The GENNID Study

A resource for mapping the genes that cause NIDDM

LESLIE J. RAFFEL, MD
DAVID C. ROBBINS, MD
JILL M. NORRIS, PHD
ERIC BOERWINKLE, PHD
RALPH A. DEFRONZO, MD
STEVEN C. ELBEIN, MD
WILFRED FUJIMOTO, MD
CRAIG L. HANIS, PHD
STEVEN E. KAHN, MB, CHB
M. ALAN PERMUTT, MD
KEN C. CHIU, MD
JAIME CRUZ, MD, MPH
DAVID A. EHRMANN, MD
R. PAUL ROBERTSON, MD
JEROME I. ROTTER, MD
JOHN BUSE, MD

OBJECTIVE — To develop a resource, consisting of comprehensive data and lymphoblastoid cell lines, of well-characterized NIDDM families that will be available to the scientific community for genetic studies of NIDDM.

RESEARCH DESIGN AND METHODS — Non-Hispanic white, Hispanic, African-American, and Japanese-American multiplex NIDDM families, with a minimum of one affected sib-pair, are being collected by the eight Harold Rifkin Family Acquisition Centers. Detailed family and medical histories are obtained from all participants. Family members with diabetes have fasting blood samples drawn, while nondiabetic family members have an oral glucose tolerance test and, when possible, insulin sensitivity and insulin secretion measurements by frequently sampled intravenous glucose tolerance testing or euglycemic insulin clamp. Lymphoblastoid cell lines are established for all participants.

RESULTS — Over 1,400 individuals from ~220 families have been studied since the start of the GENNID (Genetics of NIDDM) program in July 1993. The goal is that by July 1997, data from 300 non-Hispanic white families, >100 Hispanic families, >100 African-American families, and 15 Japanese-American families will have been collected.

CONCLUSIONS — The identification of the genes responsible for NIDDM may now be achievable, but only if sound phenotypic data are linked to genetic material from a large number of well-described multiplex families. The GENNID project of the American Diabetes Association is creating a comprehensive resource that will expedite the identification of the genetic basis of NIDDM.

NIDDM imparts a heavy burden on our public's health and results in significant personal and societal costs. In 1992, patients with diabetes constituted 4.5% of the U.S. population but accounted for 14.6% of total U.S. health care expenditures ($105 billion) (1). Approximately one in seven health care dollars is currently spent on caring for individuals with diabetes, and 95% of these patients have NIDDM (1). When the cost to individuals in terms of lost wages for affected individuals and their family members is also considered, the financial burden of this disease is staggering.

The impact of NIDDM goes far beyond economic issues. The disease shortens the life span and is associated with a variety of disabling complications, such as atherosclerosis, renal failure, and blindness. The incidence of NIDDM is rising, and the disease is especially prevalent among the aged, the poor, and many ethnic minorities (2–4). As the U.S. population continues to age and becomes more multiethnic, NIDDM will have an even greater public health impact. Thus, identifying the underlying cause of NIDDM, with the goals of preventing the disease and developing more effective therapy, is becoming increasingly important.

In the U.S., the prevalence of NIDDM in Hispanic adults is three to five times that in non-Hispanic whites, making diabetes one of the most common chronic diseases in this rapidly growing population (5–13). It is thought that this increased rate of NIDDM results from native Amerindian gene admixture in the Hispanic population (14,15). Japanese Americans have also been reported to have increased diabetes rates, with up to 20% of Nisei (second-generation) men and women affected (16,17). In the African-American population, the prevalence of NIDDM is 60% higher than in non-Hispanic whites (18). Although part of these differences in disease prevalence may be due to environmental factors, excess diabetes prevalences have been reported even when efforts were made to control for environmental risks (18).

Several lines of evidence implicate genetic factors in the etiology of NIDDM. These data include the differential concordance rates between monozygotic and dizygotic twins, increased risk associated with a positive family history, marked differences in prevalence among various ethnic groups, and the fact that at least some

From the Cedars-Sinai Research Institute and University of California (L.J.R., J.I.R.), Los Angeles, California; Medlantic Research Institute (D.C.R.), Washington, DC; University of Colorado Health Sciences Center (J.M.N.), Denver, Colorado; University of Texas Health Science Center (E.B., C.L.H.), Houston, and University of Texas Health Science Center (R.A.D., J.C.), San Antonio, Texas; Department of Veterans Affairs Medical Center and University of Utah (S.C.E.), Salt Lake City, Utah; University of Washington (W.F., S.E.K.) and Department of Veterans Affairs Medical Center (S.E.K.), Seattle, Washington; Washington University School of Medicine (M.A.P., K.C.C.), St. Louis, Missouri; University of Chicago (D.A.E.), Chicago, Illinois; University of Minnesota (R.P.R.), Minneapolis, Minnesota; and University of North Carolina School of Medicine (J.B.), Chapel Hill, North Carolina.

K.C.C. is currently at the University of California Los Angeles School of Medicine, Los Angeles, California.

Address correspondence and reprint requests to Leslie J. Raffel, MD, Associate Director, Common Disease Genetics Program, Division of Medical Genetics, Cedars-Medical Center, 8700 Beverly Blvd., SSB-378, Los Angeles, CA 90048. E-mail: lraffel@mailgate.csmc.edu.

Received for publication 11 December 1995 and accepted in revised form 18 April 1996.

CAP, College of American Pathologists; FBS, fetal bovine serum; FSIVGTT, frequently sampled intravenous glucose tolerance test; GENNID, Genetics of NIDDM; K_g, glucose disappearance constant; OGTT, oral glucose tolerance test; R_a, rate of glucose appearance; R_d, rate of glucose disappearance; S_I, insulin sensitivity; S_G, glucose effectiveness.

Table 1—*Enrollment criteria*

1. Each family must have a minimum of two siblings with documented NIDDM. Diabetes is to be documented using National Diabetes Data Group criteria. If both parents of the index cases are known to have diabetes or have been told by a health care professional that they have "borderline" diabetes, the family should not be enrolled.
2. Access to three first-degree relatives (in addition to the two NIDDM siblings) is required. First-degree relatives include parents, siblings, and offspring. Whenever possible, all first-degree relatives should be studied. If additional family members with NIDDM are studied, attempts should be made to study their first-degree relatives as well. Family members must be >18 years of age to be enrolled. (1993–1995)
3. The nondiabetic spouse of at least one of the NIDDM siblings should be studied as a control. If no such spouse is available, an unrelated nondiabetic spouse of another participating family member may be substituted. (1993–1995)
4. In addition to the minimum two affected siblings, other affected siblings and the parents (diabetic or nondiabetic) should be studied when available. If both parents are not available for study, up to two nondiabetic siblings should be studied as available. (1995–)

NIDDM cases can be attributed to alterations in single genes (19,20). In dizygotic twin pairs, the concordance rate is approximately the same as that for all siblings, while the monozygotic twin concordance rate is significantly higher, approaching 100% in some studies. A positive family history for diabetes is reported by 25–50% of people with diabetes, as compared with no more than 15% of nondiabetic individuals (19). Diabetes prevalence varies greatly among ethnic groups, and certain Native American populations, such as the Pima Indians, have among the highest rates of NIDDM in the world. The significant positive relationship between the prevalence of NIDDM among populations and the proportion of Amerindian genes strongly suggests that genetic risk factors play an important role in these populations (15,21).

Some progress is being made in identifying genetic alterations that can cause NIDDM. For example, it has been demonstrated that mutations in the insulin gene, the insulin receptor gene, the glucokinase gene, and a mitochondrial tRNA gene can all cause diabetes (23–31). However, these defects are responsible for only a small percentage of NIDDM cases. In spite of intensive research efforts, the genetic and molecular bases for most NIDDM cases remain unknown.

At least part of the reason for previous failures to localize genes contributing to NIDDM risk in the general population is attributable to the lack of DNA and clinical data on a sufficiently large sample of informative pedigrees with NIDDM. With technological developments in analyzing the structure of the human genome and the construction of genetic linkage maps, the tools now exist to unravel the genetic basis for almost all forms of NIDDM. Such research will undoubtedly rely on genetic material derived from cell lines from families in which diabetes is frequent. This paper describes methods being used by GENNID (Genetics of NIDDM), the American Diabetes Association's multicenter project designed to develop such a resource.

THE GENNID PROJECT— The American Diabetes Association and its scientific membership recognize the compelling need to identify and characterize multiplex NIDDM families for the purposes of genetic studies. GENNID was designed to establish such a comprehensive, well-characterized resource. GENNID is identifying multiplex NIDDM pedigrees from several racial and ethnic groups (non-Hispanic whites, Hispanics, African-Americans, and Japanese-Americans). Historical data, physical exam findings, basic metabolic testing, and transformed lymphoblastoid cell lines are being collected. The genetic material and the extensive database characterizing each donor are now available to scientists with plans to study the genetics of NIDDM using both linkage and association approaches. It is hoped that this valuable, robust, and carefully gathered resource will be a vital key to understanding the complex genetic basis of NIDDM.

Sampling design

The primary goal of GENNID is to develop a family-based resource to identify genes important in the etiology of commonly occurring forms of NIDDM. It was essential to maximize the probability that informative families would be recruited. To achieve this goal, multiplex families (i.e., families with a minimum of two documented sibling cases of NIDDM available for study) were necessary. At the same time, however, there was concern that recruiting only highly selected, very large families with many cases of NIDDM would enrich the sample with rare single-gene forms of NIDDM. The decision was therefore made to establish minimum entry criteria for a family and then recruit all families meeting those criteria. Although methods exist for analysis of affected relatives of varying relationships, the most straightforward approaches involve the use of affected sib-pairs. Therefore, the enrollment criteria for GENNID (Table 1) require the availability of at least two siblings (the index cases) affected with NIDDM, using National Diabetes Data Group criteria (either a fasting plasma glucose concentration of 7.8 mmol/l [≥140 mg/dl] on more than one occasion or a plasma glucose concentration of 11.1 mmol/l [≥200 mg/dl] in the 2-h sample and in at least one other sample during an oral glucose tolerance test [OGTT]) (32). Between July 1993 and early 1995, enrollment also required at least three available additional first-degree relatives (parents, siblings, or offspring) of either of these affected individuals (regardless of diabetes status) and one nondiabetic ethnically matched individual (usually the unaffected spouse of one of the index cases) to act as an environmental control, who could be used for marker association (i.e., case control) studies. Examples of pedigree structures meeting these criteria are illustrated in Fig. 1.

Beginning in early 1995, enrollment criteria were changed to maximize the collection of informative affected sibling pairs because the consensus of the GENNID Oversight Committee (consultants to the American Diabetes Association) was that the majority of investigators who would want to use the GENNID resource currently employ affected-sib-pair analytic approaches. Therefore, current ascertainment criteria are *1*) the availability of at least two siblings (the index cases) affected with NIDDM according to National Diabetes Data Group criteria (32) and 2) the availability of one or both parents *or*, if neither or only one parent can be studied, the availability of at least two additional siblings. A maximum of two nondiabetic siblings can be included; any

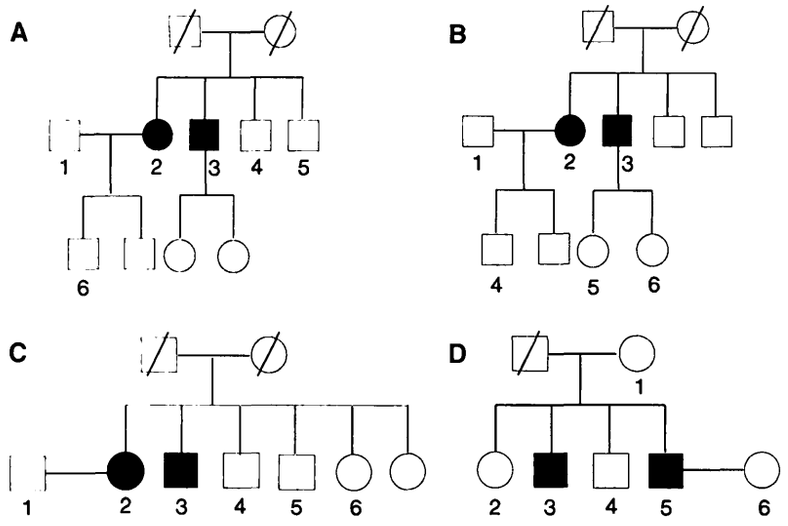

Figure 1—*Example pedigrees meeting minimum phase 1 GENNID entry criteria (1993–1995). Numbers indicate the six individuals being studied to have the family qualify.* ●, ■, *individuals with NIDDM.*

number of additional siblings with NIDDM can be studied. Examples of pedigrees meeting these criteria are illustrated in Fig. 2.

It is likely that a variety of genes and/or alleles are important in causing diabetes in different ethnic groups. Thus, GENNID Family Acquisition Centers were established to allow inclusion of multiple ethnic and racial groups. During the first 2 years, four of the eight GENNID Family Acquisition Centers primarily ascertained non-Hispanic white families, three centers recruited Hispanic families, and one center recruited African-American families. One center recruiting non-Hispanic whites also recruited Japanese-American families. Since July 1995, those centers recruiting Hispanic families have also been recruiting non-Hispanic white and African-American families, and the centers recruiting non-Hispanic white families have also been recruiting African-American families. The ultimate goal is to collect 300 non-Hispanic white families, 100 African-American families, and 100 Hispanic families (see below). Since many families will have more than the minimum of two affected siblings, these numbers represent an underestimate of the total number of sibling pairs anticipated to be available.

Sample size and power. Determining the minimum number of families needed for a genetic study of this type is problematic because the number of contributing genes and the frequencies and effects of alleles at these genes are unknown. Estimates of power and sample sizes also vary depending on whether a candidate gene or a systematic mapping approach is used, along with the degree of polymorphism of the DNA markers. The use of qualitative traits (e.g., presence/absence of diabetes) or quantitative traits (e.g., insulin response to a glucose load) and the methods of analysis selected will also influence power. In this section, we provide estimates of power under multiple statistical models.

Using a qualitative trait and a candidate gene approach, the sample size required to establish linkage with nonparametric approaches can be quite small when the gene has an effect that is moderate to large, even in the face of significant genetic heterogeneity (Table 2) (33). The power of the sib-pair linkage test is a function of the number of loci, the size of allelic effects, the frequency of the allele, and the distance between the markers and the gene of interest. In practice, qualitative sib-pair approaches can effectively demonstrate linkage between marker loci and disease susceptibility loci using relatively modest sample sizes (34,35). For example, Cudworth and Woodrow (35) reported evidence for linkage of the HLA region with insulin-dependent diabetes in 15 affected sib-pairs, one affected sib-trio, and one affected sib-quad, showing increased sharing of haplotypes identical by descent among affected siblings (35).

Systematic mapping approaches, in addition to candidate gene methods, will likely be used by many investigators; thus, the sample size must be large enough to allow for the demonstration of linkage under such conditions. Factors that must be taken into account include the density of the linkage map, the availability of parents (for determination that alleles are identical by descent and not just identical by state), the degree of polymorphism of the markers used, and the number of genes responsible for a given phenotype (36). Although it is unlikely that the GENNID database will be large enough to allow identification of *all* susceptibility loci, a significant number of such loci should be detectable, particularly in the non-Hispanic white population.

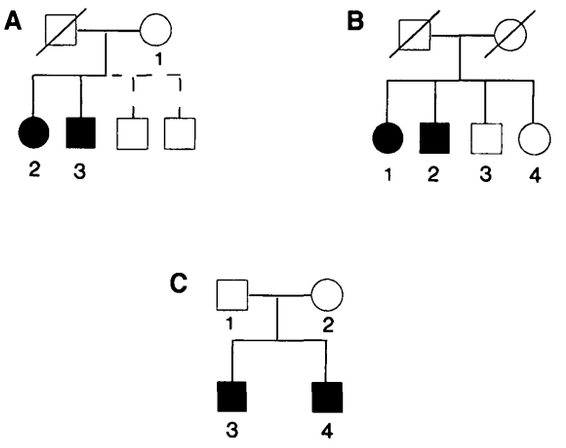

Figure 2—*Example pedigrees meeting phase 2 GENNID entry criteria (1995-). Numbers indicate the individuals being studied to have the family qualify. In a family with one parent available for study (A), up to two unaffected siblings can be studied. In all three types of family structure, an unlimited number of additional diabetic siblings can be studied.* ●,■, *individuals with NIDDM.*

Table 2—*Sample sizes for a qualitative sib-pair linkage analysis, assuming multiple loci acting independently with incomplete dominant expression, power = 0.90, and a one-sided significance = 0.001*

Number of disease loci	Sib-pair or trio	Required sample size
1	1, 1	52
	1, 1, 1	22
	0, 1, 1	28
	0, 0, 1	45
	0, 1	101
2	1, 1	222
	1, 1, 1	92
	0, 1, 1	115
	0, 0, 1	181
	0, 1	417
4	1, 1	903
	1, 1, 1	360
	0, 1, 1	451
	0, 0, 1	707
	0, 1	1,675

Depending on the number of loci contributing to a phenotype and the combinations of affected and unaffected siblings available in each family, the sample size required to demonstrate linkage varies. This table gives examples of the sample sizes necessary to show linkage with the types of sibling combinations that are being collected in GENNID. The penetrance function for one disease locus is f(A1A1) = f(A1a1) = 0.8 and f(a1a1) = 0. For penetrance function for two disease loci, see National Diabetes Data Group (32), Table 2. For the siblings studied, 1, affected; 0, unaffected. Adapted from Ref. 32.

With quantitative sib-pair analysis, factors that affect power include the heritability of the trait analyzed, the frequency of the gene that controls the trait in the population, the number of alleles at the marker locus, the number of sib-pairs, the recombination fraction between the marker and disease allele, and the level of significance (Table 3) (37). In addition, several studies have demonstrated the effectiveness of relatively small sample sizes (37,38). Boerwinkle et al. (38), for example, showed that most variation in plasma lipoprotein (a) concentration is due to differences in the apolipoprotein (a) gene using a total of 288 individuals (72 sib-pairs).

The GENNID protocol

As diagrammed in Fig. 3, individuals enrolled in the GENNID project undergo several evaluations. Index cases and other family members with documented NIDDM have fasting blood samples collected to measure glucose, insulin (unless receiving insulin therapy), and lipids and to obtain lymphoblastoid cell lines. Family members not known to have diabetes additionally have an OGTT. Those whose OGTTs demonstrate normal or impaired glucose tolerance also, when possible, undergo testing to assess insulin sensitivity and insulin secretion. At most centers, a frequently sampled intravenous glucose tolerance test (FSIVGTT) is performed, while at the University of Texas, San Antonio, a euglycemic insulin clamp is used.

Questionnaires. Although preliminary pedigree information is collected in the process of determining a family's eligibility, more comprehensive information is collected at the study visit. Information regarding age, sex, relationship to the index cases, height, weight, and medical history is collected on a one-page Family Form for each individual possible in three generations, beginning with the index cases' parental generation. In addition to this screening information, a more detailed Family/Medical History Questionnaire is completed by all study participants. Information regarding ethnic origin, education, occupation, weight, activity level, tobacco and alcohol use, medical conditions, medication usage, and diet is collected. Women are also asked about their pregnancy (including history of gestational diabetes), contraceptive, and menopause history. Individuals with NIDDM are asked their date of diagnosis, along with symptoms and weight at diagnosis. Both current and previous treatment for diabetes and detailed information concerning diabetic complications is recorded.

Physical examination. A limited physical examination is performed on all study participants, including weight, height, and waist and hip circumferences. Three consecutive measurements are taken for each to assure accuracy. Pulse and blood pressure after the participant has been sitting quietly for 5 min are also measured. Three consecutive measurements of systolic and diastolic blood pressures are taken 30 s apart on the right arm using a conventional mercury sphygmomanometer.

Initial blood sampling. Subjects are requested to fast and to refrain from smoking and medication use for at least 10 h before the study visit. Nondiabetic subjects are instructed to consume at least 300 g/day of complex carbohydrate on the 3 days before the study. All subjects have two fasting blood samples drawn 15 min apart using an indwelling catheter. These samples are used for lymphocyte isolation and lipid, glucose, and insulin measurements. The catheter is left in situ for the OGTT on all nondiabetic individuals and those with suspected but undocumented diabetes. After the administration of 75 g of glucose (Trutol, Custom Laboratories, Baltimore, MD), postchallenge blood samples are drawn at 30, 60, 90, and 120 min for glucose and insulin measurement.

FSIVGTT. Family members with normal or impaired glucose tolerance are asked to return a second time for an FSIVGTT. After placement of two intravenous catheters, three basal samples are collected and glucose, 11.4 g/m^2 body surface area as a 50% solution, is then infused over 1 min. After 20 min, tolbutamide, 125 mg/m^2 body surface area (Orinase, Upjohn, Kalamazoo, MI), is given to improve parameter identifiability (40,41). Blood samples are collected at 2, 3, 4, 5, 6, 8, 10, 14, 19, 22, 24, 27, 30, 40, 50, 70, 90, 120, 150, 180, 210, and 240 min after

Table 3—*Calculated sample size requirements for quantitative sib-pair analysis with power = 0.95 and significance = 0.05*

	Sample size required	
Heritability	Codominance	Dominance
0.5	355	400
0.7	141	178
0.9	66	98

Adapted from Ref. 34.

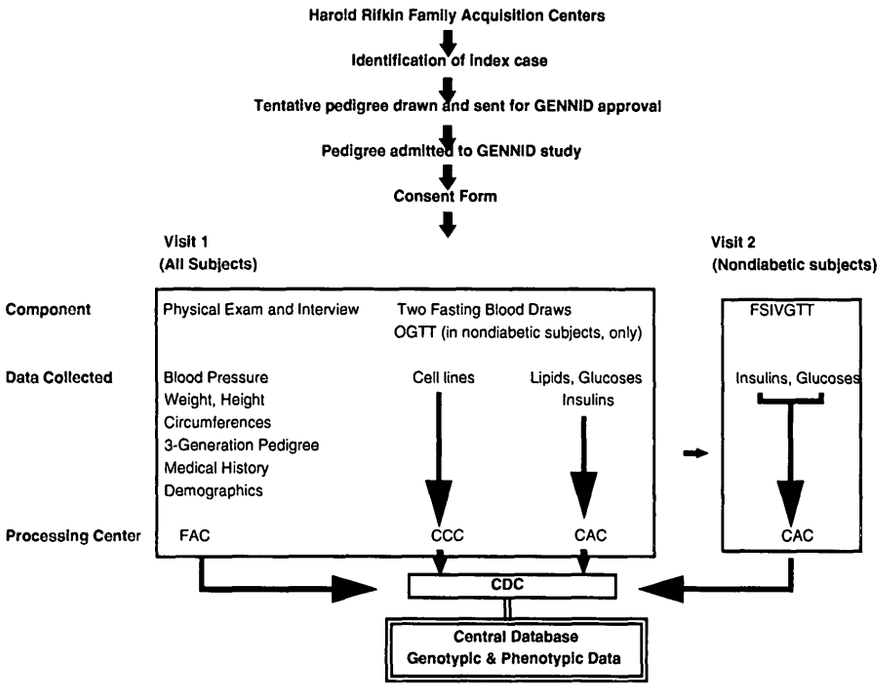

Figure 3—*Flow chart for data collection and processing in GENNID Study. CAC, Central Assay Center; CDC, Central Data Center; CCC, Central Cell-line Center; FAC, Family Acquisition Center.*

glucose injection. Insulin sensitivity (S_I) and glucose effectiveness (S_G) are derived using the minimal model of glucose kinetics (42). The first-phase insulin response is calculated as the mean incremental insulin response above basal for the first 10 min after glucose injection, and the glucose disappearance constant (K_g) is determined as the slope of the regression between time and the natural logarithm of the glucose concentrations measured at 10, 14, and 19 min.

Euglycemic insulin clamp. The euglycemic insulin clamp is used to assess insulin sensitivity in individuals studied at the University of Texas, San Antonio center (43). Two intravenous catheters are placed, and the hand with the sampling catheter is placed into a heated (60°F) box to achieve arterialization of venous blood (44). A prime constant infusion of [3-^{3}H]glucose (Du Pont-NEN, Boston, MA), 20 μCi × (fasting plasma glucose ÷ 100), is given for 270 min. After a 150-min basal tracer equilibration period, a 120-min euglycemic insulin clamp is performed (43). The rate of insulin (Humulin-R, Eli Lilly, Indianapolis, IN) infusion is 40 mU/m^2 · min. Glucose concentration is measured every 5 min, and a 20% dextrose infusion is adjusted to maintain the basal plasma glucose concentration (43). During the last 40 min of the basal period and throughout the insulin clamp, blood is drawn every 5–15 min to measure plasma [3-^{3}H]glucose specific activity and insulin concentration.

During the last 40 min of the basal period, a steady-state plateau of [3-^{3}H]glucose specific activity (disintegrations per minute per milligram) is achieved, and hepatic glucose production is determined by dividing the continuous [3-^{3}H]glucose infusion rate disintegrations per minute per minute) by the steady-state plateau of tritiated glucose specific activity (disintegrations per minute per milligram). After the administration of insulin, non–steady-state conditions exist and rates of glucose appearance (R_a) and disappearance (R_d) are calculated using Steele's equation (45). The rate of hepatic glucose production (mg/m^2 · min) is calculated by subtracting the rate of exogenous glucose infusion from the rate of total glucose appearance (R_a) determined by the isotope-tracer technique. Total body glucose uptake (mg/m^2) is calculated as the sum of the exogenous glucose infusion plus the rate of residual hepatic glucose production. Insulin sensitivity is expressed as a percentage of that obtained in nondiabetic, normal weight, ethnically and age-matched individuals. It is this latter percentage, along with the means and coefficients of variation for the glucose and insulin values at baseline and during the insulin infusion, that is stored in the main GENNID database.

Studying out-of-town family members. When some members of a family do not live near a Family Acquisition Center, arrangements are made for samples to be drawn locally. Whenever feasible, samples are obtained at one of the MetPath laboratories, which have agreed to assist with GENNID. If no MetPath laboratory is in the area, arrangements are made with the individual's physician. Modified protocols are used for these studies: blood is drawn by venipuncture, without an intravenous catheter, only a single fasting sample is obtained, and blood is collected only at 60 and 120 min in the OGTT. The Family/Medical History Questionnaire is completed for individuals studied off-site, but there is no physical examination. Notation is made in the main GENNID database to identify those subjects who are studied off-site.

Laboratory methods

Glucose, lipids, and insulin. Blood samples for glucose, insulin, and lipid measurements and for storage are collected in EDTA on ice. Samples are processed, and the plasma is aliquoted into airtight vials and stored at −20°C or lower until shipment to the Central Assay Center (Medlantic Research Institute, Penn Medical Laboratory, Washington, DC).

Glucose concentrations in EDTA plasma samples are measured in duplicate using a hexokinase method (Glucose/HK, Boehringer Mannheim, Indianapolis, IN). PreciCal EXLS-85 is used as the calibration solution (at set point 191 mg/dl). Three controls included in each glucose assay run (60 per tray) are Precitrol-N and Precitrol-A Control and National Institute of Standards in Technology traceable glucose standards (400 and 200 mg/dl). A 20-sample precision run occurs at least once every 20 working days, and the mean between-run coefficient of variation for glucose is <2%. Technical errors (variance between blinded pairs) are <4%. External calibration is achieved by participation in the College of American Pathologists (CAP) Quality Control Program.

Triglycerides are measured using a free-glycerol blanking method (Triglycerides/GB, Boehringer Mannheim). HDL cholesterol is measured after heparin–

manganese sulfate precipitation. Plasma total and HDL cholesterol are measured using Boehringer Mannheim reagents (Cholesterol/HP). The plasma pools are used as fixed-mean controls based on Abel-Kendall determined concentrations of cholesterol. Mean coefficient of variation for plasma cholesterol is 1.9%. Technical error is routinely <6% for cholesterol and <8% for triglycerides. External calibration is achieved by participation in the CAP Interlaboratory Comparison Program. The lipid measurements are certified by participation in the Centers for Disease Control–National Heart, Lung and Blood Institute Lipid Standardization Program and the Northwest Lipid Laboratory Proficiency Program.

Insulin measurements are made by radioimmunoassay using reagents from Linco Research, St. Louis, MO. Antibody 1012 is used as the first antibody. The 50% tracer displacement point (B_{50}) for proinsulin and its conversion intermediates is ~20% of that of intact human insulin. The Linco insulin standard is calibrated against the World Health Organization 1988 International Standard. Nine commercial control pools are included in each assay. Samples are run in duplicate and results are accepted if replicates agree within 15% of one another. The laboratory participates in an international sample exchange with 16 other reference laboratories and also in the CAP Comparison Program. The limit of quantitation is ~1.8 μU/ml, and the limit of detection is ~0.14 μU/ml. The coefficient of variation is <9% and the technical error is <14%.

Lymphoblastoid cell lines. Blood samples are collected in ACD (acid-citrate-dextrose) tubes and shipped overnight at room temperature to the Central Cell Line Facility (Coriell Institute for Medical Research, Camden, NJ). They are subjected to Ficoll-pacque centrifugation to isolate lymphocytes. B-cell transformation is performed aseptically using the B95-8 marmoset strain of Epstein-Barr virus with phytohemagglutinin as the mitogen. Transformation and subsequent expansion of transformed lymphocytes occurs in antibiotic-free RPMI 1640 cell culture medium with 15% heat-inactivated fetal bovine serum (FBS) and L-glutamine. When the cell culture expansion yield is at least 5×10^7 cells, 1-ml aliquots containing 5×10^6 cells in RPMI 1640, 30% heat-inactivated serum (50% FBS and 50% iron-supplemented calf serum), and 5% dimethylsulfoxide as the cryoprotectant are frozen in flame-sealed borosilicate glass ampules and placed in liquid nitrogen storage.

Before distribution, one vial of each freeze is recovered and checked for viability (trypan blue dye exclusion protocol) and growth potential. Sterility checks are performed on both freeze pools and freeze recoveries by dispensing aliquots of spent medium into trypticase soy broth, Sabouraud dextrose broth, and tryptose phosphate broth and onto blood agar plates, incubating at both 30 and 37°C. Mycoplasma testing consists of microbiological culture methods, the direct-fluorescence Hoechst-stain assay, and a polymerase chain reaction protocol using primers to mycoplasma rRNA (16s subunit). Also before distribution, DNA polymorphism testing to verify the identity of the cell lines with respect to the family relationships detailed in the pedigree structure will be performed to ensure that misidentified or contaminated cell lines will not be distributed.

The GENNID data

GENNID data can be grouped into two main categories: family data and individual data. The purpose of the family databases is to collect, store, and retrieve information specific to pedigrees, including the identity of and genetic relationship among pedigree members. Family data are collected using the Family Form and double entered into a working database. Each family is identified by a unique identification number. Family data are stored in two separate but interconvertable formats that are commonly used by most genetic analysis software. A record in the first format is defined by a pedigree, and the data on that pedigree (e.g., identification information of the pedigree members) is stored in fields for that record. A record in the second format is defined by an individual. For each individual record, the identification numbers of his/her mother and father are stored in separate fields. In this seemingly simple way, a large pedigree of nearly any configuration can be retrieved.

The purpose of the individual databases is to collect, store, and retrieve the information collected on each participant. All GENNID data are entered, managed, and stored using Paradox (DOS version 4.0). To assure quality of the database, each Family Acquisition Center (Family/Medical History Questionnaire and physical examination) and laboratory is responsible for double entry of their data. The data are then shipped on diskette to the Central Data Center, where data are extracted and compared, and discrepancies are transmitted back to the Family Acquisition Centers for resolution. Questionnaire and laboratory data are merged into the global database (housed in the GENNID Central Data Center at the University of Texas Health Science Center, Houston) only after all quality assurance steps have been taken. Research databases containing a justified subset of variables can then be constructed from this global database to address specific scientific interests.

Quality assurance

Key elements in the collection of quality data in GENNID include clear definitions and procedures as outlined in the protocol manual by the GENNID Steering Committee. Standardized data collection forms are also used by each Family Acquisition Center, as described above. Monthly conference calls, alternating between study coordinators and principal investigators, are used to discuss study progress, problems, and solutions, as are biannual meetings of the GENNID Steering Committee. Before the initiation of data collection, a centralized training session was conducted to ensure that the data collection and processing procedures were applied uniformly and correctly by study personnel. This session was attended by principal investigators and study coordinators, both of whom were responsible for training additional personnel at the study site.

Site visits were conducted 4–6 months after the start of data collection. The site-visit teams, which consisted of representatives from the Family Acquisition Centers, observed study personnel under normal operating conditions for adherence to protocol; and discrepancies, problems, and ways for improvement were discussed with the entire staff of the center being visited. A formal report was then written and distributed to the GENNID Steering Committee. Site visits of the Central Laboratory, Cell Line, and Central Data Centers were also conducted by outside expert consultants. When deemed necessary, second site visits of selected centers were conducted to verify

that the recommended corrections had been implemented.

Timetable and scope of the study

The aim of the GENNID project when initially funded was for each of the eight Family Acquisition Centers to study a minimum of 150 individuals from 10–25 families during the 2-year period from 1 July 1993 to 30 June 1995. This would result in a minimum of 600 non-Hispanic white individuals from 40–100 families, 450 Hispanic subjects from 30–75 families, 150 African-American individuals from 10–25 families, and 150 Japanese-American subjects from 10–25 families. As detailed above (see sampling design), the GENNID Oversight Committee recommended that GENNID be extended but that changes be made in recruitment to maximize smaller pedigrees with a maximum number of affected sib-pairs. As currently planned, by July 1997, GENNID will have collected 300 non-Hispanic white families, >100 Hispanic families, >100 African-American families, and 15 Japanese-American families. With the exception of the Japanese-American families, which were all recruited using the original sampling scheme, ~25% of the families in each ethnic group will be larger families collected during the first 2 years and 75% will be nuclear families collected in the second phase of the GENNID project.

Access to the cell lines and data

It is the intention of the American Diabetes Association to release GENNID phenotypic data, pedigrees, cell lines, and plasma samples to all investigators with scientifically worthy proposals. Completion of data entry and quality assurance of the database will have been accomplished for each pedigree before data or cell lines pertaining to that pedigree are released. It is anticipated that investigators who use the GENNID resource will contribute their genotype data (once published) to the GENNID database and that access to such data by multiple investigators will facilitate the rapid identification of the genes contributing to NIDDM susceptibility. Details of the application process, procedures, and costs associated with use of the GENNID resources will be published in the near future in American Diabetes Association publications.

CONCLUSIONS— The identification of the genes responsible for NIDDM may now be achievable. The genes for NIDDM will only be found if sound phenotypic data are linked to genetic material from well-described multiplex families. The GENNID project of the American Diabetes Association is creating such a database. It is our hope this comprehensive resource will expedite the identification of the genetic basis of NIDDM.

Acknowledgments— Support for GENNID is provided in part by generous contributions from the Harold Rifkin Fund, the Allied Signal Corporation, and Glaxo Wellcome, Inc. The assistance of MethPath Laboratories, where testing was performed at no cost on a number of the subjects included in the GENNID project, is also greatly appreciated. We thank the Upjohn Company for donating the tolbutamide (Orinase) used in the FSIVGTTs. This work is also supported in part by the Medical Research Service of the Department of Veterans Affairs; the Texas/Mexico Border Health Services Delivery Project; National Institutes of Health Grants HL47052 and EY06231; and General Clinic Research Center Grants RR-00037, M01-RR-00064, M01-RR-00055, and RR-00425–26.

APPENDIX

GENNID PRINCIPAL INVESTIGATOR: John Buse, MD

GENNID PROJECT OFFICERS: R. Paul Robertson, MD, and Simeon Taylor, MD, PhD

HAROLD RIFKIN FAMILY ACQUISITION CENTERS:

Cedars-Sinai Medical Center. Principal Investigator: Leslie J. Raffel, MD. Co-investigators: Jerome I. Rotter, MD; Anne Peters, MD; and Mayer Davidson, MD. Study Team: Jane Brown, RN, and Lauren Ficks, MD.

University of Chicago. Principal Investigator: David A. Ehrman, MD. Study Team: Mattie Ellison, RN; Robert J. Sobel, MD; and Jody Dushay.

University of Colorado Health Sciences Center. Principal Investigator: Jill M. Norris, PhD. Study Team: Emelin Espinoza, MSN; Shirley Gonzales; Rodney Sandoval; and Roberta Vialpando.

University of Texas Health Science Center, Houston. Principal Investigator: Craig L. Hanis, PhD. Study Team: Hilda Guerra.

University of Texas Health Science Center, San Antonio. Principal Investigator: Ralph A. DeFronzo, MD. Co-investigator: Jaime Cruz, MD, MPH. Study Team: John Kincade, RN; James King, RN; and Olga Flores.

University of Utah. Principal Investigator: Steven C. Elbein, MD. Study Team: Kim Wegner and Cindy Miles, RN.

University of Washington. Principal Investigators: Wilfred Fujimoto, MD, and Steven E. Kahn, MB, ChB. Study Team: Valerie G. Larson, ARNP, BSN; Kwang-Wen Chen, MD; Ronald L. Prigeon, MD; Michael E. Roder, MD; Lisa M. Stone, MD; Jennifer W. Cockburn; Patricia A. Kramer; Pamela Y. Yang; Jennifer H. Stephens; Pamela J. Asberry, RN, BSN; and Stein Wang

Washington University School of Medicine. Principal Investigator: M. Alan Permutt, MD. Co-investigators: Ken C. Chiu, MD, and Philip Behn, MD. Study Team: Michele Sadler, RN, and Mary Rita Shanahan, RN.

CENTRAL DATA CENTER:

University of Texas Health Science Center, Houston. Principal Investigator: Eric Boerwinkle, PhD. Database and Statistical Team: Kim Lawson and Zhiying Wang.

CENTRAL ASSAY CENTER:

Medlantic Research Institute. Principal Investigator: David C. Robbins, MD. Co-investigator: Barbara V. Howard, PhD. Laboratory Team: Kathy Gallagher, MS, and Marilyn Cedorette, MT-ASCP.

CENTER CELL-LINE CENTER:

Coriell Institute for Medical Research. Principal Investigator: Richard A. Mulivor, PhD.

IVGTT ANALYSIS CENTER:

University of Washington. Principal Investigator: Steven E. Kahn, MB, ChB.

AMERICAN DIABETES ASSOCIATION:

Inquiries regarding GENNID should be sent to the American Diabetes Association, Research Department, 1660 Duke St., Alexandria, VA 22314.

References

1. Rubin RJ, Altman WM, Mendelson DN: Health care expenditures for people with

diabetes mellitus, 1992. *J Clin Endocrinol Metab* 78:809A–809F, 1994
2. Hamman RF: Genetic and environmental determinants of non-insulin dependent diabetes mellitus (NIDDM). *Diabetes Metab Rev* 8:287–338, 1992
3. Carter JS, Wiggins CL, Becker TM, Key CR, Samet JM: Diabetes mortality among New Mexico's American Indian, Hispanic, and non-Hispanic white populations, 1958–1987. *Diabetes Care* 16:306–309, 1993
4. Geiss LS, Herman WH, Goldschmid MG, DeStefano F, Eberhardt MS, Ford ES, German RR, Newman JM, Olson DR, Sepe SJ, Steven JM, Vinicor F, Wetterhall SF, Will JC: Surveillance for diabetes mellitus: United States, 1980–89. *MMWR CDC Surveillance Summaries* 42:1–20, 1993
5. Haffner SM, Rosenthal M, Hazuda HP, Stern MP, Franco LJ: Evaluation of three potential screening tests for diabetes in a biethnic population. *Diabetes Care* 7:347–353, 1984
6. Gardner LL, Stern MP, Haffner SM, Gaskill SP, Hazuda HP, Relethford JH, Eifler CW: Prevalence of diabetes in Mexican Americans: relationship to percent of gene pool derived from Native American sources. *Diabetes* 33:86–92, 1984
7. Drury TF, Powell AL: Prevalence, impact and demography of known diabetes in the United States. In *National Center for Health Statistics Advance Data #114*. Washington, DC, U.S. Govt. Printing Office, 1986, p. 1–16
8. Stern MP: Diabetes in Hispanic Americans. In *Diabetes in Americans: Diabetes Data Compiled 1984*. Washington, DC, U.S. Govt. Printing Office, 1985 (NIH publ. no. 85–1468), p. IX–IV
9. Stern MP, Gaskill SP, Hazuda HP, Gardner LL, Haffner SM: Does obesity explain excess prevalence of diabetes among Mexican Americans? Results of the San Antonio Heart Study. *Diabetologia* 24:272–277, 1983
10. Stern MP, Rosenthal M, Haffner SM, Hazuda HP, Franco LJ: Sex difference in the effects of sociocultural status on diabetes and cardiovascular risk factors in Mexican Americans. *Am J Epidemiol* 120:834–851, 1984
11. Hamman RF, Mayer EJ, Moo-Young GA, Hildebrandt W, Marshall J, Baxter J: Prevalence and risk factors for diabetic retinopathy in non-Hispanic whites and Hispanics with NIDDM: San Luis Valley Diabetes Study. *Diabetes* 38:231–237, 1989
12. Perez-Stable EJ, McMillan MM, Harris M, Juarez RZ, Knowler WC, Stern MP, Haynes SG: Self-reported diabetes in Mexican Americans: HHANES 1982–1984. *Am J Public Health* 79:770–772, 1989
13. Hanis CL, Ferrell RE, Barton SA, Aguilar L, Garza-Ibarra A, Tulloch BR, Garcia CA, Schull WJ: Diabetes among Mexican-Americans in Starr County, Texas. *Am J Epidemiol* 118:659–672, 1983
14. Stern MP, Haffner SM: Type II diabetes and complications in Mexican-Americans. *Diabetes Metab Rev* 6:29–45, 1990
15. Hanis CL, Hewett-Emmett D, Bertin TK, Schull WJ: Origins of U.S. Hispanics: implications for diabetes. *Diabetes Care* 14: 618–627, 1991
16. Fujimoto WY, Leonetti DL, Kinyoun JL, Newell-Morris L, Shuman WP, Stolov WC, Wahl P: Prevalence of diabetes mellitus and impaired glucose tolerance among second-generation Japanese-American men. *Diabetes* 36:721–729, 1987
17. Fujimoto WY, Leonetti DL, Bergstom RL, Kinyoun JL, Stolov WC, Wahl PW: Glucose intolerance and diabetic complications among Japanese-American women. *Diabetes Res Clin Prac* 13:119–130, 1991
18. Cowie CC, Harris MI, Silverman RE, Johnson EW, Rust KF: Effect of multiple risk factors on differences between blacks and whites in the prevalence of non-insulin-dependent diabetes mellitus in the United States. *Am J Epidemiol* 137:719–732, 1993
19. Rotter JI, Vadheim CM, Rimoin DL: Diabetes mellitus. In *The Genetic Basis of Common Disease*. King RA, Rotter JI, Motulsky AG, Eds. New York, Oxford Univ. Press, 1992, p. 413–481
20. Hanson RL, Elston RC, Pettitt DJ, Bennett PH, Knowler WC: Segregation analysis of non-insulin-dependent diabetes mellitus in Pima Indians: evidence for a major-gene effect. *Am J Hum Genet* 57:160–170, 1995
21. Sievers ML, Fisher JR: Diabetes in North America Indians. In *Diabetes in America: Diabetes Data Compiled 1984*. Washington, DC, U.S. Govt. Printing Office, 1985 (NIH publ. no. 85–1468)
22. Bogardus C: Insulin resistance in the pathogenesis of NIDDM in Pima Indians. *Diabetes Care* 16:228–231, 1993
23. Robbins DC, Blix PM, Rubenstein AH, Kanazawa Y, Kosoka K, Tager HS: A human proinsulin variant at argenine 65. *Nature* 291:679–681, 1981
24. Tager HS: Lilly Lecture 1983: abnormal products of the human insulin gene. *Diabetes* 33:693–699, 1984
25. Oohashi H, Ohgawara H, Nanjo K, Tasaka Y, Cao Q-P, Chan SJ, Rubenstein AH, Steiner DF, Omori Y: Familial hyperproinsulinemia associated with NIDDM: a case study. *Diabetes Care* 16:1340–1346, 1993
26. Kadowaki T, Bevins E, Cama A, Ojama K, Marcus-Samuels B, Kadowaki H, Beitz L, McKeon C, Taylor SI: Two mutant alleles of the insulin receptor gene in a patient with extreme insulin resistance. *Science* 240:787–790, 1988
27. Vionnet N, Stoffel M, Takeda J, Yasuda K, Bell GI, Zouali H, Lesage S, Velho G, Iris F, Passa P, Froguel P, Cohen D: Nonsense mutation in the glucokinase gene causes early-onset non-insulin-dependent diabetes mellitus. *Nature* 356:721–722, 1992
28. Van den Ouweland JMW, Lemkes HHPJ, Ruitenbeck W, Sandkuijl LA, de Vijlder MF, Struyvenberg PAA, van de Kamp JJP, Maassen JA: Mutation in mitochondrial tRNA Leu(UUR) gene in a large pedigree with maternally transmitted type II diabetes mellitus and deafness. *Nature Genet* 1:368–371, 1992
29. Ballinger SW, Shoffner JM, Hedaya EV, Trounce I, Polak MA, Koontz DA, Wallace DC: Maternally transmitted diabetes and deafness associated with a 10.4 kb mitochondrial DNA deletion. *Nature Genet* 1:11–15, 1992
30. Kadowaki T, Kadowaki H, Mori Y, Tobe K, Sakuta R, Suzuki Y, Tanabe Y, Sakura H, Awata T, Goto Y-I, Hayakawa T, Matsuoka K, Kawamori R, Kamada T, Horai S, Nonaka I, Hagura R, Akanuma Y, Yazaki Y: A subtype of diabetes mellitus associated with a mutation of mitochondrial DNA. *N Engl J Med* 330:962–968, 1994
31. Reardon W, Ross RJM, Sweeney MG, Luxon LM, Pembrey ME, Harding AE, Trembath RC: Diabetes mellitus associated with a pathogenic point mutation in mitochondrial DNA. *Lancet* 340:1376–1379, 1992
32. National Diabetes Data Group: Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 28:1039–1057, 1979
33. Sribney WM, Swift M: Power of sib-pair and sib-trio linkage analysis with assortative mating and multiple disease loci. *Am J Hum Genet* 51:773–784,1992
34. Walker A, Cudworth AG: Type I (insulin-dependent) diabetic multiplex families: mode of genetic transmission. *Diabetes* 29:1036–1039, 1980
35. Cudworth AG, Woodrow JC: Evidence for HL-A-linked genes in "juvenile" diabetes mellitus. *Br Med J* 3:133–135, 1975
36. Hanis CL, Lawson KS, Boerwinkle E: Searching the genome for NIDDM genes: power of affected sibling pairs (Abstract). *Diabetes* 44 (Suppl. 1):41A, 1995
37. Amos CI, Elston RC, Wilson AF, Bailey-Wilson JE: A more powerful robust sib-pair test of linkage for quantitative traits. *Gen Epidemiol* 6:435–449, 1989
38. Boerwinkle E, Leffert CC, Lin J, Lackner C, Chiesa G, Hobbs HH: Apolipoprotein (a) gene accounts for greater than 90% of the variation in plasma lipoprotein (a) concentrations. *J Clin Invest* 90:52–60, 1992

39. Warden CH, Daluiski A, Bu X, Purcell-Huynh DA, DeMeester C, Shieh BH, Puppione DL, Gray RM, Reaven GM, Chen Y-DI, Rotter JI, Lusis AJ: Evidence for linkage of the apolipoprotein A-II locus to plasma apolipoprotein A-II and free fatty acid levels in mice and humans. *Proc Nat Acad Sci USA* 90:10886–10890,1993
40. Beard JC, Bergman RN, Ward WK, Porte D Jr: The insulin sensitivity index in man: correlation between clamp-derived and IVGTT-derived values. *Diabetes* 35:362–369, 1986
41. Prigeon RL, Kahn SE, Porte D Jr: Reliability of error estimates obtained with the minimal model: implications for measurements in physiological studies. *Am J Physiol* 266:E279–E286, 1993
42. Bergman RN, Ider YZ, Bowden CR, Cobelli C: Quantitative estimation of insulin sensitivity. *Am J Physiol* 236:E667–E677, 1979
43. DeFronzo RA, Tobin JD, Andres R: Glucose clamp technique: a method for quantifying insulin secretion and resistance. *Am J Physiol* 237:E214–E223, 1979
44. McGuire E, Helderman J, Tobin J, Andres R, Berman M: Effects of arterial versus venous sampling on analysis of glucose kinetics in man. *J Appl Physiol* 41:565–573, 1976
45. Steel R: Influences of glucose loading and of injected insulin on hepatic glucose output. *Ann NY Acad Sci* 82:420–430, 1959