Mapping Genes for NIDDM

Design of the Finland–United States Investigation of NIDDM Genetics (FUSION) Study

OBJECTIVE — To map and identify susceptibility genes for NIDDM and for the intermediate quantitative traits associated with NIDDM.

RESEARCH DESIGN AND METHODS — We describe the methodology and sample of the Finland–United States Investigation of NIDDM Genetics (FUSION) study. The whole genome search approach is being applied in studies of several different ethnic groups to locate susceptibility genes for NIDDM. Detailed description of the study materials and designs of such studies are important, particularly when comparing the findings in these studies and when combining different data sets.

RESULTS — Using a careful selection strategy, we have ascertained 495 families with confirmed NIDDM in at least two siblings and no history of IDDM among the first-degree relatives. These families were chosen from more than 22,000 NIDDM patients, representative of patients with NIDDM in the Finnish population. In a subset of families, a spouse and offspring were sampled, and they participated in a frequently sampled intravenous glucose tolerance test (FSIGT) analyzed with the Minimal Model. An FSIGT was completed successfully for at least two nondiabetic offspring in 156 families with a confirmed nondiabetic spouse and no history of IDDM in first-degree relatives.

CONCLUSIONS — Our work demonstrates the feasibility of collecting a large number of affected sib-pair families with NIDDM to provide data that will enable a whole genome search approach, including linkage analysis.
Mapping genes for NIDDM

insulin response and insulin resistance, we have chosen the Finnish population as the target for our study. In this paper, we describe the rationale, design, and sample characteristics of the Finland—United States Investigation of NIDDM Genetics (FUSION) study.

RESEARCH DESIGN AND METHODS

Family ascertainment criteria

Relatives of individuals who were diagnosed with NIDDM at an early age have a higher disease risk than relatives of individuals with a later onset of disease (18). However, there is a relatively greater risk of IDDM among the younger age-groups. Therefore, to maximize the genetic component while minimizing the likelihood of identifying subjects with IDDM, we ascertained probands in whom diabetes was diagnosed between 35 and 60 years of age. A given family was eligible for participation in the FUSION study if 1) the proband or another affected sibling was diagnosed with NIDDM between 35 and 60 years of age, 2) there was no history of IDDM in first-degree relatives, 3) the proband had one or more full siblings diagnosed with NIDDM at any age, and 4) at least one parent was apparently nondiabetic, with preference given to families with living parents or parents who had lived a long life without known diabetes. The last criterion was included to minimize bilineality. In this paper, we use the term index case to indicate the first proband or affected sibling in each family who met our study inclusion criteria.

Sources and types of families

There are approximately 150,000 known NIDDM patients in Finland. The probands in the FUSION study were identified primarily from the National Hospital Discharge Registry (NHDR), which includes records since 1970 of all hospitalized patients with diabetes, and from previous studies carried out by the National Public Health Institute in Finland. From the NHDR, we identified all patients who were hospitalized with a diagnosis of NIDDM in Finland during 1987–1993.

A total of 33,029 potential probands were initially identified in 1994, 29,807 of them from the NHDR. A total of 22,312 screening questionnaires were sent to the NIDDM patients living relatively close to one of the 21 clinics we established for this study. The patients were also sent a short description of study eligibility criteria, including our minimum requirement of a diabetic sib pair. A total of 9,626 individuals returned the screening questionnaire. In addition, 14 eligible families voluntarily contacted us and were included in the study.

Families in the FUSION study were classified as nuclear or extended. In the nuclear families, all the affected siblings and any living parents were studied. At least one affected sib pair (ASP) was required in each family. In the extended families, in addition to the nuclear family members, we studied a nondiabetic spouse and at least two offspring of an affected sibling, fulfilling the proband criteria described above.

Control subjects

The primary purpose of the control subjects in the FUSION study was to provide allele frequency estimates in unaffected individuals for comparison with affected individuals. All our control subjects were born in the year 1925 and had been examined with oral glucose tolerance tests (OGTTs) in 1990 as part of the South Finland Aging Study (19). Those who had normoglycemic OGTT values by the World Health Organization’s (WHO) criteria were given a second OGTT as part of the FUSION study in 1995, and the 248 individuals with normal glucose tolerance according to both OGTTs were included as control subjects for the FUSION study.

Diabetes status

The diagnosis of diabetes was accepted if an individual was receiving drug treatment for diabetes or if the blood glucose values obtained from the medical records met current WHO criteria (20). Fasting blood glucose was determined in all individuals at the FUSION examination. If patients were treated with diet alone and fasting blood glucose was <7.0 mmol/l, the diagnosis of diabetes was confirmed by OGTT. The age at diagnosis of NIDDM was based on self-reported information. The inferred diabetes status of parents was based primarily on information reported by the affected siblings because nearly all their parents were deceased.

Cases of probable IDDM were identified to exclude them from future analyses. An individual was classified as having probable adult-onset IDDM if 1) insulin treatment was started within 10 years of disease diagnosis, autoantibodies to GAD were detected, and the fasting C-peptide was <0.30 nmol/l; or if 2) insulin treatment was started within 4 years of diagnosis and the fasting C-peptide was <0.30 nmol/l. Medical record review was carried out for all affected individuals whose C-peptide was <0.50 nmol/l, who had antibodies to GAD, or who were treated with insulin and for whom C-peptide and/or GAD antibody samples were not assayed in the FUSION study to confirm their age of diagnosis and the age when insulin treatment was started.

Study clinics and examinations

The FUSION study was carried out in 21 cities distributed throughout much of Finland. The study nurses from each clinic attended a 1-week training seminar in Helsinki. Blood collection and all clinical measurements were carefully standardized.

All study subjects underwent a screening examination during a single visit. In addition, nondiabetic spouses and offspring returned for a frequently sampled intravenous glucose tolerance test (FSIGT) during a second visit. At the screening examination, blood pressure, pulse, weight, height, and waist and hip circumferences were measured on all participating family members and control subjects. Also, data on medical and lifestyle history, medication, current health status, physical activity, diet, smoking history, alcohol consumption, gestational medical history, and birth weight were collected from all family members and control subjects.

Blood samples for DNA, fasting glucose and insulin, and serum lipids were drawn for all participating family members. In addition, OGTTs were carried out in nondiabetic family members. OGTTs were performed according to WHO criteria (75-g oral load of glucose). Serum C-peptide and autoantibodies to GAD were assayed in the affected family members. An overnight urine sample was collected from each diabetic family member to assess urinary albumin-to-creatinine ratio.

To butamide-modified FSIGTs were performed in nondiabetic spouses and nondiabetic offspring during a second visit in six clinics. The reduced sampling protocol, in which 14 blood samples for glucose and insulin were collected over the course of a 3-h test (21), was employed. FSIGT data were analyzed, using the Minimal Model method (22), to derive quantitative measures of insulin sensitivity (S) and glucose effectiveness (Sg). Insulin secretion was assumed as the acute insulin response...
to glucose (AIR_{g}) calculated as described by Ward et al. (23).

Assays
Blood glucose concentration was initially measured at the study clinic using a glucose monitoring device (One Touch; LifeScan, Milpitas, CA). Plasma glucose, plasma insulin, and serum C-peptide samples were frozen and shipped on dry ice to the central laboratory at the University of Southern California in Los Angeles, California, for measurement. Plasma glucose was measured on an autoanalyzer (YSI-2700; YSI, Yellow Springs, OH) using the glucose oxidase method. Interassay coefficients of variation (CVs) were 2.9 and 2.3% for the low and high quality control (QC) pools, respectively. The correlation between One Touch and plasma glucose values was 0.94 and 0.96 for fasting glucose and postload glucose values, respectively. Plasma insulin was measured by radioimmunoassay (RIA) using dextran-charcoal separation (24). Antibody and tracer for the insulin RIA were purchased from Novo Nordisk (Bagsvaerd, Denmark). Interassay CVs were 11 and 13% for low and high QC pools, respectively. Serum C-peptide was determined by using a kit purchased from Linco (St. Charles, MO) with interassay CVs of 14 and 5% for low and high QC pools, respectively.

Autoantibodies to GAD were measured at the University of Washington in Seattle, Washington, by fluid-phase immunoprecipitation assay (25) using recombinant human GAD65 radiolabeled by in vitro translation (TNT; Promega, Madison, WI) in the presence of [35S]methionine (Amersham, Arlington Heights, IL) (26). The assay had high sensitivity and 100% specificity in an international GAD antibody workshop (27) and an interassay CV of 15%.

Serum total and HDL cholesterol and triglycerides were determined in Helsinki from fresh samples using an enzymatic assay method (CHOD-PAP; Monotest; Boehringer Mannheim, Indianapolis, IN). The interassay CVs for total cholesterol, HDL cholesterol, and triglycerides were 1.2, 2.2, and 1.6%, respectively. Urine creatinine and albumin concentrations were also measured at the National Public Health Institute in Helsinki. The creatinine in urine was determined using the method of Jaffe (28), and the urinary albumin concentration was determined by the immunoturbidimetric method, using reagents from Roche (Montclair, NJ) (29). Interassay CVs for urine creatinine and albumin were 2.4 and 5.1%, respectively.

DNA was isolated from 30 ml of fresh whole blood in EDTA using the Purgene D-50K DNA isolation kit (Genta Systems, Plymouth, MN) at the National Public Health Institute. The isolated samples were frozen and shipped on dry ice to the National Human Genome Research Institute in Bethesda, Maryland.

Genotyping
The genetic markers being used for genetic mapping in the FUSION study are modified ABI sets (Applied Biosystems Division of Perkin-Elmer, Foster City, CA) consisting of approximately 365 dinucleotide repeats divided into 28 panels, each of which contains 7 to 19 markers generating a 10-cm (cm) average density map. Incorrect genotypes can arise from polymorphism in the allele sizes, which can lead to inflated error rates or reduced power to detect linkage. Thus, a genotyping system that has a low frequency of errors is a particularly important tool in the study of complex diseases, in which gene effects may often be weak. Because the commercially based software available from ABI is not very efficient at automatically adjusting for gel-to-gel variation in allele sizing, we developed software to do these tasks automatically. In addition, we optimized our genotyping system for precise sizing and binning of alleles (30). The automation also allows us to run a series of internal checks to improve the binning accuracy and reproducibility. The blinded presence of samples in duplicate allowed us to estimate a genotype-specific error rate of 0.034% (2 of 5,846) for the last 64 markers typed in our laboratory.

Because of previous findings suggesting an association between HLA haplotypes and NIDDM, HLA-A, -B, -Bw, -DR and -DQ genotypes were also determined serologically using 180 HLA antisera (120 class I and 60 class II) (31) in 110 extended families in Helsinki.

Statistical analysis plan
The variety of phenotypic data, together with the uncertain etiology of NIDDM and its associated quantitative traits, requires a multifaceted statistical analysis strategy. In the initial phase of genome screening for NIDDM, we will employ primarily on mode-of-inheritance-free methods of linkage analysis that compare estimated allele sharing identical by descent (IBD) in the ASPs to expectations under the hypothesis that no disease locus is present (32–34). By parameterizing our models in terms of probability of IBD sharing in the ASPs, we will be able both to search for linkage and to build exclusion maps of regions that appear unlikely to harbor NIDDM genes of specified levels of genetic effect. Similar approaches can be taken for the NIDDM-associated quantitative traits using a regression or variance components framework (35, 36).

In the initial genome scan, we will follow up even rather modest suggestions of linkage by typing additional markers in the region. Our simulations (32) suggest that if significant linkage is found in a genome scan of 300–400 markers and provides good power to detect loci of even modest effect. For regions identified as likely to harbor a disease gene, we will also employ mode-of-inheritance–based methods of linkage analysis as well as combined segregation and linkage analysis in hopes of more accurately assessing the existence and the effect of the putative disease loci. In addition, we will emphasize linkage disequilibrium mapping as a complementary strategy of disease gene localization, particularly once initial evidence for linkage is obtained.

RESULTS
Family ascertainment
Among the 9,626 respondents to the screening questionnaire, there were 1,835 potential probands with onset of diabetes between 35 and 60 years of age who had at least one affected sibling and not both parents known to be diabetic. These potential probands were ranked according to their family information, with preference given to families with evidence of at least one nongenetic parent and multiple early-onset affected siblings. A total of 1,382 potential probands living near one of the 21 FUSION study clinics were contacted by phone to confirm their eligibility and willingness to participate. Of these, 470 potential probands or their affected siblings were unable or unwilling to participate. An additional 286 diabetic patients were found to be ineligible during the first interview. Thus, from the list of potential probands and volunteers who contacted us, we enrolled 640 affected individuals from 577 separate families in the FUSION study. In addition, three families were enrolled in the study after the index case refused to participate.
Characteristics of the study subjects
All index cases, siblings, and control subjects were born in Finland. Of the 577 studied index cases, diabetes status could not be confirmed in 7 individuals by fasting plasma glucose, OGTT, or review of the medical records. In addition, 20 index cases were classified as having probable adult-onset IDDM, leaving 550 index cases meeting FUSION study criteria for NIDDM (Table 1). A total of 733 siblings of the index cases with self-reported history of NIDDM were studied, and 638 of these siblings met FUSION study criteria for NIDDM. A total of 2,376 individuals were included in the study. Of the index cases and affected siblings, 55 and 45% were men, respectively. The mean age of onset of diabetes was 49 years (range, 25–69) in male index cases and 51 years (30–72) in female index cases. The age of onset for affected siblings was 53 (29–74) and 54 (22–80) years in men and women, respectively (Table 1).

In the FUSION study, we have investigated a total of 580 Finnish families, ranging in size from 1 to 13 members. Of this total, 533 families have at least two siblings affected with diabetes, resulting in a total of 795 ASPs (Table 2). In 495 of these 533 families, no cases of probable adult-onset IDDM were identified among the individuals examined in the FUSION study or among the first-degree relatives of the index cases or affected siblings, resulting in a total of 740 ASPs.

Of the index cases and affected siblings, 55 and 45% were men, respectively. The mean age of onset of diabetes was 49 years (range, 25–69) in male index cases and 51 years (30–72) in female index cases. The age of onset for affected siblings was 53 (29–74) and 54 (22–80) years in men and women, respectively (Table 1).

In the FUSION study, we have investigated a total of 580 Finnish families, ranging in size from 1 to 13 members. Of this total, 533 families have at least two siblings affected with diabetes, resulting in a total of 795 ASPs (Table 2). In 495 of these 533 families, no cases of probable adult-onset IDDM were identified among the individuals examined in the FUSION study or among the first-degree relatives of the index cases or affected siblings, resulting in a total of 740 ASPs.

Offspring were studied in 210 extended families, and FSIGTs were carried out in 79% of these families. Altogether, 602 FSIGTs were successfully done for nonaffected family members in families with no IDDM cases. An FSIGT was completed for at least two nondiabetic offspring in 156 families with a nondiabetic spouse confirmed by OGTT (Table 2).

Of the affected index cases and affected siblings, 55 and 45% were men, respectively. The mean age of onset of diabetes was 49 years (range, 25–69) in male index cases and 51 years (30–72) in female index cases. The age of onset for affected siblings was 53 (29–74) and 54 (22–80) years in men and women, respectively (Table 3).

Of the index cases and siblings with NIDDM, 12% were treated with diet alone, 46% with oral hypoglycemic agents (OHAs), 21% with insulin alone, and 20% with a combination of insulin and OHAs. The distribution of treatment modes corresponds well with that seen in Finnish diabetes clinics in general (37).

There were no notable differences in the means of the weight-related variables between affected index cases and siblings within either sex (Table 3). Of the male index cases and siblings, 40% were clearly overweight (BMI >30 kg/m²); 51% of the female cases and siblings met NIDDM were overweight. The control subjects were leaner: only 18% of men and 24% of women were clearly overweight.

Of the index cases and their affected siblings, 7% had a fasting C-peptide value of ≤0.30 nmol/l. Of these diabetic subjects, 42% were classified as having probable adult-onset IDDM. Antibodies to GAD were detected in 3.2% of the affected index cases and in 4.1% of the affected siblings. After excluding the subjects classified as having probable IDDM, 10 (1.9%) of the 517 affected index cases and 13 (2.2%) of the 594 affected siblings showed GAD antibody positivity. In normoglycemic control subjects, GAD autoantibodies were detected in 2 (1.7%) of the 120 studied individuals. The mean fasting C-peptide values between 1.5 and 1.7 nmol/l for affected index cases and siblings (Table 4), respectively, are typical for individuals with NIDDM.

There were no notable differences between index cases and siblings in serum lipid and blood pressure values (Table 5). HDL cholesterol levels were considerably higher, and triglyceride levels were lower, in control subjects than in affected individuals. Based on the questionnaire data, 60% of the male index cases and 51% of male

### Table 1—Diabetes affection status in all individuals

<table>
<thead>
<tr>
<th>Individuals</th>
<th>Confirmed</th>
<th>Probable</th>
<th>Nonaffected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studied</td>
<td>NIDDM</td>
<td>adult-onset IDDM</td>
<td></td>
</tr>
<tr>
<td>Index case</td>
<td>577</td>
<td>550</td>
<td>20</td>
</tr>
<tr>
<td>Sibling</td>
<td>733</td>
<td>638</td>
<td>22</td>
</tr>
<tr>
<td>Father</td>
<td>10</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Mother</td>
<td>32</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>Spouse</td>
<td>221</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Offspring</td>
<td>555</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Control</td>
<td>248</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>2,376</td>
<td>1,240</td>
<td>43</td>
</tr>
</tbody>
</table>

Data are n.

### Table 2—Distribution of the family sizes in the FUSION study

<table>
<thead>
<tr>
<th>Studied family members (n)</th>
<th>Total number of families</th>
<th>Families with ≥2 siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>All family members</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of families</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected siblings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of families</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After excluding families</td>
<td></td>
<td></td>
</tr>
<tr>
<td>with IDDM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offspring in extended</td>
<td></td>
<td></td>
</tr>
<tr>
<td>families</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of families</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offspring with sampled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSIGT†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of families in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>which spouse also</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sampled</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse not sampled</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are n. Control subjects were not included in these data. *Three index cases refused to participate after the family was enrolled; additional affected siblings met study criteria for proband in these families. †In families with no IDDM cases.
affected siblings had hypertension; 25% of the male control subjects reported hypertension. In women, the contrast between affected individuals and control subjects was even greater: 65 and 60 vs. 24% for female index cases, affected siblings, and control subjects, respectively.

The Minimal Model–derived FSIGT values for nondiabetic spouses and offspring are shown in Table 6. Nondiabetic women were generally more insulin sensitive than nondiabetic men. In men, this finding was partly compensated for by a more robust acute insulin response. However, the mean disposition index remained higher in females than in males. Also, the nonaffected spouses were more insulin resistant than their offspring.

### CONCLUSIONS

In a study of the determinants of a complex genetic disease
Mapping genes for NIDDM

Table 5—Serum lipid and blood pressure values in studied subjects in families with no IDDM cases

<table>
<thead>
<tr>
<th></th>
<th>Serum total cholesterol (mmol/l)</th>
<th>Serum HDL cholesterol (mmol/l)</th>
<th>Serum triglycerides (mmol/l)</th>
<th>Systolic blood pressure (mmHg)</th>
<th>Diastolic blood pressure (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected index cases</td>
<td>5.7 ± 1.3 (5.5)</td>
<td>1.03 ± 0.29 (1.00)</td>
<td>2.6 ± 2.3 (2.0)</td>
<td>152 ± 21 (150)</td>
<td>87 ± 11 (88)</td>
</tr>
<tr>
<td>Affected siblings</td>
<td>5.6 ± 1.3 (5.6)</td>
<td>1.02 ± 0.27 (1.00)</td>
<td>2.5 ± 2.6 (1.9)</td>
<td>150 ± 22 (148)</td>
<td>85 ± 12 (84)</td>
</tr>
<tr>
<td>Affected fathers</td>
<td>4.6 ± 1.7 (5.6)</td>
<td>1.12 ± 0.36 (1.11)</td>
<td>1.1 ± 0.6 (1.2)</td>
<td>151 ± 8 (152)</td>
<td>84 ± 5 (82)</td>
</tr>
<tr>
<td>Nonaffected fathers</td>
<td>5.5 ± 0.7 (5.7)</td>
<td>1.06 ± 0.14 (1.11)</td>
<td>1.4 ± 0.2 (1.5)</td>
<td>145 ± 30 (146)</td>
<td>83 ± 17 (85)</td>
</tr>
<tr>
<td>Nonaffected spouses</td>
<td>5.5 ± 0.8 (5.5)</td>
<td>1.12 ± 0.30 (1.07)</td>
<td>1.6 ± 0.9 (1.4)</td>
<td>150 ± 20 (148)</td>
<td>85 ± 11 (84)</td>
</tr>
<tr>
<td>Nonaffected offspring</td>
<td>5.3 ± 1.0 (5.2)</td>
<td>1.19 ± 0.29 (1.16)</td>
<td>1.6 ± 1.0 (1.2)</td>
<td>130 ± 13 (130)</td>
<td>82 ± 10 (80)</td>
</tr>
<tr>
<td>Controls subjects</td>
<td>5.7 ± 0.9 (5.6)</td>
<td>1.28 ± 0.29 (1.26)</td>
<td>1.5 ± 0.7 (1.4)</td>
<td>149 ± 18 (150)</td>
<td>88 ± 9 (88)</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affected index cases</td>
<td>5.9 ± 1.3 (5.8)</td>
<td>1.15 ± 0.32 (1.08)</td>
<td>2.5 ± 1.7 (2.1)</td>
<td>156 ± 24 (154)</td>
<td>83 ± 11 (82)</td>
</tr>
<tr>
<td>Affected siblings</td>
<td>5.9 ± 1.1 (5.8)</td>
<td>1.16 ± 0.30 (1.12)</td>
<td>2.3 ± 1.6 (1.8)</td>
<td>157 ± 25 (154)</td>
<td>84 ± 11 (84)</td>
</tr>
<tr>
<td>Affected mothers</td>
<td>5.7 ± 1.1 (5.9)</td>
<td>1.17 ± 0.25 (1.14)</td>
<td>2.4 ± 1.8 (1.8)</td>
<td>151 ± 20 (160)</td>
<td>79 ± 16 (80)</td>
</tr>
<tr>
<td>Nonaffected mothers</td>
<td>6.0 ± 1.4 (6.3)</td>
<td>1.38 ± 0.44 (1.36)</td>
<td>1.4 ± 0.4 (1.4)</td>
<td>143 ± 18 (134)</td>
<td>80 ± 6 (80)</td>
</tr>
<tr>
<td>Nonaffected spouses</td>
<td>6.1 ± 1.1 (6.0)</td>
<td>1.43 ± 0.35 (1.36)</td>
<td>1.4 ± 0.6 (1.2)</td>
<td>146 ± 22 (142)</td>
<td>86 ± 11 (86)</td>
</tr>
<tr>
<td>Nonaffected offspring</td>
<td>4.9 ± 0.9 (4.9)</td>
<td>1.36 ± 0.30 (1.34)</td>
<td>1.1 ± 0.5 (0.9)</td>
<td>123 ± 14 (120)</td>
<td>77 ± 10 (78)</td>
</tr>
<tr>
<td>Control subjects</td>
<td>6.4 ± 1.0 (6.4)</td>
<td>1.54 ± 0.30 (1.55)</td>
<td>1.4 ± 0.6 (1.2)</td>
<td>147 ± 16 (148)</td>
<td>86 ± 10 (86)</td>
</tr>
</tbody>
</table>

Data are means ± SD (median).

such as NIDDM, careful attention should be paid to all elements of study design. These elements include selection of the study population, ascertainment and exclusion criteria for probands and families, and the choice of phenotypes and genetic markers for study. Currently, there are several ongoing efforts to map the susceptibility genes for NIDDM using the whole genome search approach (38–42) and many others using the candidate-gene approach. To permit valid comparison of results from different studies and to assess the meaning of consistency or inconsistency of these results, proper understanding of study design, study methods, and family samples is of critical importance. In this report, we have described the sampling design and study methods we are using, and the sample of families obtained, in the FUSION study.

Advantages and relevance of a genetic study of NIDDM in Finland. There are several advantages in using a Finnish population for carrying out genetic studies. The vast majority of Finland’s 5.1 million inhabitants belong linguistically and ethnically to the Finno-Ugrian ethnic group with some admixture with the Swedish population in the coastal areas of the country. Finns are closely related to other populations of northern Europe (43), and it has been estimated that the bottleneck in the Finnish population dates back 4,000 years (44). During the first centuries after the end of migration in about A.D. 800, linguistic, geographic, and cultural barriers kept the Finns isolated from the surrounding areas (45). This small founding population and relative homogeneity increases the likelihood that individual genes predisposing to NIDDM can be identified, as has been demonstrated for genes responsible for several monogenic hereditary diseases (46–48). Relative genetic homogeneity is likely to have resulted in fewer NIDDM genes segregating than in outbred populations such as those of the U.S. or the U.K., so that individual NIDDM-susceptibility genes should be easier to detect. This distinction is important, because the detection of genes for a complex disease generally requires large samples of families, and the power to detect linkage is inversely related to the magnitude of the effect of the gene on disease risk.

NIDDM in Finland is clinically similar to NIDDM in other European populations, and the prevalence of NIDDM in Finland is similar to that of most other European populations (49). Thus, it is reasonable to expect that NIDDM susceptibility genes identified

Table 6—FSIGT variables in subjects in families with no IDDM cases

<table>
<thead>
<tr>
<th></th>
<th>Subjects (n)</th>
<th>AIRG (µU/ml × 10 min)</th>
<th>Sg (10^{-5} min^{-1} -1^{-1})</th>
<th>Sg (× 100 min^{-1})</th>
<th>Disposition index (Sg × AIRG)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonaffected spouses</td>
<td>49</td>
<td>239 ± 193 (169)</td>
<td>4.6 ± 2.6 (4.3)</td>
<td>1.6 ± 0.5 (1.5)</td>
<td>987 ± 789 (817)</td>
</tr>
<tr>
<td>Nonaffected offspring</td>
<td>233</td>
<td>225 ± 174 (189)</td>
<td>7.3 ± 4.4 (6.5)</td>
<td>1.6 ± 0.5 (1.6)</td>
<td>1340 ± 890 (1,158)</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonaffected spouses</td>
<td>108</td>
<td>235 ± 156 (198)</td>
<td>6.4 ± 3.5 (5.8)</td>
<td>1.7 ± 0.6 (1.6)</td>
<td>1348 ± 913 (1,131)</td>
</tr>
<tr>
<td>Nonaffected offspring</td>
<td>212</td>
<td>206 ± 125 (177)</td>
<td>7.9 ± 4.6 (7.0)</td>
<td>1.9 ± 0.6 (1.8)</td>
<td>1465 ± 942 (1,327)</td>
</tr>
</tbody>
</table>

Data are means ± SD (median). AIRG, acute insulin response to glucose; Sg, glucose effectiveness; Sg, insulin sensitivity.
Previous studies
Until recently, genetic studies of NIDDM have focused on the candidate-gene approach. Such studies in Finland have included investigations of the HLA system on chromosome 6 (50), and genes for glycogen synthase (51), glucokinase (52,53), hexokinase II (54,55), GLUT4 (56), fatty acid binding protein (57), apolipoprotein D (58), the insulin receptor substrate 1 (59,60), and the \( \beta_{2} \)-adrenergic receptor (61). Except for a very significant association between certain HLA haplotypes and NIDDM (50), these studies have not revealed major positive findings.

Recently, results from two whole genome searches for NIDDM have been reported, one based on 330 Mexican-American ASP families from Starr County, Texas (38), the second on 26 moderate-sized Finnish families from western Finland (39). Hanis et al. (38) reported evidence for linkage in several chromosomal regions, with the strongest evidence at the anonymous marker D2S125 on chromosome 2q, and they were able to exclude the existence of a single major locus for NIDDM throughout most of the genome. Mahtani et al. (39) failed to find compelling evidence for linkage anywhere in the genome with their complete sample. However, when they restricted attention to the six families with the lowest mean 30-min insulin level during an OGTT among the affected individuals, they did find evidence for linkage to the MODY3 region of chromosome 12q.

Rationale for a whole genome search
We have chosen to carry out a whole genome search for genes for NIDDM with linkage analysis using a map with an average resolution of 10 cM. This approach has the advantage of assaying the entire genome. The alternative candidate-gene approach is attractive because of its low cost. However, an approach based solely on candidate genes is unlikely to detect all or even most NIDDM susceptibility loci, because many such genes may be novel with no presently known homologs.

Once a candidate region is identified by linkage analysis, association analysis to identify linkage disequilibrium may help narrow the region of interest. At that stage, identification of candidate genes within the region may prove useful, an approach that has been called positional candidate cloning (62).

Approaches to maximize power to detect linkage
We have taken several steps intended to maximize power to detect NIDDM genes. First, we have restricted family ascertainment to probands with the relatively early age of diagnosis of clinical diabetes of 35–60 years, because NIDDM is more strongly familial in relatives of early-onset probands (18) and environmental factors are likely to have greater effect at older ages.

Second, we have chosen a sampling design and analytic approach based on ASPs. Compared with the alternative of extended pedigrees, ASPs are easier to ascertain. Furthermore, analysis of ASPs does not require ad hoc specification of mode of inheritance. Instead, allele sharing between affected siblings is compared with expectations assuming no linkage, so that excess allele sharing suggests linkage.

Third, we have sought to minimize the number of families in which both parents are affected with NIDDM. Although such bilineal families can provide linkage information, for most genetic models, linkage information is substantially less than that provided by families with no affected parents or with one affected parent (63). To minimize bilineality, we gave preference to families with living parents and to those families in which one or both parents had lived a long life without known diabetes. Despite these efforts, it is likely that some of our families are bilineal, because only 42 parents in 577 families were available for a study visit, and because in one-sixth of the FUSION study families, neither parent had lived past the age of 61 years without known diabetes.

Fourth, we sought to avoid families with IDDM. Our decision was based on a wish to minimize heterogeneity of the clinical phenotype. During the natural course of NIDDM, the endogenous insulin secretion diminishes, often leading to a requirement for exogenous insulin (64,65). This situation can make it difficult to distinguish NIDDM from adult-onset IDDM, especially with the current trend in NIDDM therapy toward more aggressive treatment with insulin (66). The magnitude of the problem is suggested by results from the Second National Health and Nutrition Examination Survey (NHANES II), in which 7.4% of all diabetic patients were classified as having adult-onset IDDM (67). We have carefully phenotyped affected individuals to identify those with likely IDDM. We used a low serum C-peptide level as a marker for low endogenous insulin secretion, and antibodies to GAD as a marker for immune process, suggesting probable adult-onset IDDM. While other markers such as ICA and ICA512 antibodies may be less prevalent in older-onset IDDM, GAD antibody prevalence remains high in this age-group (68). Together with thorough medical record review, we classified 43 affected individuals in 39 FUSION families as probable adult-onset IDDM. These families will be excluded from most analyses.

Fifth, we have carried out extensive phenotyping in the families. Many cases of NIDDM are asymptomatic and can only be detected by testing glucose tolerance. In Finland, the proportion of such undetected cases of all NIDDM patients is estimated to be 25–40% (69,70). Thus, as much as possible, family members of NIDDM patients were tested for glucose tolerance to confirm their glucose tolerance status. In addition, we carried out FSIGTs together with Minimal Model analysis (22,71) in the unaffected offspring and spouses of the index case or of an affected sibling in a subset of our families. This approach estimates the individual physiological functions that determine the ability of the human organism to dispose of carbohydrate nutrients, including insulin sensitivity, glucose effectiveness, and acute insulin response. Knowledge of these phenotypes provides the opportunity to separately map the responsible genes for each component of the glucose homeostatic system. Perhaps even more important is that it will provide information to help assess the roles of the genes localized by linkage analysis of NIDDM.

In our study, we found that the nonaffected offspring of NIDDM patients were more insulin sensitive than nonaffected spouses of the patients. There are several factors that may explain this finding. 1) Insulin sensitivity may decline with age (72), although not all studies show the association between biological age and insulin resistance when body weight and distribution of body fat has been taken into account (73,74). In the study by Chen et al. (72), 57- to 82-year-old men had 63% lower insulin sensitivity than 18- to 36-year-old men matched for body weight and fasting glucose. In our study, male spouses had 37% lower insulin sensitivity, and female spouses 19% lower insulin sensitivity, than the offspring. 2) Obesity, abdominal obesity, and hyperglycemia are all well-known factors associated with insulin resistance (73,75). The spouses in our study were more obese than the offspring. Also, waist-to-hip ratios were higher, indicating a higher degree of
Mapping genes for NIDDM

abdominal obesity in spouses. Moreover, spouses had slightly higher glucose values in the OGTT than their offspring. 3) Of the spouses, 25% reported diabetes in their parents, suggesting a genetic predisposition to diabetes in some of the unaffected spouses, with the proportion being approximately the same as expected in the Finnish general population (49,69,70).

Sixth, we have sampled a large number of ASP families. Our power calculations suggest that even 400 ASPs typed on a 10-cM map provide 82% power to detect a gene that confers a 1.8-fold excess risk to siblings of affected individuals at an LOD score of 3.0 (32). Even for a 1.4-fold excess risk, there is 81% power to detect genes at an LOD score of 1.0. Our large sample should also provide an excellent basis for extending our linkage and association results reported by other groups.

Limitations of the sampling design
Several of the choices we made to maximize power to detect NIDDM genes may have limited the set of NIDDM genes that we will be able to detect. By choosing a relatively homogeneous population such as the Finns, we may have limited the number of segregating genes available for detection. By selecting early diagnosed probands primarily from the NHDHR, we are preferentially sampling more severely affected individuals. If the genes for early, more severe NIDDM are different from those for later, less severe disease, we will have increased our power to detect one set of genes at the expense of the other. We believe that increased power to identify a subset of NIDDM genes more than offsets loss of the theoretical opportunity to detect all such genes. By requiring the presence of an ASP, we might be identifying genes for familial NIDDM rather than for all NIDDM. However, given the high degree of familiality of NIDDM in all populations that have been studied, this problem does not appear to be significant, particularly because genes for familial disease often turn out to have a more general importance (76–78).

Several of these choices also made it more difficult to identify and recruit acceptable families. In Finland, only 3% of symptomatic NIDDM patients have their disease onset before the age of 40 years, and only 37% have it before the age of 60 years (64,79). Our requirement for a second affected sibling and zero or only one parent known to be affected, and our exclusion of families with IDDM, further restricted the pool of suitable families. Thus, screening more than 22,000 NIDDM patients resulted in only about 1,550 individuals (7%) meeting our study criteria. Even a lower proportion of suitable families was reported from the English study of Cook et al. (80), in which primary screening criteria were less stringent than ours. Their conclusion was to initiate large-scale collaborative multicenter studies to collect the family material necessary for the study of susceptibility genes for NIDDM. In contrast, the FUSION study has demonstrated that it is feasible, although not easy, to collect a large number of ASP families with NIDDM within one population even after applying careful selection criteria.

Future directions
Given our success in family collection and phenotypic analysis in the FUSION study, we are moving forward with FUSION 2, in which we hope both to extend our current FUSION families and to obtain additional families. In all our current FUSION families, we will type unaffected siblings to allow more accurate inference of identity by descent in the ASPs for purposes of linkage analysis. In our current FUSION nuclear families, we will sample a spouse and offspring of one of the affected siblings for purposes of haplotype determination and linkage disequilibrium mapping. We also plan to identify and sample a second set of Finnish families as a replication set to allow us to immediately test interesting chromosomal regions in a second, comparable set of families. In addition, we will obtain a sample of parent-offspring trios, in which the offspring is affected with NIDDM and the parents are living, to allow family-based association studies (81).

Even given a large initial sample of families and a second replication sample, detection of NIDDM susceptibility genes by linkage analysis will depend on a combination of hard work and good luck. Most critical will be the identification of the various susceptibility genes on disease risk. With a combined sample of perhaps 800 ASPs, we would have 69% power to detect a locus with the impact estimated by Hanis et al. (38) for their chromosome 2q locus at an LOD score of 3.0. However, for a locus that confers a 1.2-fold excess risk to siblings, even 800 ASPs would be expected to provide an LOD score of only 1.59 (32). Thus, in genetic studies of NIDDM and other common, complex genetic diseases, sharing results with other workers in the field and combining data from comparable samples will be critical to our success in identifying disease susceptibility genes.

Acknowledgments—This project was made possible by intramural funds from the National Human Genome Research Institute (NHGRI) (OH95-C-N030) and by R01 HG00376. Family studies were approved by the Institutional Review Board at the National Institutes of Health (SPA S-5737-05) and at the National Public Health Institute in Helsinki, Finland. R.M.W. and C.D.L. were supported by training grant T32 HG00040 from the NIH; R.M.W. is currently supported by individual postdoctoral award F32 DK09525. We are grateful to the National Public Health Institute for funding the lipid assays and HLA typing. The work of Paula Nyholm, Jouko Sundvall, Tuula Tenkula, and Sanemila Vilkki is appreciated. We also thank Upjohn Pharmaceuticals for kindly donating tolbutamide (Orinase Diagnostic) for the FSIGTs.

We gratefully acknowledge the excellent work of Gunther Birznieks, Peter Chines, William Eldridge, Zarir Karanjawala, Julie Knapp, Colin Martin, Anjenee Musick, Tiffany Musick, Joseph Rayman, Shane Shapiro, Alistair So, Joyce Tannenbaum, Catherine Te, Ray Whiten, and Alyson Witt at NHGRI, and Edward Trager and Peggy White at the University of Michigan.

References


45. Hästbacka J, de la Chapelle A, Källial, Sis-
Mapping genes for NIDDM


