Two recent publications reported association of common polymorphisms in the P2 promoter of hepatocyte nuclear factor 4α (HNF4α) (the MODY1 gene) with risk for type 2 diabetes. We attempted to reproduce this putative association by genotyping 11 single nucleotide polymorphisms (SNPs) spanning the HNF4α coding region and the P2 promoter in >3,400 patients and control subjects from Sweden, Finland, and Canada. One SNP that was consistently associated in the two previous reports (rs1884613, in the P2 promoter region) also trended in the same direction in our sample, albeit with a lower estimated odds ratio (OR) of 1.11 (P = 0.05, one-tailed). We genotyped this SNP (rs1884613) in an additional 4,400 subjects from North America and Poland. In this sample, the association was not confirmed and trended in the opposite direction (OR 0.88). Meta-analysis of our combined sample of 7,883 people (three times larger than the two initial reports combined) yielded an OR of 0.97 (P = 0.27). Finally, we provide an updated analysis of haplotype structure in the region to guide any further investigation of common variation in HNF4α. Although our combined results fail to replicate the previously reported association of common variants in HNF4α with risk for type 2 diabetes, we cannot exclude an effect smaller than that originally proposed, heterogeneity among samples, variation in as-yet-unmeasured genotypic or environmental modifiers, or true association secondary to linkage disequilibrium (LD) with as-yet-undiscovered variant(s) in the region. 

Type 2 diabetes is extremely common, affecting 5–10% of people in the industrialized world. Twin studies document the role of inherited genetic factors, whereas secular trends demonstrate the powerful role of environmental and behavioral influences. In the search for genetic factors, both monogenic and polygenic forms of the disease are recognized, with only a few percentage of cases attributable to monogenic forms and the rest presumed to be attributable to a more complex cause. Maturity-onset diabetes of the young (MODY) is an autosomal dominant form of type 2 diabetes, characterized by early onset and a defect in the function of the pancreatic β-cells. Six genes are known to cause MODY (1–6), and mutations in these genes are thought to act by causing abnormal regulation of transcription in β-cells, creating a defect in metabolic signaling of insulin secretion and/or β-cell mass (7). Another single-gene defect that causes type 2 diabetes is maternally inherited diabetes and deafness, caused by mutations in the mitochondrial DNA (8). Although the molecular identification of the MODY genes has illuminated the pathophysiology of diabetes, inherited variation in these genes has not previously been demonstrated to contribute in a causal manner to the common, late-onset form of type 2 diabetes.

Attempts to identify genes that are involved in the common form of type 2 diabetes have relied on genome-wide linkage studies and candidate gene association studies. Although genetic association studies for complex traits have been challenged by irreproducibility (9), several common genetic variants now have been widely replicated as influencing risk for type 2 diabetes, including the Pro12Ala polymorphism in the peroxisome proliferator-activated receptor-γ, the E23K polymorphism in the Kir6.2 gene (both reviewed in 10,11), and variation in the region of Calpain10 (12,13). For each of these genes, multiple large studies...
independently observed association under a consistent model of genotype-phenotype correlation, meta-analysis of all published studies gave a significant positive result, and subsequent large studies replicated the model proposed by the initial meta-analysis. Experience with these and other complex disease genes has led to a model in which common variants may have modest effects, such that very large study populations and substantial statistical significance (much greater than the typically used thresholds of \( P < 0.05 \) or \( P < 0.01 \)) are required to achieve reproducible results (11,14–18).

It is not yet clear how often the same genes that cause monogenic forms of common diseases (e.g., MODY) also harbor common variants of more modest effects that contribute to late-onset forms of disease. Recently, two groups independently obtained evidence suggesting that common variants near the P2 promoter of hepatocyte nuclear factor-4α (HNF4α), the MODY1 gene, are associated with risk for type 2 diabetes (19,20). The gene for HNF4α spans 29 kb on chromosome 20q13.1–13.2 (21). The P2 promoter, located ~46 kb upstream of the P1 promoter and the coding start site, controls the predominantly expressed splice form in the β-cell (22,23). A number of linkage studies have observed suggestive linkage peaks in the region of HNF4α, further highlighting this region for study (24–32).

In the study by Love-Gregory et al. (20), the strongest associations were for a single nucleotide polymorphism (SNP) near the P2 promoter (rs1884614; OR 1.45; nominal \( P = 0.008 \)) and another in a 3' intron (rs3818247; OR 1.49; nominal \( P = 0.003 \)). Ten SNPs in the Silander et al. study (19) were nominally associated (below threshold \( P = 0.05 \)), the strongest of which was rs1884613, also near the P2 promoter (OR 1.34; \( P = 0.01 \)). Each study included ~1,200 case patients and control subjects and genotyped 16 and 13 SNPs, respectively.

The similarity in results between the Love-Gregory et al. (20) and Silander et al. (19) studies and the high previous probability of a gene involved in MODY diabetes and in a region of putative linkage all support the hypothesis that variation in HNF4α might influence risk for the common form of type 2 diabetes. Nevertheless, given the history of nonreplication in genetic association studies (9,33), all such hypotheses need to be explored for consistency in additional samples. We studied SNPs in HNF4α in an additional 7,883 patients (approximately three times the combined study sample of Love-Gregory et al. [20] and Silander et al. [19]) to extend knowledge of genotype-phenotype correlation at this important locus.

**TABLE 1**

Clinical characteristics of patient samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Type</th>
<th>Sex (M/F)</th>
<th>Age (years)</th>
<th>BMI (kg/m²)</th>
<th>Fasting plasma glucose (mmol/l)</th>
<th>2-h OGTT PG (mmol/l) or HbA₁c (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discordant sibs</td>
<td>Diabetes/severe IGT sib</td>
<td>280/329</td>
<td>65 ± 10</td>
<td>29 ± 5</td>
<td>9.3 ± 3.3</td>
<td>14.3 ± 5.6</td>
</tr>
<tr>
<td></td>
<td>NGT sib</td>
<td>275/305</td>
<td>62 ± 10</td>
<td>26 ± 3</td>
<td>5.4 ± 0.4</td>
<td>6.0 ± 1.1</td>
</tr>
<tr>
<td>Scandinavia</td>
<td>Diabetes/severe IGT</td>
<td>252/219</td>
<td>60 ± 10</td>
<td>28 ± 5</td>
<td>9.8 ± 3.4</td>
<td>15.0 ± 5.3</td>
</tr>
<tr>
<td></td>
<td>NGT</td>
<td>254/217</td>
<td>60 ± 10</td>
<td>27 ± 4</td>
<td>6.2 ± 1.8</td>
<td>6.8 ± 2.8</td>
</tr>
<tr>
<td>Canada</td>
<td>Diabetes</td>
<td>70/57</td>
<td>53 ± 8</td>
<td>29 ± 5</td>
<td>6.4 ± 1.8</td>
<td>12.8 ± 2.1</td>
</tr>
<tr>
<td></td>
<td>NGT</td>
<td>70/57</td>
<td>58 ± 8</td>
<td>29 ± 4</td>
<td>5.1 ± 0.6</td>
<td>6.1 ± 1.1</td>
</tr>
<tr>
<td>Sweden</td>
<td>Diabetes/severe IGT</td>
<td>267/247</td>
<td>66 ± 12</td>
<td>28 ± 4</td>
<td>8.5 ± 2.5</td>
<td>6.5 ± 1.5</td>
</tr>
<tr>
<td></td>
<td>NGT</td>
<td>267/247</td>
<td>66 ± 12</td>
<td>28 ± 4</td>
<td>4.8 ± 0.7</td>
<td>ND</td>
</tr>
<tr>
<td>GCI U.S.</td>
<td>Diabetes</td>
<td>644/582</td>
<td>63 ± 11</td>
<td>33 ± 7</td>
<td>9.8 ± 3.0</td>
<td>8.0 ± 3.1</td>
</tr>
<tr>
<td></td>
<td>NGT</td>
<td>644/582</td>
<td>61 ± 10</td>
<td>27 ± 5</td>
<td>5.1 ± 0.9</td>
<td>ND</td>
</tr>
<tr>
<td>GCI Poland</td>
<td>Diabetes</td>
<td>422/387</td>
<td>62 ± 10</td>
<td>30 ± 5</td>
<td>8.9 ± 4.0</td>
<td>7.9 ± 1.3</td>
</tr>
<tr>
<td></td>
<td>NGT</td>
<td>422/387</td>
<td>59 ± 7</td>
<td>26 ± 4</td>
<td>4.8 ± 1.2</td>
<td>ND</td>
</tr>
</tbody>
</table>

Data are presented as means ± SD. Plasma glucose (PG) was measured at baseline (fasting) and 2 h after an oral glucose tolerance test (OGTT). IGT, impaired glucose tolerance; NGT, normal glucose tolerance; ND, not determined.

**RESEARCH DESIGN AND METHODS**

The characteristics of four of our patient samples have been described elsewhere (11). They include 1,189 Scandinavian siblings who were discordant for type 2 diabetes; two Scandinavian case-control samples that contained 942 and 1,028 subjects, respectively; and 254 subjects from the Sagnenay Lac-St. Jean region in Quebec. The case-control samples were individually matched for age, BMI, and geographic region. The type 2 diabetic patients met World Health Organization 1998 criteria for type 2 diabetes. Severe impaired glucose tolerance was defined as 10.0 mmol/l >120 min, blood glucose ≥8.5 mmol/l. An oral glucose tolerance test was performed when fasting plasma glucose was <11 mmol/l. Age of onset is available for 641 individuals in this study, with a mean ± SD of 53.0 ± 11.7 years.

This study also includes analysis of two additional case-control samples from Genomics Collaborative, Inc. (GCI), which have previously been described (34). The first sample contains 2,452 individuals of U.S. white ancestry, and the second includes 2,018 individuals from Poland. To qualify to be enrolled in the GCI diabetes study, participants were required to meet American Diabetes Association criteria for definition of type 2 diabetes and to be treated by the enrolling physician for diabetes at the time of study entry. Control subjects are healthy individuals who have no known history of chronic diseases and are required to have had a normal fasting plasma glucose <7 mmol/l within the last year. Case patients and control subjects were matched one-to-one on the basis of sex, age (>2 years), and self-reported ethnicity. For this study, both case patients and control subjects were required to report two parents and four grandparents with the same self-reported ethnicity. In addition, we matched on country of birth for the participant, both parents, and all four grandparents as much as possible in the case-control matching. The phenotypic characteristics of all samples are described in Table 1.

**Genotyping.** Genotyping was performed as previously described by primer extension of multiplex products with detection by matrix-assisted laser desorption ionization time of flight mass spectroscopy using a Sequenom platform (35,36). The average genotype completeness for working assays was 96%. The genotyping consensus error was determined to be 0.25%, using both duplicate genotypes (2,761 comparisons) and errors in Mendelian inheritance.

**Statistical analysis.** To determine the allelic association of each particular SNP with type 2 diabetes, we used simple \( \chi^2 \) analysis in the case-control samples and the Discordant Allele Test in the sibling pairs (37). Results for the subsamples were combined using Mantel-Haenszel meta-analysis of the odds ratios (ORs) (33). Power calculations were performed using the program of Purcell et al. (38), available at http://statgen.iop.kcl.ac.uk/gpc. Homogeneity among studies was tested using a Pearson \( \chi^2 \) goodness-of-fit as previously described (33).
TABLE 2
Association of HNF4α tag SNPs with type 2 diabetes in 3,413 patients

<table>
<thead>
<tr>
<th>SNP</th>
<th>Minor allele</th>
<th>OR</th>
<th>P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs412828*</td>
<td>C</td>
<td>0.99</td>
<td>0.42</td>
<td>0.88–1.11</td>
</tr>
<tr>
<td>rs1884613*</td>
<td>G</td>
<td>1.11</td>
<td>0.05</td>
<td>0.97–1.27</td>
</tr>
<tr>
<td>rs2144908*</td>
<td>A</td>
<td>1.09</td>
<td>0.12</td>
<td>0.94–1.25</td>
</tr>
<tr>
<td>rs4364072</td>
<td>G</td>
<td>0.98</td>
<td>0.71</td>
<td>0.87–1.10</td>
</tr>
<tr>
<td>rs3092370</td>
<td>A</td>
<td>0.97</td>
<td>0.62</td>
<td>0.88–1.08</td>
</tr>
<tr>
<td>rs1800963*</td>
<td>A</td>
<td>1.01</td>
<td>0.92</td>
<td>0.90–1.13</td>
</tr>
<tr>
<td>rs3212184</td>
<td>G</td>
<td>1.02</td>
<td>0.74</td>
<td>0.91–1.14</td>
</tr>
<tr>
<td>rs1885088*</td>
<td>A</td>
<td>1.05</td>
<td>0.22</td>
<td>0.94–1.20</td>
</tr>
<tr>
<td>rs1800961</td>
<td>T</td>
<td>0.92</td>
<td>0.58</td>
<td>0.68–1.23</td>
</tr>
<tr>
<td>rs10258583*</td>
<td>G</td>
<td>0.92</td>
<td>0.15</td>
<td>0.82–1.03</td>
</tr>
<tr>
<td>rs3818247*</td>
<td>G</td>
<td>0.98</td>
<td>0.35</td>
<td>0.87–1.10</td>
</tr>
</tbody>
</table>

*SNP was also genotyped by both the Love-Gregory et al. (20) and Silander et al. (19) studies. †SNP was found to have a significant association to type 2 diabetes in at least one previous study. One-tailed P values are given for the SNPs with the dagger symbol (?); all others are two-tailed.

Haplotype structure. To evaluate the haplotype structure of the HNF4α region, we genotyped 119 SNPs from dbSNP in a multigenerational panel of 12 Centre d’Etude du Polymorphisme Humain (CEPH) pedigrees that contained 96 chromosomes. These SNPs span 108 kb, from ~17 kb upstream of the P2 promoter to ~16 kb downstream of the end of the HNF4α 3’ UTR. SNPs were initially selected on the basis of an evenly spaced grid across the region, with additional SNPs added on the basis of the extent of linkage disequilibrium (LD). Nineteen (16%) of the SNPs attempted were technical failures (failing either Hardy Weinberg equilibrium or to attain a 75% genotyping percentage), and 42 (42%) of the remaining 100 SNPs either were monomorphic in this population or had a minor allele frequency (<5%), totaling a final set of 58 working, high-frequency SNPs. The average spacing between these 58 SNPs is 1.9 kb, with the largest interval being 12.2 kb (in a region of strong LD). Haplotype blocks were determined as described previously (11).

SNP evaluation. To determine how selected SNPs for our study and the Love-Gregory et al. (20) and Silander et al. (19) studies captured variation in the HNF4α region, we evaluated all genotyped SNPs for their correlation to one another in the CEPH samples described above. Specifically, we recorded the maximal pairwise r2 value of each tag SNP to the complete set of other variants typed in the region. On the hypothesis that any of these variants could be a putative causal variant or proxy thereof, we report the fraction of all variants typed in the region. On the hypothesis that any of these variants could be a putative causal variant or proxy thereof, we report the fraction of all such sites; at a density of one SNP per 1.9 kb, most (but not all) of the tag SNPs typed in the patient sample. We note that this is a non-conservative estimate of power, because we have not typed all SNPs in the region but only the one per 1.9 kb found in dbSNP. Because the total number of SNPs >5% frequency is only approximately one per 500 bp on average across the human genome (39,40), our survey likely has captured approximately one-fourth of all such sites; at a density of one SNP per 1.9 kb, most (but not all) of the remaining sites will typically show a high r2 value to one of the sites already observed (P.d.B., M.J.D., and D.A., unpublished observations).

RESULTS
We began by genotyping 11 SNPs in 3,413 type 2 diabetic patients and matched control subjects from Scandinavia and Canada previously described (11,18), (Table 1). The markers were selected before the publication of Love-Gregory et al. (20) and Silander et al. (19) but were found to include each of the SNPs significantly associated in both of the previous studies or a proxy thereof with r2 > 0.9 (online appendix [available at http://diabetesjournals.org]) in an initial panel of 96 CEPH chromosomes. In addition to the SNPs that were relevant to the two previous papers, we genotyped the common missense polymorphism T130I and seven other SNPs from the HNF4α region. The single SNP association results are presented in Table 2. For previously associated SNPs, we report one-tailed P values (with the P value corresponding to an OR in the same direction as the previous report); for SNPs that were not previously published, two-tailed P values are reported.

Ten of the 11 genotyped SNPs had a nominal P > 0.1, including the missense SNP previously associated in another report (41). However, SNP rs1884613 near the P2 promoter (the strongest result in the Silander et al. study [19] and the second best in the Love-Gregory et al. [20] analysis) showed nominally significant evidence for replication, with a one-sided P = 0.05 uncorrected for the multiple tests performed. The OR in our samples was considerably lower than that in the two previous reports (1.11 vs. 1.34 and 1.38). The previous studies used family-based samples (some of which showed some evidence for linkage in this region), and such genetic loading could potentially account for the higher OR. In our family-based subsample, many of which were multiplex, the observed OR was 1.09, similar to our Scandinavian unrelated case-control sample (OR 1.11).

Because these results were consistent with the previous results but not definitive, we genotyped the most consistently associated SNP, rs1884613, in an additional sample of 4,470 patients (type 2 diabetic patients and control subjects) from North America and Poland. These patient samples have been previously published (34,42) and replicate both the previously associated peroxisome proliferator-activated receptor-γ P12A SNP (42) and the Kir6.2 E23K variant (34). Association testing of the single SNP in this additional sample did not support the previous model but trended in the opposite direction from previously described studies (OR 0.88). When we performed a meta-analysis of the 7,883 patients who were studied in this report, the estimated OR for rs1884613 was 0.97 (P = 0.27; Table 3).

Failure to replicate a previous result could possibly be due to lack of power for replication or heterogeneity between the original patient sample and the one used for replication. Power calculations (38) indicate that our combined sample of 7,883 individuals provides 98% power (at a significance level of 0.05) to detect an OR of 1.34 as proposed for rs1884613 by Silander et al. (19). Of course, the OR in an initial report is often overestimated as a result of the so-called “winner’s curse”; under the scenario that the true effect (if one exists) is smaller than in the original reports, our power is less than these estimates suggest (11,18,42).

TABLE 3
Association of SNP rs1884613 with type 2 diabetes in individual samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Frequency (case/control)</th>
<th>OR</th>
<th>P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discordant sibs</td>
<td>0.19/0.18</td>
<td>1.09</td>
<td>0.89</td>
<td>0.75–1.59</td>
</tr>
<tr>
<td>Canada</td>
<td>0.12/0.17</td>
<td>0.67</td>
<td>0.25</td>
<td>0.41–1.12</td>
</tr>
<tr>
<td>Scandinavia</td>
<td>0.20/0.19</td>
<td>1.11</td>
<td>0.73</td>
<td>0.80–1.39</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.19/0.15</td>
<td>1.25</td>
<td>0.12</td>
<td>0.99–1.58</td>
</tr>
<tr>
<td>GCI U.S.</td>
<td>0.15/0.17</td>
<td>0.85</td>
<td>0.07</td>
<td>0.73–0.99</td>
</tr>
<tr>
<td>GCI Poland</td>
<td>0.17/0.18</td>
<td>0.93</td>
<td>0.66</td>
<td>0.79–1.09</td>
</tr>
<tr>
<td>Combined</td>
<td><strong>0.97</strong></td>
<td><strong>0.59</strong></td>
<td><strong>0.90–1.06</strong></td>
<td></td>
</tr>
</tbody>
</table>

Results from the Scandinavian and Canadian subsamples shown in Table 2 are listed individually here for rs1884613. The U.S. and Polish samples are from GCI. Frequencies are given for the minor (G) allele. Results from all samples are combined by Mantel-Haenszel meta-analysis. All P values are two tailed.
When we considered the similar results obtained in both the Silander et al. (19) and our Scandinavian samples, we wondered whether heterogeneity across samples (as a result of ethnicity or geography) could be responsible for the lack of replication in the North American/Polish samples. A statistical test for heterogeneity in the OR of rs1884613 across all of our subsamples demonstrated borderline evidence for heterogeneity.
that we have typed (one per 1.9 kb) likely represents approximately one-fourth of all of the common (>5%) SNPs in a typical region of the genome (approximately one per 500 bp) (39,40).

All three groups’ SNP sets performed very similarly (Table 4); this is not surprising, as the three groups used very similar sets of SNPs on the basis of what was available in dbSNP 1 year or more ago. We found that none provided particularly complete coverage of the current LD map. Specifically, the average correlation coefficient ($r^2$) of the genotyped tag SNPs to the remaining markers in the LD map averaged only 0.51–0.54, and only 28–35% of typed SNPs were captured with an $r^2 > 0.8$. This means that the studies of HNF4α performed to date (including the current study) are as yet far from complete in their evaluation of all of the currently known common variants in the region for association with diabetes risk. For groups that are interested in doing a more comprehensive study, a set of tag SNPs selected using this updated haplotype map is given in the online appendix. Finally, we note that the extent of LD is considerably less at this locus than in the genome as a whole (http://www.hapmap.org/) (36).

### DISCUSSION

Given 25,000 or more genes in the human genome, the a priori probability that any given gene carries mutations that influence a trait of interest is very low. Where a gene is already known to play a functional role in the biological process of interest or, better still, is causative in a monogenic form of the disease (as in HNF4α), the a priori expectation of a causal role clearly rises. By any measure, the two recently published association results relating common variants in HNF4α to risk for type 2 diabetes are particularly intriguing. Severe mutations in the gene clearly cause a monogenic form of type 2 diabetes, and at least one SNP showed similar and nominally significant results that were consistent across two studies. However, both studies were only modestly significant, and so only if the previous probability for such variants in HNF4α were very high would the chance of a true positive be greater than that of a false positive (43). Also weighing in the equation is the lack of a demonstrated functional consequence of the putative associated alleles, although this may be a common situation for complex genetic diseases.

To try to demonstrate the reproducibility of the genotype-phenotype correlation, we genotyped 11 tag SNPs in 3,413 individuals and tested the single site that was most consistently associated across studies in 4,470 additional patients. Initially, we observed a modestly significant replication of association (one-tailed $P = 0.05$, not corrected for the multiple hypotheses examined) for the rs1884613 SNP, consistent in direction with previously published studies, albeit with a much weaker estimated effect size (OR 1.11, as compared with ~1.3–1.4 in Love-Gregory et al. [20] and Silander et al. [19]). The larger North American and Polish study trended in the opposite direction, however, such that meta-analysis of all of our results fails to confirm that the variation at the HNF4α P2 promoter is associated with increased risk for type 2 diabetes.

There are several possible explanations for our inability to replicate consistently the findings of Love-Gregory et al. (20) and Silander et al. (19). First, it is possible that vari-
ation at this locus does have an effect on disease risk and that our results in the U.S./Poland sample are false negative. Given 4,400 people, it is unlikely that an effect as large as in the original reports (OR 1.34) would be missed. However, it is well understood that the true effect of an association is often more modest than estimated in the original studies, as a result of the so-called “winner’s curse” (33). We estimate that ~8,000 case patients would be necessary to have 80% power to detect an OR of 1.15, and so all of the published results are consistent with a somewhat smaller effect.

Second, it is possible that the initial associations could have been statistical fluctuations, as the corrected P values in the original report were ~0.01. Although it might seem unlikely that two studies should get the same result with a P value of ~0.01, we note that a large number of such studies are being performed, and there was no a priori hypothesis that this SNP and gene would be the one to show up in these particular two studies.

A third potential explanation is that the previously reported association signals are real but that there is heterogeneity among populations, and the variant in question is not associated with risk in the U.S. and Poland samples as a result of unmeasured variation in genetic, environmental, or behavioral modifiers. For example, the Scandinavian and Canadian samples are matched for BMI, whereas the GCI samples are not. Although both study designs are valid, it is possible that the individuals in these samples could have a different mix of predisposing genetic risk factors. Moreover, although we failed to demonstrate significant evidence of heterogeneity, it is possible that with increased data, the trend toward heterogeneity (P = 0.06) might be documented as significant. For distinguishing between these hypotheses, additional large positive studies should be obtained or proof of significant heterogeneity documented.

There is general importance to the eventual outcome of genetic association testing in the genes implicated in monogenic forms of type 2 diabetes such as MODY1, the other MODY genes, and the mtDNA. During the past decade, many genes that cause Mendelian forms of common diseases have been identified. Examples include MODY, maternally inherited diabetes and deafness, >20 inherited forms of blood pressure regulation (44), early-onset breast cancer (BRCA1, BRCA2, ATM), Alzheimer’s (APP, PS1, PS2), and others. It is as yet unclear how often the genes that cause the Mendelian form of these diseases also contribute to the common form of disease. The extent to which the genes for common and rare forms of disease turn out to be overlapping will inform understanding of the diversity of the biological and evolutionary paths to a shared clinical endpoint.

ACKNOWLEDGMENTS

J.N.H. is a recipient of a Burroughs Wellcome Career Award in Biomedical Sciences. D.A. is a Charles E. Culpeper Scholar of the Rockefeller Brothers Fund and a Burroughs Wellcome Fund Clinical Scholar in Translational Research, the latter of which supported this work. L.G., T.T., and the Botnia Study are principally supported by the Sigrid Juselius Foundation, the Academy of Finland, the Finnish Diabetes Research Foundation, the Folkhalsan Research Foundation, EC (BM4-CT95-0662, GIFT), the Swedish Medical Research Council, the JDF Wallenberg Foundation, and the Novo Nordisk Foundation.

We thank the Botnia research team for clinical contributions and the members of the Altshuler, Hirschhorn, and Group laboratories for helpful discussions.

REFERENCES


38. Purcell S, Cherry SS, Sham PC: Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19:149–150, 2003


