[13] First, an empirical null distribution was estimated for each tissue by calculating association *P* values between SNPs and expression traits in data sets with randomized sample labels (patient identification numbers). Next, the FDR at level α was controlled by determining the *P* value threshold P_α such that the ratio of the average number of associations with *P* value $<P_\alpha$ under the null distribution to the number of associations with *P* value $<P_\alpha$ in the real data was equal to α. We found that the empirical null distribution was stable and 3 permutation runs per tissue were sufficient for convergence, indicating the distributions among the expression traits were largely exchangeable. At an FDR of 15%, the nominal genome-wide *P*-value thresholds for the Kruskal–Wallis test were 2.9×10^{-7} (AAW), 2.9×10^{-7} (IMA), 6.0×10^{-7} (liver), 3.5×10^{-7} (SM), 4.0×10^{-7} (SF), 2.3×10^{-7} (VF), and 6.0×10^{-7} (WB). *Cis*-acting expression single-nucleotide polymorphisms (eSNPs) were defined as SNPs within 1 Mb of the transcription start or end sites of the affected gene. All other eSNPs were defined as *trans*-acting. Because multiple testing corrections considering all SNPs (≈10^{10} pairwise tests) are too conservative for identifying *cis*-eSNPs (≈7×10^{6} pairwise tests), the above procedure was repeated, this time testing only SNPs within 1 Mb of the transcription

Table 1. eSNPs Identified From 7 Tissues in 109 Coronary Artery Disease Patients of the Stockholm Atherosclerosis Gene Expression Study (False-Discovery Rate=15%)

Tissue	No. of eSNPs	No. of Genes	No. of Gene-eSNP Pairs	*P*-Value Cutoff
Cis-eSNP				
AAW	3579	799	3716	2.5×10^{-4}
IMA	4979	1122	5304	2.9×10^{-4}
SF	3867	876	4024	3×10^{-4}
VF	4012	923	4324	2.4×10^{-4}
SM	5046	1095	5315	3×10^{-4}
Liver	10 180	2158	10 927	4×10^{-4}
WB	15 012	2984	16 952	6.7×10^{-4}
Total *cis*-eSNP	29 530	6450	34 611	NA
Total shared *cis*-eSNP	7429	1839	6986	NA
Total tissue-specific *cis*-eSNP	22 101	4611	27 625	NA
Trans-eSNP				
AAW	115	77	124	2.5×10^{-4}
IMA	155	130	184	2.9×10^{-4}
SF	99	50	99	3×10^{-4}
VF	112	85	124	2.4×10^{-4}
SM	134	175	223	3×10^{-4}
Liver	386	257	421	4×10^{-4}
WB	598	531	760	6.7×10^{-4}
Total *trans*-eSNP	1494	1222	1832	NA
Total shared *trans*-eSNP	44	48	35	NA
Total tissue-specific *trans*-eSNP	1450	1174	1797	NA

Cis- and *trans*-acting eSNPs identified in 7 different tissues and their corresponding number of genes and *P* value cutoff. AAW indicates atherosclerotic arterial wall; eSNP, expression single-nucleotide polymorphism; IMA, internal mammary artery; NA, not applicable; SF, subcutaneous fat; SM, skeletal muscle; VF, visceral abdominal fat; and WB, whole blood.

start or end site for each gene. At the threshold of FDR=15%, nominal *P*-value thresholds for the Kruskal–Wallis test for *cis*-acting eSNPs were 2.5×10^{-4} (AAW), 2.9×10^{-4} (IMA), 5.5×10^{-4} (liver), 3.0×10^{-4} (SM), 3.0×10^{-4} (SF), 2.4×10^{-4} (VF), and 6.7×10^{-4} (WB). This resulted in a total of 29530 *cis*-eSNPs for 6450 genes, corresponding to 8156 eQTLs (the eSNP most strongly associated with each gene), and 1494 *trans*-acting eSNPs for 1222 genes (Table 1; Methods in the Data Supplement). Of 29530 *cis*-eSNPs, 15012 eSNPs were identified in WB, which, at least in part, likely is attributed to the relatively larger sample number of this tissue (n=102). However, on the whole, variation in STAGE sample sizes (n=63–88 besides WB) seemed to have subtle, if any, effect on the number of identified eSNPs. For regional plots, LocusZoom[14] was used for selected CAD/MI risk loci, plotting for each SNP in the QC set its lowest association *P* value over all genes in all tissues and using the Europe population, hg19 genome assembly, and 1000 Genomes[15] to calculate linkage disequilibrium (LD).

Clustering Based on Tissue mRNA Profiles and eSNP Sets

We analyzed 555 mRNA profiles (considered as 19610-dimensional vectors) by multidimensional scaling using the Euclidean distance and Sammon's nonlinear mapping criterion. Tissues were also clustered according to eSNP overlap by constructing a phylogenetic tree from 29530-dimensional vectors recording the presence/absence of an eSNP in a tissue, using the neighbor-joining method with average linkage and the Jaccard distance measure.

Genic Regulatory and Evolutionary Genomic Features

Enrichments of eSNPs at genomic regions with certain features were compared to the bulk (n=963) of verified risk loci[16] identified in GWA studies[3] and present in our QC SNP set, using chromosome-bound circular permutations.[17] The null hypothesis of the permutation method is that eSNPs and non-eSNPs are equally likely to be located (ie, found) within each genomic annotation. The real data were compared against 20000 permuted genomes generated from the QC SNP set, which preserved the internal structure of the genome in terms of relative distance between SNPs and the observed clustering of genomic features while analyzing the LD structure (LD partners, $r^2>0.9$) around all SNPs (Methods in the Data Supplement).[17] The permuted genomes were used to calculate the odds ratios of risk enrichment/depletion of the real data in genic and regulatory regions, evolutionarily conserved regions, and regions with distinctly different histone modification patterns (Table I in the Data Supplement). Human umbilical vein endothelial cells were the most CAD-relevant cell type for which histone modification data were available.

CAD/MI Risk Enrichment

To examine eQTLs and tissue-defined subsets thereof for relative enrichment in CAD/MI risk,[18,19] we used the MIGen[11] and WTCCC[12] GWA data sets. The queried eQTL set (real set >10 eQTLs; Table II in the Data Supplement) was matched against SNPs of the GWA data set and expanded with 1000 Genomes[15] to include SNPs in that data set in strong LD ($r^2>0.8$) with eQTLs. The expanded set was compared with 10000 SNP sets of equal size randomly sampled from each GWA data set with matching chromosomal distribution. The fold enrichment was computed from the percentage of disease-associated SNPs ($P<0.05$) in the expanded set compared with the average percentage of disease-associated SNPs in the random sets. The enrichment *P* value was determined from the percentage of random sets with the same fold enrichment as the real set (or higher; Methods in the Data Supplement).

eQTL-Associated Gene Sets

For highly risk-enriched sets of eQTLs, many associated gene sets was identified. Such gene set was identified by first seeking genes associated in *cis* or in *trans* with each of the eQTLs of interest across the 7 tissues at a nominal *P* value for the Kruskal–Wallis test ($P<1\times10^{-3}$). Next, highly coexpressed genes active in the same tissue as eQTL-gene pair was identified forming the final gene set (Pearson correlation coefficient, >0.85). The threshold parameters were those yielding optimal gene set sizes from all combinations of association *P*-value thresholds of 5×10^{-4}, 1×10^{-3}, and 5×10^{-3} and Pearson correlation threshold of 0.8, 0.85, and 0.9. eQTLs and associated gene sets in a tissue were assessed for molecular functions in Gene Ontology categories.[20] Functional-enrichment *P* values were determined with Fisher exact test and Benjamini and Hochberg FDR corrections for multiple testing.

Small Interfering RNA Perturbation of THP-1 Macrophages Incubated With Acetylated Low-Density Lipoproteins

As previously described,[21] THP-1 monocytic cells were differentiated into macrophages using phorbol myristate acetate and thereafter transfected with small interfering RNA targeting each master regulator (one at a time) or mock-transfected (Table III in the Data Supplement). Next, the cells were incubated with acetylated low-density lipoproteins (50 μg/mL) for 48 hours allowing foam cells to form. The cholesterol-ester accumulation in the foam cells were determined by measuring and subtracting free cholesterol from total cholesterol adjusted for protein concentration (Methods in the Data Supplement).

Results

The STAGE Study and Tissue Gene Expression Patterns

The STAGE cohort consisted of typical CAD cases with well-managed hyperlipidemia and hypertension (>90% on β-blocker and plasma-lipid-lowering regimens); 20% had diabetes mellitus (Table IV in the Data Supplement). The expression patterns of the 19610 genes across 7 tissues in the STAGE study segregated the tissues into well-defined groups, reflecting unique tissue functions (Figure 1A). Gene expression patterns in AAW and nonatherosclerotic IMA were closely related but, importantly, well segregated; gene expression patterns in SF and VF were more intermixed (Figure 1A; Figures I and II in the Data Supplement).

Most *Cis*-Acting eSNPs Affect Gene Expression Only in One Tissue

The allelic distributions of 29530 SNPs within 1 Mb of the transcription start site (TSS) were associated with mRNA levels of 6450 unique genes (*cis*-eSNPs [eSNPs]; FDR 15%; Methods in the Data Supplement; Table 1; and Figure III in the Data Supplement). Of the identified 29530 eSNPs, 22101 were identified in only one tissue (tissue-specific eSNPs) and 7429 in ≥2 tissues (multitissue eSNPs). Similar proportions of tissue-specific versus shared eSNPs have been reported.[22–25] Also identified were 1494 *trans*-eSNPs (>1 Mb from the TSS), mainly in WB (38%). Within these sets, we identified 39 eSNPs matching SNPs previously associated with CAD/MI risk in GWA studies,[4] whereof 21 exact matches and 18 in strong LD ($r^2>0.8$; Table V and Figure IV in the Data Supplement). On average, the multitissue eSNPs were closer to the TSS than the tissue-specific eSNPs (Figure 1D and 1E), as reported.[22,26] Whether the multitissue eSNPs were present in 2 or more tissues did not affect their positions relative to the TSS (Figure 1D). The notion that eSNPs more distal to the TSS have tissue-specific effects on gene expression is

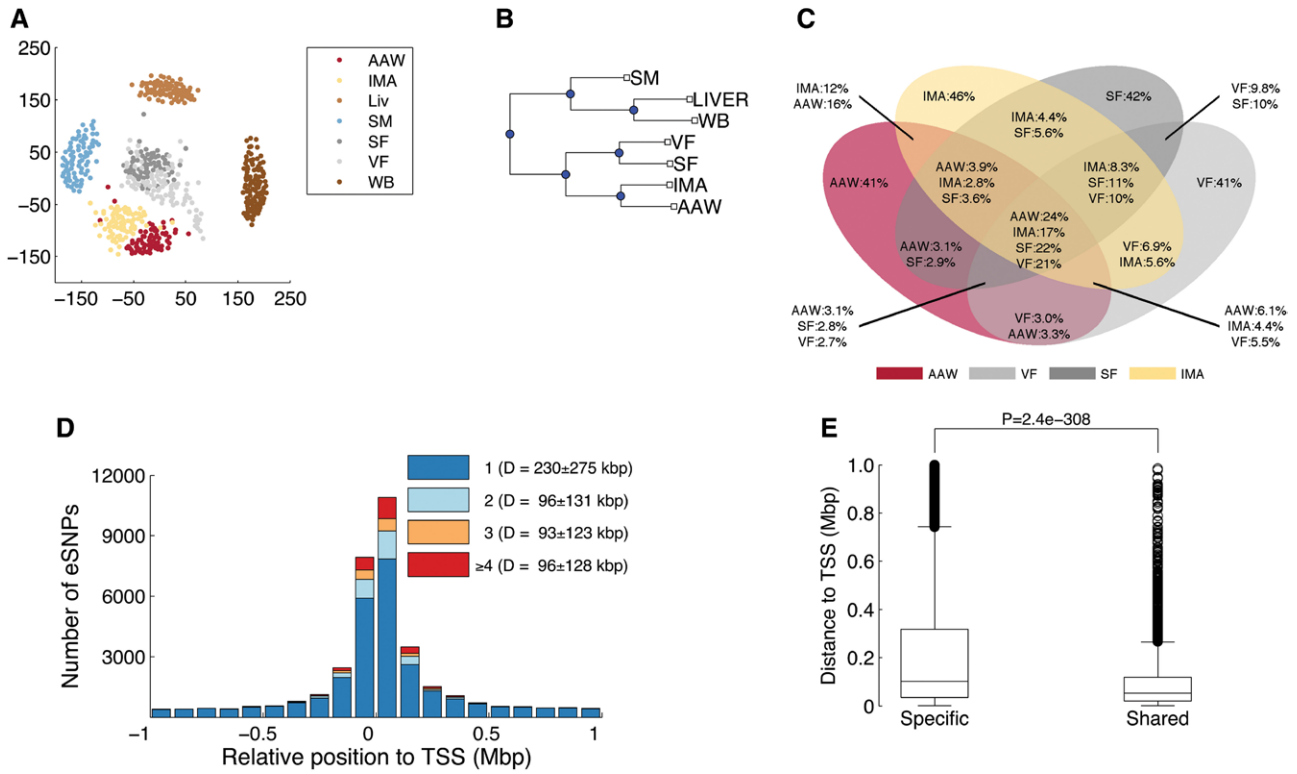

Figure 1. Characteristics of gene expression single-nucleotide polymorphisms (eSNPs) of the Stockholm Atherosclerosis Gene Expression cohort. eSNPs were identified from associations between allelic values of SNPs and gene expression traits in each tissue. **A**, Sammon plots of tissue-specific global gene expression. Each dot is a 2D representation of 19 610 gene expression signals (HuRSTA-2a520709, Affymetrix). Tissues profiled were atherosclerotic arterial wall (AAW; n=68 arrays), internal mammary artery (IMA; n=79), liver (n=77), skeletal muscle (SM; n=78), subcutaneous fat (SF; n=63), visceral fat (VF; n=88), and whole blood (WB; n=102). **B**, Tree view of hierarchically clustered eSNP sets identified in all 7 tissues. **C**, Venn diagram showing the portion of specific and multitissue eSNPs in AAW, IMA, SF, and VF. Percentages indicate fractions of expression quantitative trait loci in 1 tissue. **D**, Frequency plot of 29 530 eSNPs in relation to the transcription start site (TSS). Blue, tissue-specific eSNPs; light blue, eSNPs shared in 2 tissues; orange, eSNPs shared in 3 tissues; red, eSNPs shared in 4 or more tissues. **E**, Box plots of the distribution of distances to the TSS for all tissue-specific and multitissue eSNPs, showing that tissue-specific eSNPs lie significantly closer to the TSS than multitissue eSNPs (Wilcoxon rank-sum test). Each box indicates the median and the 25th and 75th percentiles of the data.

consistent with the fact that tissue-specific enhancer regions are commonly more distal to the TSS. Similarly, the notion that eSNPs close to the TSS have tissue-independent effects on gene expression (across 2 or more tissues) is consistent with the idea that the TSS of a gene is similar across tissues.

Shared eSNPs in Arterial Wall and Fat Deposits

eSNPs in arterial wall samples (AAW, IMA) and fat deposits (VF and SF) formed a common cluster, as did those in liver and WB; eSNPs in SM formed an isolated cluster (Figure 1B). In fact, ≤60% of the eSNPs in AAW, IMA, VF, and SF were shared (Figure 1C). Thus, in CAD, diseased and healthy arteries seem to share most regulatory SNPs with fat depots, fewer with liver and WB, and fewest with SM. Venn diagrams for eQTLs in all combinations of 4 tissues are shown in Figure V in the Data Supplement.

Multitissue eQTLs Carry More Inherited CAD Risk Than Tissue-Specific eQTLs

The strongest eSNP association of a gene is referred to as its eQTL. To assess the importance of different sets of tissue-specific and multitissue eQTLs for inherited risk of CAD/MI, we used the MIGen[11] and WTCCC[12] GWA data sets to analyze risk enrichment (Table II in the Data Supplement). According to both data sets, the entire set of *cis*-eQTLs was risk enriched by 1.15- to 1.2-fold (Table II in the Data Supplement). The set of *trans*-eQTLs was not risk enriched (not shown).

CAD is a complex disease that is particularly dependent on the molecular status of the metabolic tissues that drive atherosclerosis and thrombosis in the vascular bed. Thus, eQTLs present only in one tissue (tissue-specific) may have different roles from multitissue eQTLs, affecting gene expression across several tissues in mediating CAD risk. We examined this possibility by assessing risk enrichments of different sets of eQTLs present in single or multiple tissues (>2 tissues). According to MIGen GWA data set, only tissue-specific eQTLs in SM, SF, and WB carried CAD/MI risk (Figure 2A, red bars). However, for multitissue eQTLs, all tissues but SM and VF were risk enriched (Figure 2B, red bars). According to the WTCCC, the risk-enrichment levels for both tissue-specific and multitissue eQTLs were higher, but multitissue eQTLs were again found to be more risk enriched than tissue-specific eQTLs (Figure 2A and 2B, blue bars).

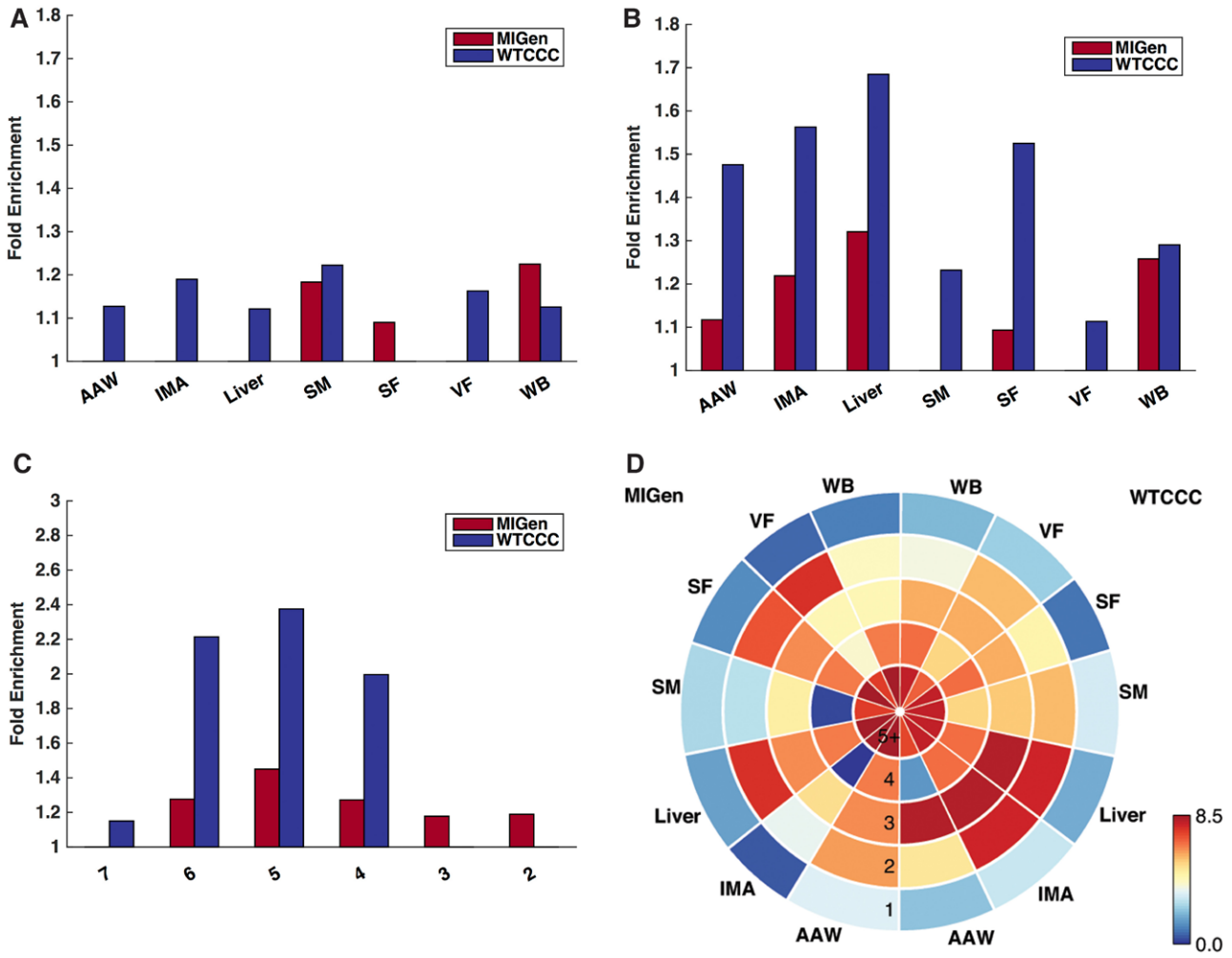

Figure 2. Inherited risk enrichment for Stockholm Atherosclerosis Gene Expression (STAGE) expression quantitative trait loci (eQTLs) according to 2 major genome-wide association (GWA) studies of coronary artery disease (CAD)/myocardial infarction (MI). Inherited CAD/MI risk enrichments of STAGE eQTLs and subsets thereof were sought as described[19] using 2 independent GWA studies: Myocardial Infarction Genetics Consortium (MIGen)[11] (red bars) and Wellcome Trust Case Control Cohort (WTCCC)[12] (blue bars). Sets of eQTLs were tested for enrichment against 10 000 random sets at a linkage disequilibrium of r^2>0.8 (see Methods in the Data Supplement). Bars indicate enrichment results that are statistically significant (P<10^{-3}). **A**, CAD/MI risk enrichments for tissue-specific eQTLs divided by tissues. **B**, CAD/MI risk enrichments for multitissue eQTLs divided by tissues. **C**, CAD/MI risk enrichments of eQTLs shared in all combinations of 2 to 7 tissues. **D**, Pie chart showing risk enrichments for eQTLs in specific tissues (outer circle, 1 pie=1 tissue) and the risk enrichments of multitissue eQTLs active in one given and a combination of 1 to 4 tissues (toward the center). Colors indicate degree of enrichment. For multitissue eQTLs, the highest enrichment for each tissue combination is indicated. AAW indicates atherosclerotic arterial wall; IMA, internal mammary artery; SF, subcutaneous fat; SM, skeletal muscle; VF, visceral fat; and WB, whole blood.

As eQTLs Are Shared Across More Tissues They Are Increasingly Risk Enriched

Because multitissue eQTLs (>2 tissues) were more risk enriched than tissue-specific eQTLs, we asked whether the CAD/MI risk of eQTLs increases with the number of tissues in which these eQTLs affect gene expression (from 2 to 7 tissues). Indeed, this proved to be the case: risk enrichment peaked in combinations of 4 to 6 tissues (1.2-fold, P<8.0×10^{-4}) according to the MIGen data set and in combinations of 6 tissues (2.7-fold, P<2.0×10^{-86}) according to the WTCCC data set (Figure 2C). Thus, eQTLs that affect gene expression across 4 to 6 tissues seem to be the most relevant as carriers of CAD risk.

When we instead only considered the highest enrichment for a combination of specific tissues (from 2 to 7), the risk-enrichment scores according to both GWA data sets generally increased from tissue-specific eQTLs (Figure 2D, outer circle, ≤1.2-fold, P<4.0×10^{-9}) to those active in combinations of 5 or more tissues (Figure 2D, inner circle, ≤8.5-fold, P<1.7×10^{-43}).

Multitissue eQTLs Are Well Conserved and Enriched at Sites of Regulatory Elements

Most verified GWA hits are enriched at sites of regulatory elements,[17,27] suggesting that disease-risk loci affect disease by modifying regulatory hot spots rather than by modifying

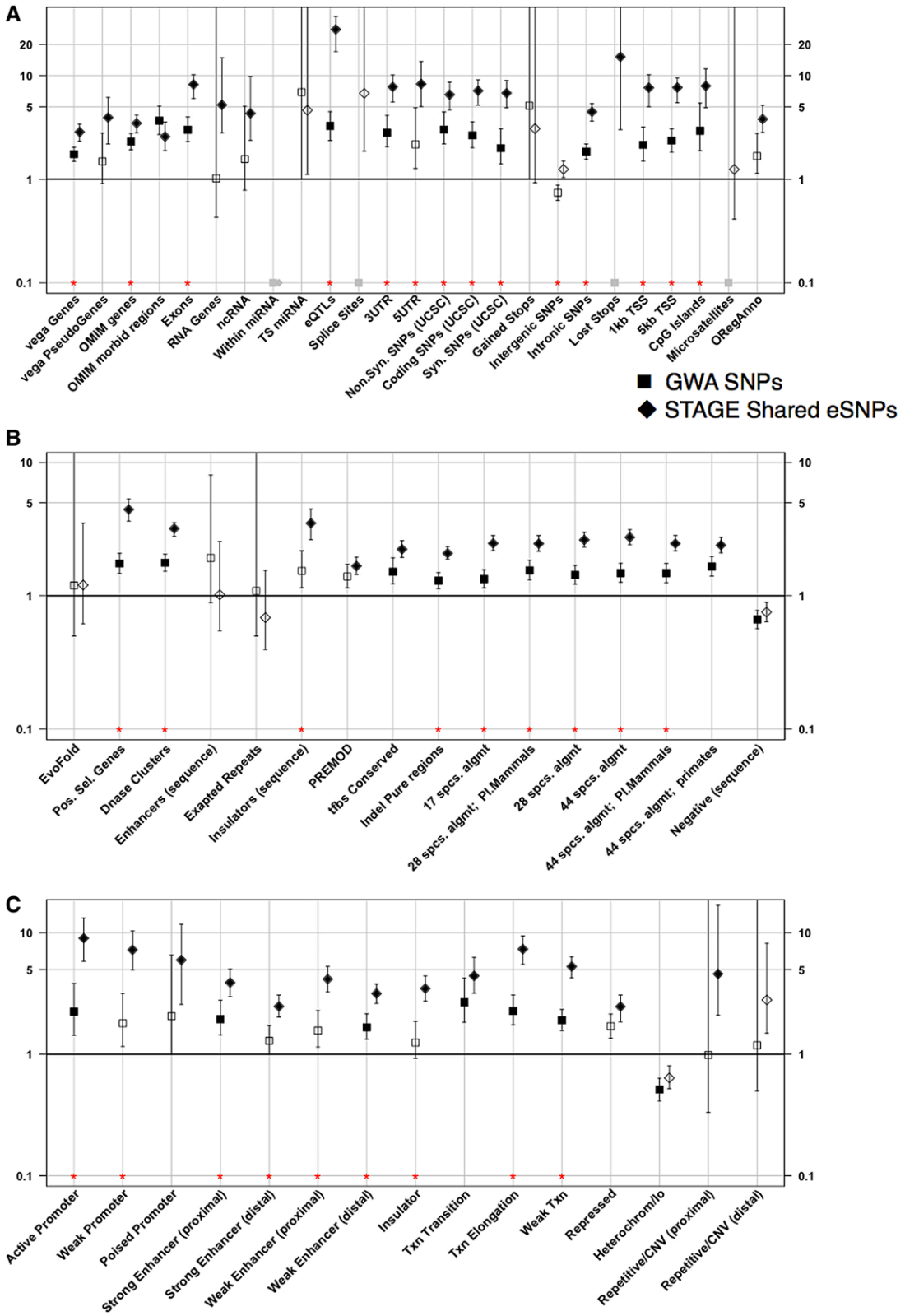

Figure 3. Gene expression single-nucleotide polymorphism (eSNP) and genome-wide association (GWA) disease locus enrichments at regulatory and evolutionarily conserved genic regions of the genome. Plots show regulatory element enrichment patterns for 29530 *cis*-eSNPs (Stockholm Atherosclerosis Gene Expression [STAGE] eSNPs) and 963 established GWA single-nucleotide polymorphisms (SNPs) compared to equally sized random eSNP groups (see Methods section) at positions in the genome with known (**A**) regulatory, (**B**) evolutionarily conserved, and (**C**) histone-modifying properties in human umbilical vein endothelial cells (HUVECs). *x*-axes show regulatory region categories (Table II in the Data Supplement); *y*-axes show fold-enrichment for each category. Bars show 95% confidence intervals. Filled symbols indicate categories with significant enrichments. Statistically significant differences between enrichments for STAGE eSNPs and GWA SNPs are indicated with red stars above each. (*Continued*).

Figure 3. (Continued) category on the *x*-axis. (A–C) Fold enrichment of STAGE eSNPs and GWA SNPs at (A) regulatory regions of the genome, (B) evolutionary conserved regions of the genome, and (C) histone modification regions of the genome in HUVECs. HUVECs were chosen as the most coronary artery disease–relevant cell type with mapped histone modification regions. CNV indicates copy-number variation; eQTL, expression quantitative trait loci; miRNA, microRNA; ncRNA, noncoding RNA; OMIM, Online Mendelian Inheritance in Man; PREMOD, Predicted Regulatory Module; TS miRNA, TargetScan miRNA; TSS, transcription start site; UCSC, University of California Santa Cruz genome browser; and UTR, untranslated region.

the expression of specific genes. Conceivably, eQTLs are enriched at regulatory sites. To investigate this possibility, we compared the regulatory element enrichment patterns of all the identified eSNPs (n=29 530) and the tissue-specific (n=22 101) and multitissue (n=7429) eSNPs with the regulatory element enrichment patterns of 963 established GWA loci[16] (GWA-SNPs) in a range of regulatory element subcategories (Table I in the Data Supplement). Strikingly and as opposed to tissue-specific eSNPs (Figure VI in the Data Supplement), the regulatory element enrichment patterns for the multitissue eSNPs were significantly enriched, and more so than those for GWA-SNPs, regardless of the subcategory (genic and regulatory regions, evolutionary and conserved regions, or regions of histone modification in human umbilical vein endothelial cells; Figure 3A–3C). This finding strongly suggests that in contrast to tissue-specific eQTLs, and to a greater extent than GWA-SNPs, multitissue eSNPs/eQTLs affect gene expression across tissues mainly by modifying DNA regulatory elements at well-conserved genomic sites. This finding is consistent with their tissue-independent effects on gene expression (ie, being multitissue eQTLs) and possibly with their greater importance as carriers of CAD risk.

Forty-Two Multitissue eQTLs Found to Be Important Carriers of CAD/MI Risk

The finding that eQTLs shared across combinations of tissues (>4) were more enriched with inherited CAD/MI risk than tissue-specific eQTLs and those shared across fewer tissues (<4) suggests that there might be specific subsets of highly shared eQTLs that are particular important carriers of CAD risk. We therefore examined eQTLs active across specific combinations of the 7 tissues. Twelve eQTLs in one combination of 5 tissues (AAW, liver SF, VF, and WB) and 30 eQTLs in a combination of 6 tissues (IMA, liver, SM, SF, VF, and WB) were found highly enriched in CAD/MI risk according to the MIGen (8.5-fold, $P=1.6\times10^{-43}$) and WTCCC (3.6-fold, $P=2.2\times10^{-30}$) GWA data sets (Tables VI and VII in the Data Supplement). Combined, these 42 eQTLs were found to be risk enriched by ≤7.3-fold (7.3-fold, $P=7.7\times10^{-20}$ in WTCCC and 2.3-fold, $P=3.1\times10^{-6}$ in MIGen). To validate the 42 eQTLs in a broader set of GWA studies of CAD, we interrogated the GWA meta-analysis of CAD undertaken by the Coronary Artery Disease Genome Wide Replication And Meta-analysis (CARDIoGRAM) consortium.[4] The CARDIoGRAM data set consists of ≈22 000 CAD cases and ≈60 000 controls and includes the WTCCC and MIGen data sets, which contribute 14% of the cases and 7% of the controls.[4] Using this GWA data set, we found that the 42 eQTLs were risk enriched by 4.2-fold ($P<1.1\times10^{-54}$). The CAD associations for the individual SNPs in this set are shown in Table VIII in the Data Supplement.

Gene Sets Downstream of Multitissue eQTLs

To assess downstream effects on gene expression governed by the multitissue eQTLs, we first identified genes associated in *cis* or in *trans* with each of the 42 eQTLs (Figure VII in the Data Supplement; $P<10^{-3}$ by Kruskal–Wallis test) and then identified highly coexpressed genes downstream of the master regulators (absolute Pearson correlation coefficient, >0.85). In this fashion, 29 gene sets (with ≥30 genes in gene set in different tissues) for 16 of the 42 eQTLs, termed as master regulator eQTLs, involving 19 *cis*-regulated genes, termed as master regulator genes, were identified. According to the CARDIoGRAM GWA data set, these 16 eQTLs were further enriched with CAD risk (8.4-fold, $P=5.5\times10^{-51}$). To put the fold enrichment of risk in perspective, we tested the risk enrichment for the 153 CAD risk SNPs established by GWAS[2] and found it to be 10-fold ($P<10^{-99}$) according to the CARDIoGRAM GWA data set.

To further evaluate the role of these 16 master regulatory eQTLs, we first used Gene Ontology[20] to assess the functional characteristics of the 29 gene sets. Nineteen gene sets returned significant functional enrichments that without exception revealed biological processes relevant to CAD: T-cell differentiation, regulation of cholesterol efflux, oxidation–reduction process, steroid metabolic process, blood vessel development, respiratory electron transport chain, and cell cycle (Table 2; Tables VI, VII, and IX in the Data Supplement).

We also examined whether clustering patients according to expression levels within each of the 29 gene sets would result in patient subgroups associated with CAD traits or clinical phenotypes. Remarkably, as many as 27 of the 29 gene sets clustered the STAGE patients according to ≥1 CAD trait or clinical phenotype (Table 2; Table IX in the Data Supplement). Heat maps of patient clusters within each of the 29 gene sets are shown in Figure VII in the Data Supplement.

Master Regulatory eQTL Genes

In view of the master regulatory role of the 16 eQTLs in CAD, it is interesting to know more about their *cis*-regulated genes (master regulators). Three of the 4 master regulators with gene sets found to be associated with the primary CAD trait—the extent of coronary atherosclerosis assessed from preoperative angiograms in the STAGE patients (CAD score in Table 2)—were identified in abdominal fat (VF): *G3BP1*, *PSORSIC3*, and *FLYWCH1*. The gene sets of *G3BP1* (a DNA-unwinding enzyme) and *PSORSIC3* (a psoriasis-susceptibility gene) were highly enriched in cell-cycle genes ($P<1\times10^{-21}$ and $P<1\times10^{-48}$). The gene set of *FLYWCH1* (a zinc finger–containing protein) was instead enriched in oxidation–reduction processes ($P<1\times10^{-11}$).

A blood (WB) gene set (n=53) with the master regulator *SNAPIN* (SNAP-associated protein) was also found to be

Table 2. Master Regulators and CAD Traits/Clinical Phenotype of the Stockholm Atherosclerosis Gene Expression Patients

CAD Traits	Master Regulatory eQTL	Master Regulatory Gene	eQTL-Gene *P* Value	eQTL-Gene FDR	Gene Set (n)	Tissue	Functional Association (*P* Value)	Clinical Phenotype (*P* Value)
Preoperative coronary angiography								
CAD score*	rs892006grp†	*G3BP1*	2.29×10^{-6}	0.05	112	VF	Cell cycle (1×10^{-21})	NA
	rs9930148†	*FLYWCH1*	2.3×10^{-4}	0.15	239	VF	Oxidation-reduction process (3×10^{-11})	NA
	rs6908994‡	*PSORS1C3*	1.78×10^{-10}	0.05	473	VF	Cell cycle (3.92×10^{-48})	NA
	rs12564925‡	*SNAPIN*	2.25×10^{-5}	0.05	53	WB	T cell differentiation (1.09×10^{-4})	Inherited diabetes mellitus (0.004), years of smoking (0.015)
Blood chemistry								
LDL cholesterol	rs7345‡	*SNAPIN*	2.5×10^{-6}	0.05	49§	VF	NA	NA
	rs12564925‡		1.93×10^{-6}					
C-reactive protein	rs7345‡	*SNAPIN*	5.55×10^{-4}	0.15	117	IMA	Steroid metabolic process (4.42×10^{-21})	Allergy (0.001)
	rs12564925‡	*SNAPIN*	2.29×10^{-4}	0.15	126	IMA	Steroid metabolic process (3.892×10^{-20})	Allergy (0.001)

eQTLs with corresponding genes defined as master regulators when associated with >30 genes (gene sets). CAD traits and clinical phenotypes are compared in subgroups of STAGE patients ($P<0.05$) defined from clustering by gene set expression values. CAD indicates coronary artery disease; eQTL, expression quantitative trait loci; FDR, false-discovery rate; IMA, internal mammary artery; LDL, low-density lipoprotein; NA, not applicable; STAGE, Stockholm Atherosclerosis Gene Expression; VF, visceral fat; and WB, whole blood.

*Assessed in preoperative coronary angiograms of the STAGE patients.

†eQTL group I (shown in Tables VI and VII in the Data Supplement).

‡eQTL group II (shown in Tables VI and VII in the Data Supplement).

§Weaker associations were also found with thyroid stimulating hormone.

associated with CAD score and enriched in genes active in T-cell differentiation ($P<1\times10^{-4}$; Table 2). This *SNAPIN* gene set was also associated with inherited diabetes mellitus, years of smoking, and physical activity in STAGE patients (Table 2; Table IX in the Data Supplement). Another *SNAPIN* gene set, identified in abdominal fat (VF), was found associated with low-density lipoprotein cholesterol levels (Table 2). For yet 2 other *SNAPIN*-associated gene sets (both identified in the arterial wall [IMA]) that were highly enriched in steroid metabolic processes ($P<1\times10^{-20}$), one set was found associated with C-reactive protein levels of the STAGE patients and the other with the diagnosis of allergy (Table 2). Master regulator eQTL genes related to other phenotypes of the STAGE patients, including plasma levels of lipids, glucose, and C-reactive protein, are presented in Table IX in the Data Supplement.

Master Regulators Associated With Extent of Coronary Atherosclerosis Validated by Small Interfering RNA in Foam Cells

The master regulatory genes *G3BP1*, *PSORSIC3*, *FLYWCH1* (in VF), and *SNAPIN* (in WB) were all associated with CAD score in STAGE patients, suggesting these genes are important regulators of the formation of atherosclerosis. A key event of atherosclerosis is the formation of foam cells. To validate these genes, we examined the effect of small interfering RNA targeting of *G3BP1*, *PSORSIC3*, *FLYWCH1*, and *SNAPIN* on cholesterol-ester accumulation in THP1-foam cells. Besides *SNAPIN*, small interfering RNA targeting of *G3BP1*, *PSORSIC3*, and *FLYWCH1* had pronounced effects on cholesterol-ester accumulation in THP-foam cells (Table 3; Table III in the Data Supplement).

Discussion

In this study, we identified eQTLs from 7 CAD-relevant tissues from the same patients and used well-established GWA data sets to assess their importance for inherited risk of CAD. Our analysis confirmed that SNPs regulating gene expression (ie, eSNPs/eQTLs) are more likely to carry disease risk than nonregulatory SNPs.[18,22–25] Importantly, we also found that eQTLs affecting gene expression across multiple tissues carried substantially more inherited CAD risk than eQTLs affecting gene expression in isolated tissues. Unlike tissue-specific eQTLs, these multitissue eQTLs affected gene expression largely by modifying well-conserved regulatory elements of the genome rather than by modifying the expression of specific genes. In fact, the overall CAD-risk enrichment of multitissue eQTLs at these regulatory sites was significantly greater than that of GWA hits reported by the Encyclopedia of DNA Elements (ENCODE) Consortium and others.[17] Thus, multitissue eQTLs are relatively well conserved and important in modulating pathophysiological responses to gene expression.

We also identified 42 multitissue eQTLs highly enriched with inherited CAD/MI risk, as shown by our analysis of 2 isolated GWA data sets and the combined data set of GWA studies of CAD from the CARDIoGRAM Consortium.[4] Sixteen of these 42 eQTLs were further enriched with inherited risk and linked to 19 master regulatory genes with

Table 3. Effects of Small Interfering RNA Inhibition on CE Accumulation in a THP1-Foam Cell Model

Gene Symbol	Control (Relative CE Levels)	siRNA Knock (Relative CE Levels)	CE Content (% Relative Control)	*P* Value	siRNA Inhibition, %
G3BP1	100±13*	138±15*	+38	2.3×10^{-6}	84
FLYWCH1	100±18*	144±30*	+44	1.7×10^{-5}	62
PSORS1C3	100±18*	130±38*	+30	6.7×10^{-3}	64
SNAPIN	100±23*	104±29*	+4	0.738	74

CE indicates cholesterol ester; and siRNA, small interfering RNA.
*Values are presented as mean±SD.

29 downstream gene sets in specific tissues. Analysis of the functional and clinical associations of these gene sets further validated the roles of the 16 multitissue eQTLs and their 19 regulatory genes in CAD. Specifically, 19 of 29 gene sets were enriched with biological processes highly relevant to CAD, and 27 sets clustered STAGE patients by several key CAD traits and clinical phenotypes, including 4 master regulatory genes linked to the extent of coronary atherosclerosis; 3 of these genes also affected the extent of foam cell formation in vitro.

By analyzing 7 tissues from the same patients, we identified 2 sets of eQTLs affecting gene expression in combinations of 5 and 6 tissues that were highly enriched in CAD/MI risk (≤8.5-fold, $P=1.6\times10^{-43}$). Previous studies of the risk enrichment of eQTLs have examined at most 4 tissues, often fewer.[18,19,22–25] The marked risk enrichment of eQTLs affecting gene expression across >4 tissues was unexpected and may reflect the complexity of CAD. Although atherosclerosis (the key process in CAD) occurs in the arterial wall, its progression and composition (which determine clinical outcomes) are also determined by blood constituents (eg, plasma cholesterol and glucose) that in turn are dependent on the molecular status in several organs particularly those involved in metabolism. Hence, eQTLs affecting specific genes in isolated tissues may indeed be less important for the development and clinical outcome of CAD than eQTLs affecting high hierarchical regulatory elements that influence gene expression across multiple vascular and metabolic tissues.

Sixteen of 42 multitissue eQTLs were linked to 19 genes with master regulatory roles in CAD (Tables 2 and 3). Unexpectedly, the downstream effects of these master regulatory eQTLs were tissue-specific and seemingly dependent on contexts in some but not all of the tissues in which the eQTLs were active. Such contexts appeared largely present in VF, which contained many gene sets, but were also in the arterial wall (AAW and IMA) and blood (WB), where most of the gene sets were identified (Tables 2 and 3).

Besides the 4 master regulatory genes associated with CAD (Table 2, *G3BP1*, *FLYWCH1*, *PSORS1C3*, and *SNAPIN*), we identified additional 15 that were associated with other CAD phenotypes (eg, plasma levels of low-density lipoprotein; Table IX in the Data Supplement). Of 19 master regulators in total, 13 (*ABHD16A*, *SNAPIN*, *S100A1*, *S100A13*, *PRRC2A*, *PHOSPHO2*, *PBXIP1*, *HSPC157*, *GPANK1*, *G3BP1*, *FLYWCH1*, *CKAP2*, and *C15orf57*) had not previously been associated with CAD or any of its subphenotypes, as judged from the Genetic Association Database.[28] *S100A1*, however, has been associated with breast cancer. The remaining 6 master regulators had been associated with CAD-related subphenotypes: diabetes mellitus type 2 (*SIGLEC12*, *VARS*, and *HLA-C*), left ventricular failure (*PSORS1C3*), plasma triglycerides (*PSORS1C3*), rheumatic heart disease (*HLA-C*), coronary aneurysm (*HLA-C*), heart rate (*LGALS4*), and blood pressure and arterial hypertension (*CTNS*; Table X in the Data Supplement). Interestingly, genetic variation in *SIGLEC12* (a master regulator associated with hypertension in the STAGE patients; Table IX in the Data Supplement) and treatment of hypertension have been linked to cardiovascular outcomes.[29]

Although our finding that multitissue eQTLs carry more inherited risk than tissue-specific eQTLs may well be explained by complex interactions between metabolic and vascular tissue in CAD development, we considered other explanations for this observation. For example, could increasingly shared (ie, multitissue) eQTLs be more strongly associated with their expression traits (have higher eQTL *P* values than tissue-specific eQTLs) and thus be a more robust subpopulation of eQTLs that therefore are more likely to be risk-enriched? This was not found to be the case: first, there was no correlation between the strength of association between SNPs and their expression traits (ie, eSNP *P* values according to STAGE data) and their level of risk enrichments (CAD association according to GWA data sets; Figure VIII in the Data Supplement). In fact, the set of eSNPs with the strongest associations with their expression traits (ie, strongest eSNP *P* values according to STAGE data) were not found to be more risk enriched than the set of SNPs with weakest associations (Table XI in the Data Supplement). Second, nor was the higher risk enrichment of multitissue eQTLs lost or substantially altered in any combination of eQTL discovery FDR threshold (10%, 15%, and 20%) and major allele frequency cutoff (0.05, 0.10, and 0.15; Figure IX in the Data Supplement).

In sum, by examining eQTLs across 7 tissues, we discovered that eQTLs active across >4 tissues are important carriers of CAD/MI risk. We also identified 16 eQTLs governing the expression of 19 master regulatory genes linked to CAD. These master regulators and particularly those related to the extent of coronary atherosclerosis also affecting cholesterol-ester accumulation in foam cells merit further attention as putative CAD targets.

Acknowledgments

We thank Stephen Ordway for editorial assistance.

Disclosures

This work was supported by Precocious Coronary Artery Disease in the 6th European Union-framework program (LSHM-CT-2007–037273), the Swedish Heart-Lung Foundation (Dr Björkegren), the Swedish Research Council (Drs Björkegren and Skogsberg), the King Gustaf V and Queen Victoria's Foundation of Freemasons (Drs Björkegren and Skogsberg), and the Swedish Society of Medicine (Drs Björkegren and Skogsberg). This work was also supported by grant from University of Tartu (SP1GVARENG, Dr Björkegren), the Estonian Research Council (Sihtasutus Eesti Teadusagentuur [ETAg], Drs Ruusalepp and Björkegren), the Freiburg Institute for Advanced Studies and Roslin Institute Strategic Grant Funding from the Biotechnology and Biological Sciences Research Council (Dr Michoel), Astra Zeneca Translational Science Centre—Karolinska Institutet (joint research program in translational science, Dr Björkegren), and Clinical Gene Networks AB (CGN). Dr Samani holds a Chair funded by the British Heart Foundation and is supported by the National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit. Drs Samani, Björkegren, Schunkert, Erdmann, Braenne, and Schadt are members of the Leducq Consortium CAD Genomics, and this work was undertaken as part of the Consortium. We are grateful to the Coronary Artery Disease Genome Wide Replication and Meta-Analysis (CARDIoGRAM) consortium for allowing access to their coronary artery disease genome-wide association meta-analysis data. Dr Björkegren is founder, main shareholder, and chairman of the board for CGN. Drs Ruusalepp and Schadt are board members and shareholders. Dr Michoel is shareholder. CGN has an invested interest in microarray data generated from the Stockholm Atherosclerosis Gene Expression cohort. The other authors report no conflicts.

Appendix

From the Cardiovascular Genomics Group, Division of Vascular Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden (H.F.A., H.A.T., J.S., J.L.M.B.); MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom (A.S.D.K.); Division of Genetics and Genomics, The Roslin Institute, University of Edinburgh, Edinburgh, United Kingdom (T.M.); Department of Pathological Anatomy and Forensic Medicine, University of Tartu, Tartu, Estonia (R.K.J., R.E., A.R., J.L.M.B.); Department of Cardiac Surgery, Tartu University Hospital, Tartu, Estonia (R.E.); Clinical Gene Networks AB, Stockholm, Sweden (A.R., J.L.M.B); Department of Genetics and Pharmacogenomics, Early Development and Discovery Sciences, Merck Research Laboratories, Boston, MA (K.-D.H.N., D.F.R.); Department of Genetics and Genomic Sciences, Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY (R.D., E.E.S., J.L.M.B.); Deutsches Herzzentrum München, Technische Universität München and German Centre for Cardiovascular Research (DZHK), Munich Heart Alliance, Munich, Germany (H.S.); Department of Cardiovascular Sciences, University of Leicester and National Institute for Health Research Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, United Kingdom (N.J.S.); Institut für Integrative und Experimentelle Genomik, Universität zu Lübeck and 2DZHK (German Centre for Cardiovascular Research), Partner Site Hamburg/Lübeck/Kiel, Lübeck, Germany (I.B., J.E.); Department of Clinical Sciences, Hypertension and Cardiovascular Disease, Clinical Research Center, Skåne University Hospital, Malmö, Sweden (O.M.); Freiburg Institute for Advanced Studies, University of Freiburg, Freiburg, Germany (J.Q., T.M.); and Department of Thoracic Surgery, Karolinska University Hospital, Stockholm, Sweden (T.I.).

References

1. Schadt EE, Björkegren JL. NEW: network-enabled wisdom in biology, medicine, and health care. *Sci Transl Med*. 2012;4:115rv1. doi: 10.1126/scitranslmed.3002132.
2. CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45:25–33.
3. Hindorff LA, MacArthur J, Morales J, Junkins HA, Hall PN, Klemm AK, et al. A catalog of published genome-wide association studies. Available at: http://Www.Genome.Gov/gwastudies. Accessed November 21, 2013.
4. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al; Cardiogenics; CARDIoGRAM Consortium. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*. 2011;43:333–338. doi: 10.1038/ng.784.
5. Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. *Hum Mol Genet*. 2011;20:R198–R205. doi: 10.1093/hmg/ddr384.
6. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, et al. Variations in DNA elucidate molecular networks that cause disease. *Nature*. 2008;452:429–435. doi: 10.1038/nature06757.
7. Orozco LD, Bennett BJ, Farber CR, Ghazalpour A, Pan C, Che N, et al. Unraveling inflammatory responses using systems genetics and gene-environment interactions in macrophages. *Cell*. 2012;151:658–670. doi: 10.1016/j.cell.2012.08.043.
8. Tegnér J, Skogsberg J, Björkegren J. Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Multi-organ whole-genome measurements and reverse engineering to uncover gene networks underlying complex traits. *J Lipid Res*. 2007;48:267–277. doi: 10.1194/jlr.R600030-JLR200.
9. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature*. 2009;461:218–223. doi: 10.1038/nature08454.
10. Hägg S, Skogsberg J, Lundström J, Noori P, Nilsson R, Zhong H, et al. Multi-organ expression profiling uncovers a gene module in coronary artery disease involving transendothelial migration of leukocytes and LIM domain binding 2: the Stockholm Atherosclerosis Gene Expression (STAGE) study. *PLoS Genet*. 2009;5:e1000754. doi: 10.1371/journal.pgen.1000754.
11. Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet*. 2009;41:334–341.
12. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–678.
13. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol*. 2008;6:e107. doi: 10.1371/journal.pbio.0060107.
14. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010;26:2336–2337. doi: 10.1093/bioinformatics/btq419.
15. The 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56–65.
16. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009;106:9362–9367. doi: 10.1073/pnas.0903103106.
17. Kindt AS, Navarro P, Semple CA, Haley CS. The genomic signature of trait-associated variants. *BMC Genomics*. 2013;14:108. doi: 10.1186/1471-2164-14-108.
18. Zhong H, Beaulaurier J, Lum PY, Molony C, Yang X, Macneil DJ, et al. Liver and adipose expression associated SNPs are enriched for association to type 2 diabetes. *PLoS Genet*. 2010;6:e1000932. doi: 10.1371/journal.pgen.1000932.
19. Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am J Hum Genet*. 2010;86:581–591. doi: 10.1016/j.ajhg.2010.02.020.
20. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–29. doi: 10.1038/75556.
21. Skogsberg J, Lundström J, Kovacs A, Nilsson R, Noori P, Maleki S, et al. Transcriptional profiling uncovers a network of cholesterol-responsive atherosclerosis target genes. *PLoS Genet*. 2008;4:e1000036. doi: 10.1371/journal.pgen.1000036.
22. Greenawalt DM, Dobrin R, Chudin E, Hatoum IJ, Suver C, Beaulaurier J, et al. A survey of the genetics of stomach, liver, and adipose gene

expression from a morbidly obese cohort. *Genome Res*. 2011;21:1008–1016. doi: 10.1101/gr.112821.110.
23. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet*. 2012;44:502–510. doi: 10.1038/ng.2205.
24. Fu J, Wolfs MG, Deelen P, Westra HJ, Fehrmann RS, Te Meerman GJ, et al. Unraveling the regulatory mechanisms underlying tissue-dependent genetic variation of gene expression. *PLoS Genet*. 2012;8:e1002431. doi: 10.1371/journal.pgen.1002431.
25. Grundberg E, Small KS, Hedman ÅK, Nica AC, Buil A, Keildson S, et al; Multiple Tissue Human Expression Resource (MuTHER) Consortium. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet*. 2012;44:1084–1089. doi: 10.1038/ng.2394.
26. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science*. 2009;325:1246–1250. doi: 10.1126/science.1174148.
27. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M. Linking disease associations with regulatory information in the human genome. *Genome Res*. 2012;22:1748–1759. doi: 10.1101/gr.136127.111.
28. Becker KG, Barnes KC, Bright TJ, Wang SA. The genetic association database. *Nat Genet*. 2004;36:431–432. doi: 10.1038/ng0504-431.
29. McDonough CW, Gong Y, Padmanabhan S, Burkley B, Langaee TY, Melander O, et al. Pharmacogenomic association of nonsynonymous SNPs in SIGLEC12, A1BG, and the selectin region and cardiovascular outcomes. *Hypertension*. 2013;62:48–54. doi: 10.1161/HYPERTENSIONAHA.111.00823.

CLINICAL PERSPECTIVE

Genome-wide association studies now have identified at least 46 DNA variants mediating the inherited risk of suffering coronary artery disease (CAD). However, these risk variants only account for a small fraction (≈10%) of the total variation of CAD risk in the population. Thus, revealing information to patients on the risk for CAD based on these 46 variants can be misleading. In the current study, we examined how DNA variants are expressed across 7 disease-relevant tissues in patients with severe CAD. Apart from identifying several genes responsible for mediating the risk of the 46 established DNA risk variants, we also identified new risk variants characterized by the fact that they are regulating gene expression across multiple tissues. Among these, we identify a set of 16 multitissue DNA variants associated with 19 master regulatory genes that seem particularly important as carriers of CAD risk. We suggest that these novel DNA risk variants may be added to the 46 established ones to help better understand particularly late risk of CAD. In addition, for 3 of these 19 master regulatory genes we found strong evidence that they are essential for arterial wall plaque development and as such may serve as new targets for CAD intervention.