Linkage disequilibrium mapping of the replicated type 2 diabetes linkage signal on chromosome 1q

Citation for published version:

Digital Object Identifier (DOI):
10.2337/db09-0081

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Diabetes

Publisher Rights Statement:
Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See http://creativecommons.org/licenses/by-nc-nd/3.0/ for details.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Linkage Disequilibrium Mapping of the Replicated Type 2 Diabetes Linkage Signal on Chromosome 1q

Inga Prokopenko,1,2 Eleftheria Zeggini,1,3 Robert L. Hanson,4 Braxton D. Mitchell,5 N. William Rayner,1,2 Pelin Akan,3 Leslie Baier,4 Swapna K. Das,6,7 Katherine S. Elliott,1 Mao Fu,5 Timothy M. Frayling,8 Christopher J. Groves,1,2 Rhian Gwilliam,3 Laura J. Scott,9 Benjamin F. Voight,10,11,12,13 Andrew T. Hattersley,8 Cheng Hu,14 Andrew D. Morris,15 Maggie Ng,16 Colin N.A. Palmer,17 Marcela Tello-Ruiz,18 Martine Vaxillaire,19 Cong-rong Wang,14 Lincoln Stein,18,20 Juliana Chan,16 Weiping Jia,14 Philippe Froguel,19,21 Steven C. Elbein,6 Panos Deloukas,3 Clifton Bogardus,4 Alan R. Shuldiner,5 and Mark I. McCarthy,1,2,22 for the International Type 2 Diabetes 1q Consortium

OBJECTIVE—Linkage of the chromosome 1q21–25 region to type 2 diabetes has been demonstrated in multiple ethnic groups. We performed common variant fine-mapping across a 23-Mb interval in a multiethnic sample to search for variants responsible for this linkage signal.

RESEARCH DESIGN AND METHODS—In all, 5,290 single nucleotide polymorphisms (SNPs) were successfully genotyped in 3,179 type 2 diabetes case and control subjects from eight populations with evidence of 1q linkage. Samples were ascertainment designed to enhance power to detect variants causal for 1q linkage. After imputation, we estimate ~80% coverage of common variation across the region ($r^2 > 0.8$).

From the 1Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.; the 2Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, U.K.; the 3Wellcome Trust Sanger Institute, Welcome Trust Genome Campus, Hinxton, Cambridge, U.K.; the 4Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, Arizona; the 5School of Medicine, University of Maryland, Baltimore, Maryland; the 6Endocrinology Section, Medical Service, Central Arkansas Veterans Healthcare System, Little Rock, Arkansas; the 7Division of Epidemiology and Metabolism, Department of Internal Medicine, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas; the 8Institute for Clinical and Biomedical Science, Peninsula Medical School, Exeter, U.K.; the 9Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan; the 10Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts; the 11Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts; the 12Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts; the 13Department of Medicine, Harvard Medical School, Boston, Massachusetts; the 14Shanghai Diabetes Institute, Department of Endocrinology & Metabolism, Shanghai Jiaotong University No. 6 People’s Hospital, Shanghai, China; the 15Diabetes Research Group, Biomedical Research Institute, University of Dundee, Dundee, U.K.; the 16Department of Medicine and Therapeutics, Chinese University of Hong Kong, Shatin, Hong Kong, SAR; the 17Biomedical Research Institute, Ninewells Hospital and Medical School, Dundee, U.K.; the 18Cold Spring Harbor Laboratory, New York, New York; the 19CNRS UMR 8906, Institute of Biology and Lille 2 University, Pasteur Institute, Lille, France; the 20Informatics & Biocomputing, Ontario Institute for Cancer Research, Toronto, Ontario, Canada; the 21Genomic Medicine, Hammersmith Hospital, Imperial College London, London, London, U.K.; and the 22Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Oxford, U.K. Corresponding author: Mark McCarthy, mark.mccarthy@drl.ox.ac.uk. Received 18 January 2009 and accepted 1 April 2009.

RESULTS—Association mapping of the 23-Mb region identified two strong signals, both of which were restricted to the subset of European-descent samples. The first mapped to the NOS1AP (CAPON) gene region (lead SNP: rs7538490, odds ratio 1.38 [95% CI 1.21–1.57], $P = 1.4 \times 10^{-6}$, in 999 case subjects and 1,190 control subjects); the second mapped within an extensive region of linkage disequilibrium that includes the ASH1L and PKLR genes (lead SNP: rs11264371, odds ratio 1.48 [1.18–1.76], $P = 1.0 \times 10^{-5}$, in a dominant model). However, there was no evidence for association at either signal on replication, and, across all data (>24,000 subjects), there was no indication that these variants were causally related to type 2 diabetes status.

CONCLUSIONS—Detailed fine-mapping of the 23-Mb region of replicated linkage has failed to identify common variant signals contributing to the observed signal. Future studies should focus on identification of causal alleles of lower frequency and higher penetrance.

Genome-wide association (GWA) analysis has provided a powerful stimulus to the discovery of common variants influencing type 2 diabetes risk, and, to date, ~20 susceptibility loci have been identified with high levels of statistical confidence (1). However, these known variants account for only a small proportion of the inherited component of disease risk (probably <10%), and the molecular basis of the majority of the genetic predisposition to type 2 diabetes has yet to be established (1).

The success of the GWA approach contrasts with the slow progress that characterized previous efforts to map susceptibility loci through genome-wide linkage (2). However, now that many of the common variants of largest effect have been identified (in European-descent populations at least), there are cogent reasons to revisit regions previously identified through genome-wide linkage. First, variants within the genomic intervals representing replicated linkage signals can be considered to have raised prior odds for a susceptibility effect, and this information can be used to prioritize GWA signals (particularly those with only modest evidence of association) for targeted replication. Second, genuine linkage signals are likely to be driven by causal variants—particularly low-frequency SNPs or copy number variants not captured by the com-
modity GWA arrays—with effect sizes larger than those currently detectable by GWA (3). Because alleles with these characteristics will have a more marked impact on individual disease predisposition than the common variants found by GWA, identification of causal variants underpinning replicated linkage signals should accelerate efforts to obtain better predictors of disease (4).

For type 2 diabetes, there appears to be only limited overlap between the regions identified by genome-wide linkage and those revealed by GWA (5). Although the discovery of TCF7L2 was prompted by a search for causal variants within a region of replicated type 2 diabetes linkage, neither the common variants in TCF7L2 nor those in HHEX and IDE (a second nearby GWA signal) account for that linkage signal (6). Thus, the discovery of TCF7L2 reflects either serendipity or the co-localization of common and rare causal variants in the same locus—the former driving the association and the latter the linkage. Similarly, whereas common variants in HNF4A have been reported to explain the chromosome 20 linkage signals seen in Finns and Ashkenazim (7,8), these associations have proved difficult to replicate (9).

Chromosome 1q (in particular the 30-Mb stretch adjacent to the centromere) ranks alongside the regions on chromosomes 10 and 20 as among the strongest in terms of the replicated evidence for genome-wide linkage to type 2 diabetes. Linkage has been reported in samples of European (U.K., French, Amish, Utah), East Asian (Chinese, Hong Kong), and Native American (Pima) origin (summarized in Supplementary Table 1, which is available in the online-only appendix at http://diabetes.diabetesjournals.org/cgi/content/full/db09-0081/DC1; ref. 2). The region concerned is gene rich and contains a disproportionately share of excellent biological candidates (2). The homologous region has also emerged as a diabetes susceptibility locus from mapping efforts in several well-characterized rodent models (10–13).

The International 1q Consortium represents a coordinated effort by the groups with the strongest evidence for 1q linkage to identify variants causal for that signal. Here, we report efforts to map causal variants using a custom linkage-disequilibrium (LD) mapping approach, predominantly based around common SNP variants, applied to a linkage-disequilibrium (LD) mapping approach, predominantly based around common SNP variants, applied to a well-powered set of multiethnic samples.

To improve power, ascertainment of type 2 diabetes cases for this study aimed to enrich for 1q causal alleles through 1) a focus on populations and samples that had shown 1q linkage; 2) selection for positive family history; and 3) for some samples, preferential recruitment on the basis of patterns of identity by descent sharing in the 1q region. For each set of case subjects, we selected a control sample of individuals from the same population. Details of the recruitment have been reported previously (14) and are summarized in the supplementary material available in the online appendix. In all, the case-control part of the study included 2,198 samples (1,000 case subjects, 1,198 control subjects) of European descent, 281 (140 case subjects, 141 control subjects) of East Asian origin, and 285 (144 case subjects, 141 control subjects) who were Native American (Pima) (supplementary Table 2). We also included a small sample of individuals of African American origin (242 case subjects, 173 control subjects) and an additional 599 Pima individuals (520 affected, 79 nondiabetic after age 45 years) from 255 families who, after combination with the Pima case-control set, were used for family-based association analyses.

These samples were submitted to dense-map SNP typing of the core region of interest (from 147.0 to 169.7 Mb [Build 35]) using a series of 1,536-plex BeadArray designs (Golden Gate, Illumina, San Diego, CA). Design of these arrays was contemporaneous with development of dbSNP and HapMap (15). Thus, whereas the first arrays were LD-agnostic and compiled using genomic localization as the primary consideration for inclusion, subsequent arrays used LD information from the CEU (European ancestry) and CHB/JPT (Asian ancestry) components of HapMap to guide SNP selection and maximize coverage of the region. In all, we designed assays for 6,023 SNPs, of which 5,290 provided reliable data in all populations after passing through our extensive quality control (see the supplementary material). We estimate that after imputation (using the appropriate set of HapMap data as a reference), coverage of the region (minor allele frequency >0.05; r² > 0.8) reaches ~80% in the European and ~72% in the East Asian samples. Coverage is harder to estimate (and imputation likely to be less valuable) in Pima and African American samples, since reference data from these populations are not available, although we estimate ~49% coverage in West Africans based on YRI data.

Genotyping quality was generally good, with over 91% of SNPs passing quality control in each population (see the supplementary material) and <0.45% of SNPs failing (P < 10⁻⁴) tests of within-sample Hardy-Weinberg equilibrium. Significant departures from expectation in the distributions of test statistics observed in the Amish and Pima samples (as revealed by QQ plots; see supplementary material) likely reflect residual relatedness between subjects from those populations. We adjusted for this (and any population stratification effects) through genomic control methods (16). Association analyses treated each study as a separate stratum and used standard meta-analysis approaches to deliver estimates of joint effect size and statistical significance (see supplementary material). A series of nested meta-analyses were performed including 1) European-descent samples only (“4-way”); 2) non–African-descent samples only (“7-way”); and 3) all samples (“8-way”).

Under an additive model with allele frequency of 0.25, our sample provides ~80% power to detect per-allele odds ratios (ORs) of >1.36 (“8-way”) or >1.43 (“4-way”) for α = 5 × 10⁻⁸. Given that the region covers ~1% of the genome, we consider this a reasonable benchmark for “region-wide” significance (equivalent to consensus genome-wide thresholds of 5 × 10⁻⁸). These power calculations are conservative: given the case ascertainment enrichment strategies used in this study, we would expect to detect variants with population-level effects in the 1.2–1.3 range. Under reasonable assumptions (three independent alleles contributing to a linkage signal with a locus-specific sibling relative risk of ~1.15), we can expect the effect size of the variants we were seeking to detect (i.e., those responsible for the 1q linkage) to be substantially greater than this (e.g., allelic OR 1.6 for a variant with risk allele frequency of 25%). Our study was therefore well powered to detect putatively causal alleles within the European and/or combined datasets.

Across these analyses, none of the SNPs showed an association with type 2 diabetes that withstood genome-wide correction (P < 5 × 10⁻⁸) (17). However, two clusters of SNPs showed association signals that approached or exceeded “region-wide” significance thresholds. The first of these, involving rs7538490 and nearby
SNPs, mapped to a 51.4-kb interval (at 160.35 Mb) within the first intron of NOS1AP (nitric oxide synthase 1 [neuronal] adaptor protein) with an estimated per allele OR (in the 4-way, European-only analysis) of 1.38 (95% CI 1.21–1.57, $P = 1.4 \times 10^{-6}$, additive model, Table 1, Fig. 1). Rs7538490 lies ~5.6 kb from one of the SNPs (rs10494366).

TABLE 1
<table>
<thead>
<tr>
<th>Association results for rs7538490 in the NOS1AP gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case subjects (n)</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>U.K.</td>
</tr>
<tr>
<td>French</td>
</tr>
<tr>
<td>Utah</td>
</tr>
<tr>
<td>Amish*</td>
</tr>
<tr>
<td>Meta-analysis: 4 European descent populations</td>
</tr>
<tr>
<td>Shanghai</td>
</tr>
<tr>
<td>Hong Kong</td>
</tr>
<tr>
<td>Pima*</td>
</tr>
<tr>
<td>Meta-analysis: 7 non-African populations</td>
</tr>
<tr>
<td>African American</td>
</tr>
<tr>
<td>Meta-analysis: 8 1qC populations</td>
</tr>
</tbody>
</table>

**Replication samples**
- Independent WTCCC sample
  - 1,495
    - 2,938
    - 0.29
    - 0.29
    - 0.97 (0.88–1.07)
    - 0.57
- W2C vs. 58BC
  - 472
    - 1,992
    - 0.27
    - 0.27
    - 1.00 (0.85–1.17)
    - 0.93
- UKT2DGDC
  - 3,932
    - 4,818
    - 0.29
    - 0.28
    - 1.03 (0.97–1.10)
    - 0.18
- Diabetes Genetics Initiative
  - 1,464
    - 1,467
    - 0.26
    - 0.26
    - 1.01 (0.87–1.16)
    - 0.82
- FUSION
  - 1,161
    - 1,174
    - 0.30
    - 0.24
    - 1.15 (0.97–1.39)
    - 0.063
| Meta-analysis: all replication samples | 8,524 | 12,389 | 1.03 (0.98–1.07) | 0.29 |
| Meta-analysis: all European descent populations | 9,523 | 13,579 | 1.06 (1.01–1.10) | 0.014 |
| Meta-analysis: all data | 10,049 | 14,034 | 1.05 (1.01–1.10) | 0.015 |

*P value for Amish and Pima was adjusted using estimated lambda for genomic control. WTCCC, Wellcome Trust Case Control Consortium.

**FIG. 1.** Single-point type 2 diabetes associations within the 1q region. This plot shows the “4-way” (European-descent samples only) meta-analysis using the additive model (Cochran-Armitage trend test). Directly typed SNPs are shown in orange and imputed SNPs in blue. The pale blue track (and secondary y-axis) represents recombination rates (for HapMap CEU) across the region. Blue diamonds represent the strongest association $P$ value in the two regions taken forward for replication. In the case of the FUSION region, the strongest association was seen for a dominant model (the equivalent additive model result is denoted with the red diamond). Only a small subset of genes within the region is denoted on the gene track.
previably shown to influence cardiac repolarization and QT interval (18); the two SNPs are in modest LD ($r^2 = 0.47$ in HapMap CEU), and rs10494366 shows some evidence for association with type 2 diabetes ($P = 3.1 \times 10^{-5}$) in the same 4-way meta-analysis.

The second signal includes $\sim 10$ SNPs in a 220-kb region of extensive LD at $\sim 152.1$ Mb. This region includes the coding sequences of the genes encoding liver pyruvate kinase (PKLR) and ash1 (absent, small, or homeotic)-like (Drosophila) (ASH1L) among others. At the lead SNP (rs11264371), the estimated OR for the 4-way analysis was 1.36 (1.18–1.58) ($P = 3.5 \times 10^{-5}$) under the additive model. The effect size and significance were marginally greater (1.48 [1.18–1.76], $P = 1.0 \times 10^{-5}$), under a dominant model (Table 2, Fig. 1).

Both signals were most prominent in the European samples, and there was no evidence that an equivalent association signal extended to the East Asian, Native American, or African American samples. Though the association $P$ values for these two signals remained strong in the 8-way meta-analysis of all data ($5.4 \times 10^{-5}$ for rs7538490 and $2.9 \times 10^{-4}$ for rs11264371, Tables 1 and 2), in each case they were driven by the larger European samples. Analyses in the larger Pima family-based association dataset also found no evidence of association (rs7538490, $P = 0.59$; rs11264372 $r^2$ of one with rs11264371 in CEU and CHB/JPT HapMap, $P = 0.79$).

Although neither signal was of sufficient effect size to be considered causal for the 1q linkage signal (the estimated sibling relative risk attributable to these loci in combination is only 1.045), we reasoned that these signals might nevertheless be pointers toward nearby causal variants (of lower frequency but higher penetrance) that were inadequately tagged by the SNPs we had typed. However, before proceeding to resequencing and fine-mapping, we first sought replication of our findings in independent datasets. Mindful that our case ascertainment strategies may have led to inflated estimates of effect size compared with those evident in unselected case subjects, we recognized that large sample sizes would be required to test the observed associations. Because the signals were clearest in European-descent samples, we focused replication on samples from Northern Europe.

First, we used GWA data from the Wellcome Trust Case Control Consortium (19,20). After removing 429 overlapping case subjects, we examined 1,495 additional type 2 diabetes case subjects and 2,938 control subjects with Affymetrix 500k data (using imputation to test for association at the lead SNPs in each interval). No evidence of association was evident (rs7538490, $P = 0.57$; rs11264371, $P = 0.33$). Similarly, analysis of GWA data from the Diabetes Genetics Initiative (21) and FUSION (22) studies provided no corroboration of either signal. Furthermore, when analyzed jointly (4,549 case subjects, 5,579 control subjects), these three studies also failed to reveal any additional common variant signals of interest ($P < 10^{-5}$) across the wider 1q region (23) and no corroboration of any of the lesser signals evident in the 1q consortium analyses.

---

### Table 2

**Association results for rs11264371 in the PKLR/ASH1L region**

<table>
<thead>
<tr>
<th>Country</th>
<th>Case subjects (n)</th>
<th>Control subjects (n)</th>
<th>Risk allele frequency in case subjects</th>
<th>Risk allele frequency in control subjects</th>
<th>Additive model OR (95% CI)</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.K.</td>
<td>444</td>
<td>443</td>
<td>0.25</td>
<td>0.20</td>
<td>1.33 (1.07–1.66)</td>
<td>0.012</td>
</tr>
<tr>
<td>French</td>
<td>219</td>
<td>244</td>
<td>0.25</td>
<td>0.23</td>
<td>1.17 (0.86–1.59)</td>
<td>0.32</td>
</tr>
<tr>
<td>Utah</td>
<td>190</td>
<td>162</td>
<td>0.26</td>
<td>0.21</td>
<td>1.34 (0.94–1.92)</td>
<td>0.095</td>
</tr>
<tr>
<td>Amish*</td>
<td>147</td>
<td>349</td>
<td>0.25</td>
<td>0.17</td>
<td>1.83 (1.19–2.81)</td>
<td>5.8 x 10^{-3}</td>
</tr>
<tr>
<td>Meta-analysis: 4 European descent populations</td>
<td>1,000</td>
<td>1,198</td>
<td></td>
<td></td>
<td>1.36 (1.18–1.58)</td>
<td>3.5 x 10^{-5}</td>
</tr>
<tr>
<td>Shanghai</td>
<td>77</td>
<td>77</td>
<td>0.77</td>
<td>0.70</td>
<td>1.47 (0.87–2.49)</td>
<td>0.15</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>63</td>
<td>63</td>
<td>0.66</td>
<td>0.73</td>
<td>0.73 (0.42–1.23)</td>
<td>0.24</td>
</tr>
<tr>
<td>Pima*</td>
<td>144</td>
<td>141</td>
<td>0.58</td>
<td>0.60</td>
<td>0.91 (0.64–1.29)</td>
<td>0.64</td>
</tr>
<tr>
<td>Meta-analysis: 7 non-African populations</td>
<td>1,284</td>
<td>1,479</td>
<td></td>
<td></td>
<td>1.24 (1.10–1.42)</td>
<td>5.6 x 10^{-4}</td>
</tr>
<tr>
<td>African American</td>
<td>242</td>
<td>173</td>
<td>0.37</td>
<td>0.34</td>
<td>1.18 (0.88–1.58)</td>
<td>0.25</td>
</tr>
<tr>
<td>Meta-analysis: 8 1qC populations</td>
<td>1,526</td>
<td>1,652</td>
<td></td>
<td></td>
<td>1.24 (1.10–1.39)</td>
<td>2.9 x 10^{-4}</td>
</tr>
<tr>
<td>Replication samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent WTCCC sample</td>
<td>1,495</td>
<td>2,938</td>
<td>0.24</td>
<td>0.24</td>
<td>1.05 (0.95–1.17)</td>
<td>0.33</td>
</tr>
<tr>
<td>W2C vs. 58BC</td>
<td>486</td>
<td>1,992</td>
<td>0.25</td>
<td>0.24</td>
<td>1.01 (0.86–1.19)</td>
<td>0.56</td>
</tr>
<tr>
<td>UK2DGC</td>
<td>3,922</td>
<td>4,819</td>
<td>0.24</td>
<td>0.24</td>
<td>1.04 (0.98–1.12)</td>
<td>0.38</td>
</tr>
<tr>
<td>Diabetes Genetics Initiative</td>
<td>1,464</td>
<td>1,467</td>
<td>0.26</td>
<td>0.27</td>
<td>0.97 (0.84–1.12)</td>
<td>1.00</td>
</tr>
<tr>
<td>FUSION</td>
<td>1,161</td>
<td>1,174</td>
<td>0.25</td>
<td>0.27</td>
<td>0.90 (0.78–1.03)</td>
<td>0.60</td>
</tr>
<tr>
<td>Meta-analysis: all replication samples</td>
<td>8,528</td>
<td>12,390</td>
<td></td>
<td></td>
<td>1.02 (0.97–1.06)</td>
<td>0.54</td>
</tr>
<tr>
<td>Meta-analysis: all European descent populations</td>
<td>9,528</td>
<td>13,588</td>
<td></td>
<td></td>
<td>1.04 (1.00–1.09)</td>
<td>0.083</td>
</tr>
<tr>
<td>Meta-analysis: all data</td>
<td>10,054</td>
<td>14,042</td>
<td></td>
<td></td>
<td>1.04 (1.00–1.09)</td>
<td>0.069</td>
</tr>
</tbody>
</table>

*P value for Amish and Pima was adjusted using estimated lambda for GC. 58 BC, 1958 British Birth Cohort; UKT2DGC, United Kingdom Type 2 Diabetes Genetics Consortium; W2C, Warren 2 cases; WTCCC, Wellcome Trust Case Control Consortium.
Finally, we genotyped the two lead SNPs (rs7538490, rs11264371) using fluorogenic 5'-nuclease (Taqman) assays in 4,572 case subjects and 6,941 control subjects from the U.K. (the UK Type 2 Diabetes Genetics Consortium and Warren 2 cases/1958 British Birth Cohort strata in Tables 1 and 2). Once again, there was no evidence of replication. Taking into account all replication samples (~8,500 case subjects, ~12,400 control subjects), there was no significant association with type 2 diabetes (rs7538490, OR 1.03 [95% CI 0.98–1.07], \( P = 0.29 \); rs11264371, OR 1.02 [0.97–1.06], \( P = 0.54 \) [additive], 1.06 [0.99–1.13], \( P = 0.10 \) [dominant]). Nominal significance was retained when these replication data were combined with the original 1q consortium case-control data (rs7538490, \( P = 0.015 \); rs11264371, \( P = 0.069 \)), but these associations are unimpressive in either the region-wide or genome-wide context. Even allowing for some heterogeneity of effect size between the primary and replication datasets (due to ascertainment differences and the “winner’s curse”), there seems to be no substantive evidence that the association signals observed in the NOSIAP and the PKLR/ASHIL region are genuinely associated with type 2 diabetes.

In summary, we have undertaken a detailed survey of common variants across the region of replicated 1q linkage, achieving coverage that exceeds that of available GWA data for the region. Despite analysis of multiple ethnic groups in samples sufficiently powered (in the European-descent component at least) to have detected common variants causal for the linkage, we found no compelling signals.

Should we conclude therefore that the original evidence for 1q linkage was false? Although this possibility cannot be discounted, it is worth considering that recent experience from GWA studies has shown that, for common susceptibility variants at least, effect sizes are modest and that none is of magnitude sufficient to generate a linkage signal detectable in achievable sample sizes. Efforts to explain the “missing heritability” for type 2 diabetes (that is, the disparity between the predisposition attributable to the known loci and independent estimates of overall heritability and familiality) are now shifting toward the search for low-frequency, medium-penetrance alleles. Alleles with these characteristics are likely to have escaped detection through the genome-wide approaches available so far, since they would be insufficiently penetrant to be detected with traditional linkage approaches applied to monogenic families and too infrequent to be reliably identified through GWA studies (4). Yet, low-frequency, medium-penetrance alleles could, particularly if several independent alleles map to the same locus, generate the kinds of linkage signals detectable in family-based studies (as is the case for NOD2/CARD15 and Crohn’s, for example) (24).

Detection of low-frequency susceptibility variants will require new approaches based around next-generation resequencing and large-scale fine-mapping. Genome-wide resequencing of large case-control samples remains economically and logistically unfeasible, but targeted resequencing of selected regions is not, and the future plans of the 1q consortium include deep resequencing of the 1q region of interest, focusing at least initially on exons and conserved sequence.

ACKNOWLEDGMENTS

The principal funding for this study was provided as a supplement to NIDDK through R01-DK073490 and as a supplement to U01-DK58026. Other major support was provided by the National Institutes of Health (T32-AG00219, R01-DK54261, R01-DK54261, K24-DK02673, K07-CAG7960, R01-DK39311, U01-DK58026, and intramural funds); the University of Maryland and Arkansas General Clinical Research Centers; the National Center for Research Resources (M01RR14288); the Department of Veteran Affairs and American Diabetes Association (U.S.); “200 Familles pour vaincre le Diabète et l’Obésité” and Association Française des Diabétiques (France); Diabetes U.K.; The Hong Kong Research Grants Committee (CUSH 4292/9M; I/04C), Chinese University of Hong Kong Strategic Grant Program (SRP9902) and Hong Kong Innovation and Technology Support Fund (ITSF/33/00) (Hong Kong); and The National Nature Science Foundation of China (39630150), Shanghai Medical Pioneer Development Project (96-3-004; 996024) and Shanghai Science Technology Development Foundation (01ZB14047) (China). We also recognize the funding support of the U.K. Medical Research Council (G0601201), the Oxford National Institute for Health Research Biomedical Research Centre, the General Clinical Research Centers Program, the Baltimore Veterans Administration Geriatric Research and Educa-
tional Clinical Center, and the Wellcome Trust (GR072960). E.Z. is a Wellcome Trust Research Career Development Fellow. For the 1958 Birth Cohort, venous blood collection was funded by the U.K. Medical Research Council, and cell line production and DNA extraction and processing was funded by the Juvenile Diabetes Research Foundation and the Wellcome Trust.

No potential conflicts of interest relevant to this article were reported.

We thank all the subjects participating in this study and those who contributed to collection of the clinical resources. In particular, we acknowledge members of the Diabetes Genetics Initiative and the Finland-U.S. Investigation of NIDDM Genetics (FUSION) for sharing data from their studies.

REFERENCES


