Brief Genetics Report

A Genome-Wide Search for Type 2 Diabetes Susceptibility Genes in West Africans

The Africa America Diabetes Mellitus (AADM) Study

Charles N. Rotimi,[1] Guanjie Chen,[1] Adebowale A. Adeyemo,[1,2] Paulette Furbert-Harris,[1] Debra Guass,[1] Jie Zhou,[1] Kate Berg,[3] Olufemi Adegoke,[1] Albert Amoah,[4] Samuel Owusu,[4] Joseph Acheampong,[5] Kofi Agyenim-Boateng,[5] Benjamin A. Eghan, Jr.,[5] Johnnie Oli,[6] Godfrey Okafor,[6] Ester Ofoegbu,[6] Babatunde Osotimehin,[2] Fayeofori Abbiyesuku,[2] Thomas Johnson,[7] Theresa Rufus,[7] Olufemi Fasanmade,[7] Rick Kittles,[1] Harold Daniel,[1] Yuanxiu Chen,[1] Georgia Dunston,[1] and Francis S. Collins[3]

The incidence of type 2 diabetes is growing rapidly, not only in developed countries but also worldwide. We chose to study type 2 diabetes in West Africa, where diabetes is less common than in the U.S., reasoning that in an environment where calories are less abundant, incident cases of type 2 diabetes might carry a proportionately greater genetic component. Through the Africa America Diabetes Mellitus (AADM) study, we carried out a genome-wide linkage analysis of type 2 diabetes in a cohort of 343 affected sibling pairs (691 individuals) enrolled from five West African centers in two countries (Ghana: Accra and Kumasi; Nigeria: Enugu, Ibadan, and Lagos). A total of 390 polymorphic markers were genotyped, and multipoint linkage analysis was conducted using the GENEHUNTER-PLUS and ASM programs. Suggestive evidence of linkage was observed in four regions on three chromosomes (12, 19, and 20). The two largest logarithm of odds scores of 2.63 and 1.92 for chromosomes 20q13.3 and 12q24, respectively, are particularly interesting because these regions have been reported to harbor diabetes susceptibility genes in several other populations and ethnic groups. Given the history of forced migration of West African populations during the slave trade, these results should have considerable relevance to the study of type 2 diabetes in African Americans. *Diabetes* 53: 838–841, 2004

From the [1]Department of Microbiology, National Human Genome Center at Howard University, College of Medicine, Washington, DC; the [2]College of Medicine, University of Ibadan, Ibadan, Nigeria; the [3]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland; the [4]Department of Medicine, University of Ghana Medical School, Accra, Ghana; the [5]Department of Medicine, University of Science and Technology, Kumasi, Ghana; the [6]Department of Medicine, University of Nigeria Teaching Hospital, Enugu, Nigeria; and the [7]College of Medicine, Endocrine and Metabolic Unit, University of Lagos, Lagos, Nigeria.

Address correspondence and reprint requests to Charles Rotimi, PhD, National Human Genome Center, Howard University, Genetic Epidemiology Unit, College of Medicine, 2216 6th St., NW, Washington, DC 20059. E-mail: crotimi@howard.edu.

Received for publication 21 July 2003 and accepted in revised form 21 November 2003.

AADM, Africa America Diabetes Mellitus; LOD, logarithm of odds.

Type 2 diabetes is now a serious international health risk. It is estimated that the number of adults with diabetes worldwide will rise from 135 million in 1995 to 300 million in 2025, with the majority of that increase occurring in the developing world (1).

It is clear that both genetic and environmental risk factors play critical roles in the development of type 2 diabetes. Across several populations and ethnic groups, excess body weight from the combined effects of increased caloric intake and sedentary lifestyle has been shown to be a major contributor to risk (2). In contrast, despite strong evidence of heritability, success in identification of specific genetic factors has been modest at best. Numerous candidate genes have been investigated, and while a few gene variants have been defined that confer increased risk of type 2 diabetes (3,4), many others have failed to be confirmed in independent studies. The more systematic approach of genome-wide linkage analysis has been applied in numerous studies on many different populations. Many regions of putative linkage have been identified in these studies, but many have not been replicated. A notable exception is a region on chromosome 20q, where linkage signals have been identified by multiple groups in multiple populations (5–10).

Using a panel of 390 polymorphic markers, scored by investigators at the Center for Inherited Disease Research, we conducted the first comprehensive search for susceptibility loci in a cohort of 343 West African families with at least two siblings with type 2 diabetes. Study participants (11) were enrolled in five centers located in two West African countries (Ghana: Accra and Kumasi; Nigeria: Enugu, Ibadan, and Lagos). Clinical characteristics of the affected subjects are displayed in Table 1. The mean age for the cohort was 53 years, and ~60% were women. The mean age at diagnosis was 46.5 years, and the mean duration of diabetes was 7.0 years. While obesity is generally less common in West Africa than in the U.S., this

TABLE 1
Demographic characteristics of study participants: the AADM study

	Men	Women
n	281	410
Duration of diabetes (years)	7.2 ± 6.6	6.6 ± 6.4
Age at diagnosis (years)	47.7 ± 10.8	46.3 ± 10.6
Age (years)	53.8 ± 10.5	53.0 ± 11.1
Height (cm)	169.6 ± 7.2	159.0 ± 7.00
BMI (kg/m^2)	25.0 ± 5.3	27.3 ± 5.6
Leptin (ng/ml)	4.5 ± 5.3	23.9 ± 19.9
Systolic blood pressure (mmHg)	134.0 ± 21.6	137.6 ± 23.9
Diastolic blood pressure (mmHg)	82.8 ± 11.5	83.3 ± 12.2
Fasting insulin (μU/ml)	19.6 ± 27.0	23.2 ± 32.0
Fasting glucose (mg/dl)	198.5 ± 95.6	202.8 ± 86.8
Fasting C-peptide (ng/ml)	1.09 ± 0.7	1.36 ± 0.78
Fat mass (kg)	15.9 ± 11.0	25.1 ± 11.4
Fat-free mass (kg)	56.0 ± 7.0	44.6 ± 5.8
Percent fat mass	20.6 ± 9.6	34.6 ± 9.1

Data are means ± SE, unless otherwise indicated. The total number of families is 343. All male versus female comparisons were significant at $P = 0.05$, except for diastolic blood pressure.

cohort of diabetes patients are still at least five BMI units heavier than the general populations of Nigeria and Ghana (12), confirming the widely reported association between excess weight and diabetes risk (1,2). Approximately 50% of these diabetic patients are also hypertensive.

Results of the multipoint linkage analysis from GENEHUNTER-PLUS for the 22 autosomal chromosomes are displayed in Fig. 1, and a summary is provided in Table 2 for markers reaching a logarithm of odds (LOD) >1.20. The strongest evidence of linkage was observed in the region between D20S480 and D20S171 (LOD score 2.63), extending over a 16-cM region on chromosome 20 (20q13.3).

A LOD score of 1.80 was also observed at position 74.4 cM between markers D20S481 and D20S480 on chromosome 20. The region between D12S2070 and D12S395 on chromosome 12 had the next highest LOD score of 1.92. We also found suggestive evidence of linkage at four other locations: chromosome 19 (LOD 1.81) in the region of marker D19S714, between D4S2623 and D4S2394 on chromosome 4 (LOD 1.37; at D17S1298 on chromosome 17 [LOD 1.26]), and between D20S470 and D20S477 on chromosome 20 (LOD 1.24).

Cognizant of the multiple analyses performed in this investigation, we conducted simulation studies to estimate the genome-wide significance of these LOD scores. The goal of these analyses was to obtain replicates using the same information content as the original dataset. In this regard, the same allele frequencies and map distances, as well as the same missing persons for each marker, were used in generating the 100 simulation datasets. Of simulation, 12.7% produced a LOD score ≥1.0, 5.2% produced a LOD score ≥1.5, 2.0% produced a LOD score ≥2.0, 0.8% produced a LOD score ≥2.5, and 0.4% produced a LOD score ≥3.0.

This is the first genome scan study to search for susceptibility genes for type 2 diabetes in West Africans. It is interesting that our strongest linkage signal is on the long arm of chromosome 20. At least 10 other studies have reported evidence for linkage for type 2 diabetes and/or type 2 diabetes–related traits in this region. Originally, Ji et al. (5) and Zouali et al. (6) reported LOD scores of 3.30 and 2.74 in American and French Caucasians, respectively. Subsequent reports have also observed linkage signals in Finns (7), Japanese (8), and Chinese (9). A recent update of the original linkage report raised the LOD score to 4.82 at 74 cM (10). While the 20q linkage peak with the highest

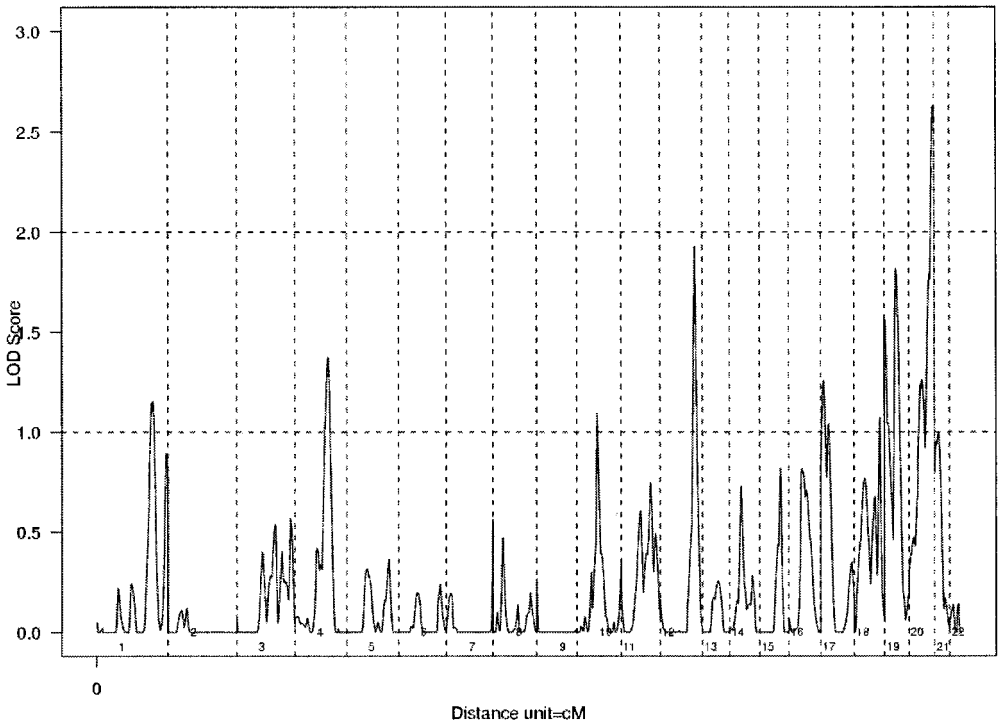

FIG. 1. Multipoint linkage analyses for genes linked to type 2 diabetes using the GENEHUNTER-PLUS and ASM programs. The *x*-axis shows distance (in centimorgans), and the *y*-axis shows the LOD score.

TABLE 2
Results of linkage analyses with type 2 diabetes: the AADM study

Chromosome	Flanking markers	Location of maximum LOD (cM)	Nearest marker	1-LOD CI (cM)	LOD score
4	D4S2623, D4S2394	124.8	D4S2394	113.8–131.8	1.37
12	D12S2070, D12S395	131	D12S2070	123.8–138.8	1.92
17	D17S1298	12.2	D17S1298	4–18.8	1.26
19	D19S714, D19S433	43.4	D19S714	36–56	1.81
20	D20S470, D20S477	44.2	D20S477	37–60	1.24
20	D20S481, D20S480	74.4	D20S480	60–78	1.80
20	D20S480, D20S171	90.8	D20S171	78–94	2.63

Map positions are in centimorgans and based on the database maintained by the Marshfield Medical Research Foundation (http://research.marshfieldclinic.org/genetics).

LOD score in this study (at 90.8 cM) is distal to what is reported in most other studies, the other linkage peaks on the same chromosome are in the same regions reported in other studies (5–10). In particular, the linkage peak at 74 cM on 20q in this study is within 1 cM of the peak reported by Klupa et al. (10) in Caucasian families. Thus, it seems likely that a diabetes gene or genes, perhaps harboring a common and ancient susceptibility variant, resides in this region.

The 12q24 linkage signal is also noteworthy in that several groups have identified evidence for a type 2 diabetes locus in this region. For example, Lindgren et al. (13) recently reported a nonparametric LOD score of 1.8 at marker D12S366 and summarized findings for several other genome scans that point to this same interval. Of note is that the *TCF1* gene, mutations for which are responsible for one type of maturity-onset diabetes of the young (MODY3), is located at 12q24 (14).

Based on these results, we plan to fine map these linkage signals with densely placed single nucleotide polymorphisms, looking for evidence of association. In the endgame of gene identification, the AADM dataset may turn out to be highly useful, as one would anticipate the region of linkage disequilibrium around a susceptibility allele to be narrower in Africans than in Europeans or Asians (15).

Type 2 diabetes affects African Americans, with an incidence about twice that of the overall population of the U.S. (16). To what extent this significant health disparity reflects diet, environment, access to health care, and cultural practices has not been determined, but it is possible that genetic factors are also involved, or perhaps even selected for in times of famine, according to the "thrifty gene hypothesis" (17). Searching for the genetic causes of type 2 diabetes in African Americans has been challenging because of the high incidence and very prominent role of obesity and sedentary lifestyle in the U.S. By focusing on the founder West African population, where incidence is much lower and lifestyle factors are less prominent, we hypothesized that affected siblings would carry a proportionately larger genetic contribution to their disease. While the LOD scores reported here do not on their own reach the level of genome-wide significance (18), the confirmation of a linkage signal on chromosome 20q is encouraging and implies that the study of this population may contribute critical clues to the understanding of this common and devastating illness.

RESEARCH DESIGN AND METHODS

The AADM study has an affected sibling pair design. Recruitment strategies and eligibility criteria for the families enrolled in this report have been described in a previous publication (11). Families were enrolled and examined from five centers in two countries in West Africa. The three centers in Nigeria (Enugu, Ibadan, and Lagos) enrolled two major ethnic groups: Ibos (28%) and Yorubas (28%). The two centers in Ghana also enrolled two major ethnic groups: Akan-Ashante (25%) and Gaa (11%). Diabetes diagnosis was based on the following criteria established by the American Diabetes Association Expert Committee: a fasting plasma glucose concentration ≥126 mg/dl (7.0 mmol/l) or a 2-h postload value in the oral glucose tolerance test ≥200 mg/dl (11.1 mmol/l) on more than one occasion. The detection of autoantibodies to GAD antibody and/or a fasting C-peptide ≤0.03 nmol/l were used to exclude probable cases of type 1 diabetes. BMI was computed from weight and height, while fat mass was measured using the bioelectric impedance technique (11). Fasting blood samples were obtained from all participants for the assessment of multiple metabolic traits, including glucose, insulin, and C-peptide, in centralized reference labs at the Diabetes Endocrinology Research Center Immunoassay Core Laboratory in Seattle, Washington. All procedures were approved by the Institutional Review Boards of the five West African universities and of Howard University, and all subjects gave informed consent.

A systematic genome-wide search for diabetes genes/loci was performed on 691 individuals from 343 families using a total of 390 short tandem repeat markers scored by the Center for Inherited Disease Research (http://www.cidr.jhmi.edu) for an average spacing of 9 cM with no gaps >20 cM. A total of 368,160 patient genotypes were released. The overall missing data rate was 4.4% (17,194/386,880 genotypes). Based on 48 blind duplicate samples, the error rate was estimated at 0.1% per genotype (34 discordant calls in 17,244 paired genotypes). A comprehensive list of markers analyzed in this report is available at http://www.cidr.jhmi.edu/markerset.html. The study has >95% power to detect a LOD score of 1, >90% power to detect a LOD score of 2, and 80% power to detect a LOD score of 3, assuming a λ_s of 1.8 and a marker spacing of 9 cM (as was the marker set typed in this study).

Statistical methods. Statistical analyses were preceded by evaluation of the veracity of genetic relationships using PEDCHK and RELTEST (19). As a result of these tests, 14 full sibs were reclassified as half sibs and retained in the analyses. One family was excluded because of a pedigree error that could not be resolved. Marker map positions were taken as that on sex-averaged genetic maps provided by the Marshfield Medical Research Foundation (http://research.marshfieldclinic.org/genetics/Map_Markers/maps).

Model-independent multipoint linkage analysis was performed using the GENEHUNTER-PLUS and ASM programs (20,21). The GENEHUNTER program estimates the nonparametric linkage score at each marker position, comparing the observed identity-by-descent sharing among all affected siblings with that expected under the null hypothesis of no linkage (20). The likelihood ratio test was used to evaluate the null hypothesis of no linkage between a locus and a disease-susceptibility gene. LOD scores were derived from the nonparametric linkage scores, as previously described (21). Maximum LOD scores are presented along with 1-unit support intervals.

REFERENCES

1. King H, Aubert RE, Herman WH: Global burden of diabetes, 1995–2025: prevalance, numerical estimates, and projections. *Diabetes Care* 21:1414–1431, 1998
2. Zimmet P, Alberti KG, Shaw J: Global and societal implications of the diabetes epidemic. *Nature* 414:782–787, 2001

3. Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl M, Nemesh J, Lane CR, Schaffner SF, Bolk S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES: The common PPAR Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26:76–80, 2000
4. Malecki MT, Jhala US, Antonellis A, Fields L, Doria A, Orban T, Saad M, Warram JH, Montminy M, Krolewski AS: Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nat Genet* 23:323–328, 1999
5. Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS: New susceptibility locus for NIDDM is localized to human chromosome 20q. *Diabetes* 46:876–881, 1997
6. Zouali H, Hani EH, Philippi A: Vionnet N, Beckmann JS, Demenais F, Froguel P: A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene. *Hum Mol Genet* 6:1401–1408, 1997
7. Ghosh S, Watanabe RM, Hauser ER, Langefeld CD, Valle T, Magnuson VL, Ally DS, Erdos MR, Balow J, Musick A, Te C, Tannebaum J, Eldridge W, Shapiro S, Martin C, Witt A, So A, Chang J, Shurtleff B, Porter R, Kudelko K, Unmi A, Segal L, Sharaf R, Blaschak-Harvan J, Tuomilehto-Wolf E, Hagopian W, Tuomilehto J, Bergman RN, Collins FS, Boenhke M: Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proc Natl Acad Sci U S A* 96:2198–2203, 1999
8. Mori Y, Otabe S, Dina C, Yasuda K, Populaire C, Lecoeur C, Vatin V, Durand E, Hara K, Okada T, Tobe K, Boutin P, Kadowaki T, Froguel P: Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate loci on 7p and 11p. *Diabetes* 51:1247–1255, 2002
9. Luo TH, Zhao Y, Li G, Yuan WT, Zhao JJ, Chen JL, Huang W, Luo M: A genome-wide search for type II diabetes susceptibility genes in Chinese Hans. *Diabetologia* 44:501–506, 2001
10. Klupa T, Malecki MT, Pezzolesi M, Ji L, Curtis S, Langefeld CD, Rich SS, Warram JH, Krolewski AS: Further evidence for a susceptibility locus for type 2 diabetes on chromosome 20q13.1-q13.2. *Diabetes* 49:2212–2216, 2000
11. Rotimi C, Dunston G, Berg K, Akinsete O, Amoah A, Owusu S, Acheampong J, Boateng K, Oli J, Okafor G, Osotimehin B, Abbiyesuku F, Johnson T, Furbert-Harris P, Kittles R, Vekich M, Adegoke O, Bonney G, Collins FS: In search of susceptibility genes for type 2 diabetes in West Africa: the design and results of the first phase of the AADM Study. *Ann Epidemiol* 11:51–58, 2001
12. Cooper R, Rotimi C, Kaufman JS, Osotimehin BO, Muna W, Kingue S, Fraser H, Forrester T, Wilks R, Cruickshank K: Prevalence of NIDDM among populations of the African Diaspora. *Diabetes Care* 20:343–348, 1997
13. Lindgren CM, Mahtani MM, Widen E, McCarthy MI, Daly MJ, Kirby A, Reeve MP, Kruglyak L, Parker A, Jeyer J, Almgren P, Lehto M, Kanninen T, Tuomi T, Groop LC, Lander ES: Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia study. *Am J Hum Genet* 70:509–516, 2002
14. Yamagata K, Oda N, Kaisaki PJ, Menzel S, Furuta H, Vaxillaire M, Southam L, Cox RD, Lathrop GM, Boriraj VV, Chen X, Cox NJ, Oda Y, Yano H, Le Beau MM, Yamada S, Nishigori H, Takeda J, Fajans SS, Hattersley AT, Iwasaki N, Hansen T, Pedersen O, Polonsky KS, Bell GI: Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature* 384:455–458, 1996
15. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D: The structure of haplotype blocks in the human genome. *Science* 296:2225–2229, 2002
16. Harris MI, Flegal KM, Cowie CC, Eberhardt MS, Goldstein DE, Little RR, Wiedmeyer HM, Byrd-Hold DD: Prevalence of diabetes, impaired fasting glucose, and impaired glucose tolerance in U.S. adults: the Third National Health and Nutrition Examination Survey, 1988–1994. *Diabetes Care* 21:518–524, 1998
17. Neel JV: Diabetes mellitus: a geneticist's nightmare. In *The Genetics of Diabetes*. Creutzfeldt W, Kobberling J, Neel JV, Eds. Berlin, Springer, 1976, p. 1–11
18. Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241–247, 1995
19. O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266, 1998
20. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363, 1996
21. Kong A, Cox NJ: Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 61:1179–1188, 1997