

Supplementary Methods

Genome Complexity Reduction

Sample DNAs should not be highly degraded nor contain PCR inhibitors, such as high concentrations of heme or chelating agents. For each individual assayed, 250 ng of genomic DNA are digested separately with 10 U of *Xba*I or *Hind*III (New England BioLabs) in volumes of 20 µL for 2 hours at 37 °C. Following heat inactivation at 70 °C for 20 minutes, 0.25 µM of *Xba*I adaptor (5'-ATT ATG AGC ACG ACA GAC GCC TGA TCT-3' and 5'phosphate –CTA GAG ATC AGG CGT CTG TCG TGC TCA TAA-3')(Affymetrix), or *Hind*III adaptor (5'-ATT ATG AGC ACG ACA GAC GCC TGA TCT-3' and 5'phosphate –AGC TAG ATC AGG CGT CTG TCG TGC TCA TAA-3')(Affymetrix) are ligated to the digested DNAs with T4 DNA Ligase (New England BioLabs) in 25 µL for 2 hours at 16 °C. The ligations are stopped by heating to 70 °C for 20 minutes, and then diluted 4- fold with water. For each ligation reaction, two to three PCRs are run in order to generate > 40 µg of PCR products. Each PCR contains 10 µL of the diluted ligation reactions (25 ng of starting DNA) in 100 µL volumes containing 1.0 µM of primer (5'-ATT ATG AGC ACG ACA GAC GCC TGA TCT-3'), 0.30 mM dNTPs, 1.0 mM MgSO₄, 5 U Platinum[®] *Pfx* Polymerase (Invitrogen), PCR Enhancer (Invitrogen) and *Pfx* Amplification Buffer (Invitrogen). 30 cycles of PCRs are run with the following cycling program: 94 °C denaturation for 15 seconds, 60 °C annealing for 30 seconds, and 68 °C extension for 60 seconds. As a check, 3 µL of PCR products are visualized on 2% TBE agarose gels to confirm the size range of amplicons. The PCR products are purified over MinElute 96 UF PCR Purification plates (Qiagen), and recovered in 40 µL of EB buffer (Qiagen). PCR yields are measured by absorbance readings at 260 nm, and adjusted to a concentration of 40 µg per 45 µL. To allow efficient hybridization to 25-mer oligonucleotide probes, the PCR products are fragmented to < 100 bp with DNase I. 0.20 U of DNase I (Affymetrix) is added to 40 µg of purified PCR amplicons in a 55 µL volume containing Fragmentation Buffer (Affymetrix) for 35 minutes at 37 °C, followed by heat inactivation at 95 °C for 15 minutes. Fragmentation products are visualized on 4% TBE agarose gels. The 3' ends of the fragmented amplicons are biotinylated by adding 214 µM of a proprietary DNA labeling reagent (Affymetrix) using Terminal Deoxynucleotidyl Transferase (Affymetrix) in 70 µL volumes for 2 hours at 37 °C, followed by heat inactivation at 95 °C for 15 minutes.

Allele Specific Hybridization to Oligonucleotide Arrays

The fragmented and biotinylated PCR amplicons are combined with 11.5 µg/mL human Cot-1 (Invitrogen) and 115 µg/mL herring sperm (Promega) DNAs. The DNAs are added to a hybridization solution containing 2.69 M tetramethylammonium chloride (TMACl), 5.77 mM EDTA, 56 mM MES, 5 % DMSO, 2.5 X Denhardt's solution, and 0.0115% Tween-20 in a final volume of 260 µL. The hybridization solution was heated to 95 °C for 10 minutes then placed on ice. After warming to 48 °C for 2 minutes, 200 µL of the hybridization solution is injected into cartridges housing the oligonucleotide arrays (Affymetrix GeneChip[®] 100K Mapping Set: 50K Array *Xba* 240 and 50K Array

Hind 240). Hybridizations are carried out at 48 °C for 16 to 18 hours in a rotisserie rotating at 60 rpm. Following the overnight hybridization, the arrays are washed with 6X SSPE and 0.01% Tween-20 at 25 °C, then more stringently washed with 0.6X SSPE and 0.01% Tween-20 at 45 °C. Hybridization signals are generated in a three step signal amplification process: 10µg/mL streptavidin R-phycoerythrin (SAPE) conjugate (Molecular Probes) is added to the biotinylated targets hybridized to the oligonucleotide probes, and washed with 6X SSPE and 0.01% Tween-20 at 25 °C; followed by the addition of 5µg/mL biotinylated goat anti-streptavidin (Vector) to increase the effective number of biotin molecules on the target; and finally SAPE is added once again and washed extensively with 6X SSPE and 0.01% Tween-20 at 30 °C. The SAPE and antibody were added to arrays in 6X SSPE, 1X Denhardt's solution and 0.01% Tween-20 at 25 °C for 10 minutes each. Following the final wash, the arrays are kept in Holding buffer (100mM MES, 1M [Na⁺], 0.01% Tween-20). The washing and staining procedures are run on Affymetrix fluidics stations. Arrays are scanned using GCS3000 scanners with AutoLoaders (Affymetrix). Scan images are processed to get hybridization signal intensity values using GCOS 2.0 software (Affymetrix). The DM genotype calling algorithm is implemented in GenoTyping Tools (GTT) (Affymetrix) and GDAS 3.0 (Affymetrix) analysis software.

Sources of Samples and Reference Genotypes

Sample DNAs from the Human Variation (African-American, Caucasian, Chinese, Japanese, Pacific, and Southeast Asian) and Polymorphism Discovery panels, as well as, CEPH and NIGMS families were purchased from the Coriell Institute for Medical Research (Camden, New Jersey). In order to make comparisons with 42 African-Americans, and 42 Caucasians; 15 Japanese, ten Chinese, ten Southeast Asian, and seven Pacific samples were grouped as 42 East Asians. 62 samples that are part of the American Diabetes Association (ADA) study on the Genetics of non-insulin dependent diabetes mellitus (GENNID) were also acquired from Coriell; these DNAs from affected families were collected and characterized by the GENNID Study Group, which includes Eric Boerwinkle, PhD, University of Texas Health Science Center; John Buse, MD, PhD, University of North Carolina; Ralph DeFronzo, MD, University of Texas Health Science Center; David Ehrmann, MD, University of Chicago; Steven C. Elbein, MD, University of Utah/University of Arkansas; Wilfred Fujimoto, MD, and Steven E. Kahn, MB, ChB, University of Washington; Craig L. Hanis, PhD, University of Texas Health Science Center; Richard A. Mulivor, PhD, and Jeanne C. Beck, PhD, Coriell Cell Repositories; Jill Norris, PhD, University of Colorado School of Medicine; M. Alan Permutt, MD, and Philip Behn, MD, Washington University School of Medicine; Leslie Raffel, MD, Cedars-Sinai Medical Center; and David C. Robbins, MD, Medlantic Research Institute. Three samples from the Autism Research Resource were obtained from Coriell as well. Sample information is listed in **Supplementary Table 5A**.

Genotyping data from the International HapMap Project (release 8, June 2004) was downloaded from www.hapmap.org in accordance with the project Data Access Policy. In cases of duplicated data from more than one HapMap center or replicate sets of HapMap genotypes for a select number of individuals, the first instance was arbitrarily chosen in the comparisons to determine concordance. Genotyping by dideoxy sequencing was performed by SeqWright.