ESM Methods

Hyperinsulinemic-euglycemic clamp procedure

During the hyperinsulinemic-euglycemic clamp [1], a priming dose of human insulin (Novolin, Clayton, NC) was followed by a constant rate (60 mU m^{-2} min^{-1}) of insulin infusion for 120 minutes, with the goal of achieving a plasma insulin concentration of 100 μIU/mL or greater. Blood was sampled every 5 minutes, and the rate of 20% dextrose co-infused was adjusted to maintain plasma glucose concentrations at 95 to 100 mg/dL. Blood samples were drawn for glucose and insulin measurement at -30, -20, -10, +100, +110, and +120 minutes.

Data cleaning for the heritability and linkage studies (HTN-IR cohort)

Pedigree relationships were confirmed using the entire genome scan data and the software RELCHECK [2, 3]. Problematic pedigrees/individuals were modified and/or removed from further analysis. In the Phase 1 sample, one complete pedigree and 1 individual in a second pedigree were removed from analysis due to gender mismatch, high genotyping failure rate, or inconsistency with Mendelian inheritance. In the Phase 2 sample, 13 likely mis-specified familial relationships were modified from 11 families. One complete pedigree and 9 individuals were removed from analysis. Mendelian inconsistency was then examined using the software PEDCHECK [4]. Markers with Mendelian inconsistencies (genotyping errors) were converted to missing. Among the cleaned genotyped sample, 513 individuals from 140 families (227 individuals from 67 families in Phase 1; 286 individuals from 73 families in Phase 2) had SSPI data and were used in the heritability and linkage analyses (ESM Figure 1).

Cardio-Metabochip design

The Cardio-Metabochip is a custom Infinium II based iSelect high throughput genotyping platform designed in collaboration with Illumina (San Diego, CA). Its purpose was to provide a method by which loci associated in GWAS with a number of traits related to cardiac and metabolic disease could be rapidly genotyped and fine mapped. In a collaborative effort between the following GWAS consortia--CARDIoGRAM, DIAGRAM, GIANT, MAGIC, Global Lipids, ICBP-GWAS and QT-IGC--investigators from each of these consortia contributed content for the chip design. Five criteria were used to determine SNPs for the chip: (a) a single set of SNPs selected for being GWAS meta-analysis signals for replication, (b) a fine mapping set of SNPs for >250 GWAS based signals (constituting 62% of the total SNPs on the chip), (c) all SNPs with genome wide significance for any trait, (d) wildcards for consortium specific purposes, and (e) a number of SNPs selected for other purposes, such as tagging copy number polymorphisms, X and Y chromosome SNPs, fingerprint SNPs, mitochondrial DNA tagging SNPs, and major histocompatibility complex SNPs. Twelve primary metabolic and atherosclerotic traits were designated as Tier 1 traits, and ~5,000 replication SNPs were selected for each of these traits (type 2 diabetes, fasting glucose, myocardial infarction, coronary artery disease, LDL-C, HDL-C, triglycerides, BMI, waist-hip ratio, systolic blood pressure, diastolic blood pressure and QT interval). A second tier of traits was determined, with ~1000 replication SNPs per trait assigned (fasting insulin, two hour glucose, HbA1c, type 2 diabetes age of diagnosis, type 2 diabetes early onset, waist circumference, height, percent fat mass, total cholesterol, platelet count, mean platelet volume and white blood count). The final tier for replication was created to include validated SNPs associated with any GWAS trait plus a proxy. This resulted in 217,697 SNPs

being submitted for design, and a final chip content of 196,725 SNPs being included on the platform.

Cardio-Metabochip genotyping and quality control in HTN-IR and MACAD

Genotyping was performed at the Medical Genetics Institute at Cedars-Sinai Medical Center using custom Infinium II technology, following the manufacturer's protocol (Illumina, San Diego, CA) [5, 6]. Of the study samples genotyped, sample quality control measures removed 16 samples for low genotyping rate (<98%) or low p10GC. Twenty six subjects were removed for errors in gender estimates, which were calculated within Genome Studio (Illumina, San Diego, CA).

Following these quality control steps, 640 HTN-IR and 678 MACAD subjects with SSPI data were identified and included in the study. The genotyping rate in these samples in HTN-IR was 99.98%, and in MACAD was 99.96%. Across the two projects 22 pairs of sample duplicates were run (representing 1% of the entire sample run as either within-plate or across plate duplicates), yielding an average reproducibility of 99.99%.

196,475 SNPs were available from Genome Studio for the quality control pipeline. A total of 37,337 SNPs were excluded due to QC parameters that included poor cluster formation and SNP failure rate >2%. Further QC excluded SNPs with minor allele frequency <1% (HTN-IR: 37,238; MACAD: 35,992), a test of Hardy-Weinberg Equilibrium with a P<0.0000001 (HTN-IR: 3; MACAD: 7), or observed heterozygosity >53% (HTN-IR: 1,135; MACAD: 709 SNPs). Call rates for SNPs passing quality control ranged from 0.98-1.0. The final number of SNPs available for analysis post quality control was about 120K SNPs.

Population stratification

A potential pitfall in association analysis is population stratification: systematic differences in ancestry associated with phenotypes that might be associated with disease. The Hispanic-American population is significantly substructured [7]; therefore, addressing the possible confounding effect of stratification is important. We computed principal components (PCs) of ancestry for unrelated founders and then projected to all family members using the program SMARTPCA, which is distributed with the software package EIGENSTRAT [8]. The principal component analysis was performed using 43K autosomal SNPs in the HTN-IR project, and 43K SNPs in the MACAD project, with SNPs selected for minimal linkage disequilibrium (r^2<0.2). Two outliers (defined as >10 standard deviations) were identified in each of the two cohorts; these subjects were excluded from association analyses, yielding sample sizes of 638 for HTN-IR and 676 for MACAD (clinical characteristics in ESM Table 1). In the HTN-IR sample, the top PC explained ~1.35% and the second top PC explained ~0.5% of genetic variance, with the remaining PCs each explaining less than 0.4% of variance. We therefore adjusted for the top two PCs in the following association analysis. In the MACAD sample, the top three PCs explained 1.11%, 0.55%, and 0.43% respectively, with the remaining PCs each explaining less than 0.4%. Therefore, the top 3 PCs were adjusted for in the following association analyses in the MACAD sample.

References

[1] DeFronzo RA, Tobin JD, Andres R (1979) Glucose clamp technique: a method for quantifying insulin secretion and resistance. Am J Physiol 237: E214-223

[2] Boehnke M, Cox NJ (1997) Accurate inference of relationships in sib-pair linkage studies. Am J Hum Genet 61: 423-429

[3] Broman KW, Weber JL (1998) Estimation of pairwise relationships in the presence of genotyping errors. Am J Hum Genet 63: 1563-1564

[4] O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63: 259-266

[5] Gunderson KL, Kuhn KM, Steemers FJ, Ng P, Murray SS, Shen R (2006) Whole-genome genotyping of haplotype tag single nucleotide polymorphisms. Pharmacogenomics 7: 641-648

[6] Gunderson KL, Steemers FJ, Ren H, et al. (2006) Whole-genome genotyping. Methods Enzymol 410: 359-376

[7] Bryc K, Velez C, Karafet T, et al. (2010) Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc Natl Acad Sci U S A 107 Suppl 2: 8954-8961

[8] Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904-909