Multiple restriction fragment length polymorphisms at the insulin receptor locus: A highly informative marker for linkage analysis
(non-insulin-dependent diabetes mellitus/insulin receptor cDNA)

S. C. ELBEIN*,†, L. CORSETTI*, A. ULRICH‡, and M. A. PERMUTT*

*Metabolism Division, Washington University Medical Center, 660 South Euclid Avenue, St. Louis, MO 63110; and †Genentech, Inc., Department of Molecular Biology, 460 Point San Bruno Boulevard, South San Francisco, CA 94080

Communicated by Paul E. Lacy, March 28, 1986

ABSTRACT

Although resistance to insulin action is a well-studied phenomenon in non-insulin-dependent diabetes and certain genetic syndromes, the role of inherited defects of the insulin receptor in these disorders is unknown. To facilitate the evaluation of that role, restriction fragment length polymorphisms (RFLPs) were identified using various portions of the insulin receptor cDNA to examine digested DNA from American Blacks, Pima Indians, and Caucasians. Five RFLPs were identified in Caucasians. Two of these were detected with a single 1.3-kilobase probe in Rsa I digests with minor allele frequencies of 0.48 and 0.23. An additional RFLP was noted with Bgl II and two more RFLPs with Sac I using a different 1.6-kilobase probe, with minor allele frequencies of 0.17 for Bgl II and 0.12 for both Sac I RFLPs. All RFLPs except for the second Sac I RFLP were present in American Blacks, while only the Rsa I RFLPs were present in Pima Indians. Pairwise analysis showed random association between all sites except for the Bgl II and second Rsa I sites, where the disequilibrium statistic, \( \Delta \), was \(-0.70\) (different from 0 at \( P < 0.001 \)). No association of any RFLP was noted with non-insulin-dependent diabetes in a small population. These studies show that this is a highly informative locus that should be important for mapping of chromosome 19p and for linkage studies.

Insulin resistance is a well studied but complex phenomenon that is common in non-insulin-dependent diabetes mellitus (NIDDM) (1), obesity (2), and certain rare genetic syndromes (3). This resistance may result from decreased number or affinity of insulin receptors, receptor antagonists, or as yet undefined defects in postbinding insulin action (4). The contribution of primary genetic defects in the insulin receptor to insulin resistance has been obscured by the secondary alterations in receptor number and in postbinding insulin action (4). Nonetheless, in some individuals with leprechaunism and type A insulin resistance associated with acanthosis nigricans, extreme resistance to insulin action appears to result from a primary defect in receptor biosynthesis (3). Even among these individuals, however, the molecular defects will probably be heterogeneous (5), and the same phenotype appears to result both from markedly decreased receptor number and from postbinding defects.

Of the many proteins that are probably involved in the action of insulin at the target tissue, the insulin receptor is the best studied. This large tetrameric glycoprotein consists of two \( \alpha \) chains, which bind insulin, and two \( \beta \) chains, which contain tyrosine kinase and autophosphorylation activity (6). The \( \beta \) subunit may couple the insulin binding to insulin action, but these events remain to be defined (6). Both chains are synthesized from a single mRNA precursor, representing a unique genetic locus (7, 8), which has been mapped by \( \text{in situ} \) hybridization and somatic cell hybrids to the distal short arm of chromosome 19 (9).

The incidence of primary defects of the receptor in diabetic and normal individuals is currently unknown, although such defects may be common. Most of the measured deficit in receptor binding in NIDDM is probably secondary (1, 2, 4), but studies have thus far been unable to distinguish postbinding defects of the receptor from postreceptor defects (3). The availability of cloned DNA for the human insulin receptor makes possible the direct examination of the role of this gene in the various syndromes of insulin resistance. This examination would be accomplished most efficiently by the use of cloned DNA to identify sufficient restriction fragment length polymorphisms (RFLPs) to evaluate the cosegregation of specific receptor alleles with diabetes or insulin resistance in most families (10). By selecting those individuals who are likely to have genetic defects at this locus, this approach limits the use of more tedious methods (such as DNA sequencing) in potentially normal alleles. To this end we have searched for common polymorphisms of the receptor gene in Caucasians, American Blacks, and Pima Indians. We report five RFLPs in Caucasians with three enzymes, resulting in a heterozygosity of 0.91.

MATERIALS AND METHODS

Selection of Subjects and DNA Preparation. DNA was extracted from nondiabetic and NIDDM individuals from Caucasian, American Black, and Pima Indian populations, as described (11). Individuals with NIDDM were diagnosed by accepted criteria (12), while nondiabetic individuals represented healthy individuals or nonhospitalized patients with other medical problems. Caucasian families with maturity onset diabetes of youth (13) and typical NIDDM were used to confirm Mendelian inheritance and to examine linkage. All family members, for whom DNA was available and who were not known to be diabetic, were screened for NIDDM.

Screening for Polymorphisms. Initial screening for DNA polymorphisms was performed in a panel of 10–20 Caucasian and American Black individuals using two probes, 12.1 and 12.2, which included the entire 5.3-kilobase (kb) cDNA (Fig. 1) (7). Digests were carried out with 5 \( \mu \)g of DNA and a 5- to 8-fold excess of restriction enzyme [Bethesda Research Laboratories, International Biotechnologies (New Haven, CT), or New England Biolabs] for 16 hr under conditions specified by the manufacturer. Digested DNA was subsequently separated on 0.8–1.2% agarose gels and blotted by Southern transfer to Zetabind (AMF Cuno) or Nytran (Schleicher & Schuell) (14). Blots were hybridized according

Abbreviations: RFLP, restriction fragment length polymorphism; kb, kilobase(s); NIDDM, non-insulin-dependent diabetes mellitus.

†Present address: Division of Endocrinology, University of Utah School of Medicine, Salt Lake City, UT 84132.
to the manufacturer's protocol, using DNA probes labeled to a specific activity of $0.5-2 \times 10^7$ cpn by nick-translation (14) with $[\beta^3P]dCTP$ (New England Nuclear). Molecular sizes were estimated using either a 1-kb DNA ladder (Bethesda Research Laboratories) or EcoRI-digested DNA labeled with $[\beta^3P]dCTP$ by using DNA polymerase, large fragment (New England Biolabs) (14). Blots were subsequently washed and re-screened with probes derived from HIR 12.1 (Fig. 1 and below). Any RFLP identified in one group was searched for in all three racial groups, but not all probe/enzyme combinations were examined for all groups. Enzymes were chosen preferentially that were more likely to yield polymorphisms based on previous studies (15).

Construction of Probes for Screening. Because the restriction pattern obtained with probe HIR 12.1 (4.3 kb) was too complex for screening most enzymes, fragments of this probe were constructed for further screening by isolation of the 4.3-kb cDNA and further digestion with the following restriction enzymes: Pst I, which gave usable fragments of 1.6 kb, 1.5 kb, 0.8 kb, and 0.3 kb; Bgl II, from which a 1.3-kb fragment was used; and Hinc II, from which a 0.7-kb fragment was isolated (Fig. 1). These smaller fragments were used directly for nick-translation, or in the case of fragments from the Pst I digest, were subcloned into the appropriate sites of pUC 19 for further use.

Confirmation of RFLPs and Establishment of Linkage Relationships. All RFLPs that appeared to have minor allele frequencies of greater than 10% and in which both alleles could be determined were confirmed by repeat digests and by digests with 10-fold excess enzyme. Potential polymorphisms that did not meet these criteria were not examined further. For each racial group, a minimum of 15 individuals, approximately half with NIDDM, were examined for each RFLP, using the probes 12.1B1.3 for Rsa I digests and 12.1P1.6 for Sac I and Bgl II digests (Fig. 1). Allelic frequencies were determined separately for each racial group. Analysis for linkage disequilibrium was performed in each group by establishing haplotypes with each pair of RFLPs in unrelated individuals who were homozygous for at least one RFLP of the pair. For Caucasians, analysis included data from complete haplotypes (see below). The significance of any deviations of the observed frequencies from those expected based on random assortment (linkage equilibrium) of the five RFLPs was tested by $\chi^2$ analysis and by calculation of the standardized disequilibrium statistic, $\Delta$ (16). Mendelian inheritance of all five RFLPs and of haplotypes was demonstrated in eight Caucasian nuclear families of which two had three generations. An additional three pedigrees were studied at loci R1, R2, SI, and S2 (see below). For all available polymorphisms, haplotypes were established from these families and frequencies of the observed and expected haplotypes were included in the pairwise analysis.

Association with Diabetes. The allelic frequencies of each RFLP were compared in nondiabetic and noninsulin-dependent diabetic members of each racial group, and the results analyzed by $\chi^2$ analysis. Diabetics were selected without regard to obesity, insulin resistance, or known family history of diabetes.

RESULTS

Initial screening for RFLPs was performed using the probes HIR 12.1 and HIR 12.2 (Fig. 1) to hybridize to a digest of DNA from American Blacks and Caucasians using the enzymes BamHI, Bcl I, Bgl II, EcoRI, HincII, HindIII, HinfI, Msp I, Pst I, Pvu II, Rsa I, Sac I, and Taq I. With probe HIR 12.2, possible RFLPs were noted with EcoRI and HincII, but these were at allelic frequencies well below 10% and were not pursued further. Probe HIR 12.1 was found to give complex patterns with numerous bands with enzymes HincII, HinfI, Msp I, Pst I, Pvu II, Rsa I, Sac I, and Taq I. Thus further screening was undertaken with five fragments of this probe (Fig. 1). From the screening with HIR 12.1 and these fragments, five RFLPs were found with the enzymes Rsa I, Bgl II, and Sac I in which minor allele frequencies were greater than 10% in Caucasians or American Blacks or both (Table 1). A search for these RFLPs in Pima Indians revealed only the Rsa I polymorphisms (see below). Additional possible polymorphisms, which appeared to be of low frequency or in which both alleles were not identified, were found with Pst I (probe 12.1P0.8), Msp I (probes 12.1P1.6 and 12.1P1.5), and Rsa I (HIR 12.1).

Rsa I. Two distinct RFLPs were identified in Rsa I digests with the probe 12.1B1.3 in all racial groups. The larger RFLP, designated R1, had two alleles of sizes 6.2 kb (+) and 7.0 kb (-) as seen in Fig. 2A. This RFLP appears to result from the presence or absence of a Rsa I site, thus generating a 0.8-kb fragment that is not detected. In American Blacks a third allelic band (allele 3) of intermediate size (6.4 kb) is found (data not shown). This additional allele probably represents the presence of a second Rsa I site 0.6 kb from one end of the 7.0-kb (-) band. Since the present probes do not permit the identification of both alleles for the second site, we have treated R1 in American Blacks as if three alleles were present.

In Pima Indians a different third polymorphic band of 3.4 kb is found, designated allele 4, which also appears to represent an additional Rsa I site 3.4 kb from one end of both the 6.2-kb and 7.0-kb fragments (data not shown). Individuals

---

**Fig. 1.** Restriction map of the 5.3-kb human insulin receptor cDNA. Shaded regions correspond to the coding portion for the $\alpha$ and $\beta$ chains. Restriction sites shown below are as follows: E, EcoRI; B, Bgl I; P, Pst I; H, HincII. Locations for each site refer to the distance from the 5' end of the cDNA in base pairs. The probes used for these studies are shown below in proper position relative to the cDNA restriction map.
allelic frequencies

<table>
<thead>
<tr>
<th>RFLP</th>
<th>Allele</th>
<th>Caucasian</th>
<th>American Black</th>
<th>Pima Indian</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>1 (+)</td>
<td>0.48 ± 0.05</td>
<td>0.67 ± 0.06</td>
<td>0.17 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>2 (−)</td>
<td>0.52</td>
<td>0.28 ± 0.05</td>
<td>0.48 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.06 ± 0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>—</td>
<td></td>
<td>0.34 ± 0.06</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>92</td>
<td>58</td>
<td>54</td>
</tr>
<tr>
<td>R2</td>
<td>+</td>
<td>0.77 ± 0.05</td>
<td>0.63 ± 0.07</td>
<td>0.63 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>88</td>
<td>38</td>
<td>32</td>
</tr>
<tr>
<td>B1</td>
<td>+</td>
<td>0.17 ± 0.04</td>
<td>0.26 ± 0.06</td>
<td>0.02 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>86</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>S1</td>
<td>+</td>
<td>0.88 ± 0.04</td>
<td>0.65 ± 0.07</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>98</td>
<td>48</td>
<td>58</td>
</tr>
<tr>
<td>S2</td>
<td>+</td>
<td>0.12 ± 0.04</td>
<td>0.02 ± 0.02</td>
<td>0.02 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>98</td>
<td>48</td>
<td>48</td>
</tr>
</tbody>
</table>

Frequencies (x ± SEM) for each allele shown for each racial group. The number of alleles scored is shown (n). For RFLP R1, alleles 1 and 2 in Pima Indians and American Blacks are the same as Caucasian alleles ‘+’ and ‘−’, respectively. Alleles 3 and 4 refer to different third alleles in American Blacks and Pima Indians, respectively. The RFLPs R1 and R2 refer to Rsa I sites, B1 to a Bgl II site, and S1 and S2 to Sac I sites.

Homzygous for all three fragment sizes were detected. The current probe probably overlaps only the 3.4-kb portion of the larger fragments, such that only this band is detected in the presence of the additional site. We have, therefore, also treated R1 in Pima Indians as a three allele locus, where alleles 1 and 2 are again equivalent to the Caucasian ‘+’ and ‘−’ alleles. The allelic frequencies for all racial groups are shown in Table 1. Using additional smaller probes, we have mapped R1 to the region between bases 1928 and 2478 of the cDNA (Fig. 1).

A second Rsa I polymorphism, designated R2, is present in all three racial groups and has alleles of 2.2 kb (+) and 2.4 kb (−) (Fig. 2A). The additional bands generated by probes 12.1P.I.5 and 12.1P.I.6 obscure the polymorphic fragments, and this RFLP is thus detected only with the probe 12.1B1.3. We have used smaller probes to map R2 to the region of bases 2744–2963 of the cDNA. Like R1, this locus is very informative, with minor allele frequencies over 0.20 (Table 1).

Table 1. Allelic frequencies

Homzygotes for both 2.2-kb (+) allele (Fig. 2, lane 1) and 2.4-kb (−) allele (data not shown) have been noted. Loci R1 and R2 were easily evaluated with the single probe 12.1B1.3. Genotypic frequencies in unrelated individuals obey Hardy–Weinberg equilibrium for both loci.

Bgl II. A single RFLP was identified in Bgl II digests of DNA from Caucasians and American Blacks using the probes HIR 12.1 and 12.1P.I.6. This polymorphism (designated B1) appears to represent an additional Bgl II site 3.4 kb from the 5' end of a 23.4-kb fragment (‘+’ allele) with the probe 12.1P.I.6. Using probe 12.1H0.7 (Fig. 1), only the large bands (23.4 kb and 20 kb) were observed. The use of additional probes has permitted mapping of B1 to the region of bases 2744–2963 of the cDNA. Homzygotes were observed for both alleles. While the minor (‘+’) allele frequency was 0.17 in Caucasians and 0.26 in American Blacks, this RFLP was observed in only 1 of 48 Pima Indian alleles (Table 1). Genotypic frequencies again agreed with Hardy–Weinberg predictions.

Sac I (Sst I). Two additional polymorphisms were identified with Sac I (or Ssr I) in Caucasians using the probe 12.1P.I.6, although in each case the minor allele frequencies were only 0.12 (Table 1). These RFLPs have been designated S1 and S2. Neither polymorphism is present to a significant degree in Pima Indians. The S1 polymorphism results in alleles of size 5.3 kb (+) and 5.8 kb (−) (Fig. 2C). This RFLP is very common in American Blacks, where the frequency of the ‘+’ allele is 0.65.

The second polymorphism (S2) is detected by the appearance in heterozygotes of bands at 7 and 2.4 kb and by the loss of the 9.4-kb band in homzygotes (Fig. 2C). Thus, the ‘+’ allele represents the presence of a new site in the 7.4-kb restriction fragment 2.4 kb from one end, in which both resulting fragments are detected by the cDNA probe. This RFLP appears to be private to Caucasian populations (Table 1). For both S1 and S2, genotypic frequencies conform to those predicted by Hardy–Weinberg equilibrium.

BamHI. A sixth polymorphism was noted in American Blacks and Caucasians with BamHI and the probe 12.1P.I.6. Two alleles were again noted, with sizes 6.9 and 7.4 kb (data not shown). In 8 American Blacks and 19 Caucasians, the genotype of this locus corresponded exactly with that of the S1 polymorphism, and the size difference between the alleles (0.5 kb) was also the same for both polymorphisms. Subse-

Fig. 2. DNA from unrelated Caucasian individuals was digested with Rsa I (A), Bgl II (B), or Sac I (C), separated on 0.8–1.2% agarose gels, transferred to Zetabind, and hybridized to 32P-labeled probe 12.1P.I.3 (A), 12.1 (B), or 12.1P.I.6 (C). Sizes are shown in kb to the left. (A) For polymorphism R1, individual 2 is homzygous for the ‘−’ allele and individual 3 is homzygous for the ‘+’ allele. Others are heterozygous. At R2, individuals 2 and 4 are homzygous for the ‘+’ allele, while individuals 1, 3, and 5 are heterozygous. Lane M is the size marker. (B) For RFLP B1, individuals 2, 4, and 5 are homzygous for the ‘+’ 23-kb allele, and individual 3 is homzygous for the ‘+’ allele (20 kb + 3.4 kb). Others are heterozygous. Lane 6 is the molecular size marker. (C) At RFLP S1, individuals 4, 5, and 6 are homzygous for the ‘+’ allele (5.3 kb), and all others are heterozygous. For RFLP S2, individual 2 is homzygous for the ‘−’ 9.4-kb allele, and individual 6 is homzygous for the ‘+’ allele (7.0 kb + 2.4 kb). Heterozygotes (1, 4, 5) have bands at 9.4 kb, 7.0 kb, and 2.4 kb.
quently, a third RFLP corresponding exactly to SI was detected with Taq I, and the size difference was again 0.5 kb. All three RFLPs occur on the same allele in three nuclear families studied, map to cDNA bases 2744–2963 (not shown), and thus may represent a 500-base-pair insertion or complete linkage disequilibrium. As expected, the BamHI polymorphism is not detected in Pima Indians. Because this polymorphism is completely redundant with SI, it provides no new information for linkage studies, and further characterization was not attempted.

**Linkage Relationships.** The amount of information provided by each additional RFLP depends on both the frequency of the minor allele and the extent to which the polymorphisms are randomly associated in a population (17, 18). The maximum additional information is gained when RFLPs are in complete random association (linkage equilibrium), while no additional information is gained when complete linkage disequilibrium exists (17), as may occur with the SI and BamHI polymorphisms described above. The exact physical distance between the RFLPs we report is unknown, but we estimate from restriction patterns that cDNA probe 12.1P1.6 spans 35–40 kb of genomic DNA and that probe 12.1B1.3 spans 11–18 kb. These estimates may be low, but the two probes overlap, so that the maximum distance between the sites reported here is probably less than 50 kb, a distance over which substantial linkage disequilibrium might be expected (16, 17). To address this issue, we first examined each pair of RFLPs from unrelated individuals, with marker phase established by homozygosity for one RFLP of the pair or by inheritance (Caucasians). The frequencies observed were compared with those predicted if the sites were in complete linkage equilibrium. For American Blacks, sufficient data was available only for pairs R1B2 (18 alleles), R1B1 (26 alleles), R1SI (30 alleles), B1SI (40 alleles), and R2SI (14 alleles). All pairs were evaluated in Caucasians (Table 2). The sites were randomly associated for all pairs studied in both racial groups, as were sites R1 and R2 in Pima Indians, except for the pair R2BI in Caucasians, which showed significant linkage disequilibrium (P < 0.001). The data for R2BI in American Blacks (8 alleles) showed a similar trend. Values for Δ, the standardized disequilibrium statistic, ranged from −0.20 to +0.18 for all other pairs and were not significantly different from 0 (P > 0.05). For pair R2BI, however, Δ was −0.70 in Caucasians and −1.0 in American Blacks (different from 0 at P < 0.001).

The degree of random association was further examined by establishing complete haplotypes from 13 Caucasian families used to confirm Mendelian inheritance. Of the 32 haplotypes expected from five RFLPs, 16 haplotypes would have an expected frequency of over 0.01 assuming random association.

Table 2. **Pairwise associations in Caucasians**

<table>
<thead>
<tr>
<th>Loci</th>
<th>Haploype</th>
<th>+ +</th>
<th>+ +</th>
<th>+ +</th>
<th>+ +</th>
<th>Sample size (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1R2</td>
<td>33 (33)</td>
<td>11 (10)</td>
<td>39 (36)</td>
<td>7 (11)</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>R1B1</td>
<td>4 (6)</td>
<td>27 (27)</td>
<td>5 (6)</td>
<td>32 (29)</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>R1SI</td>
<td>32 (34)</td>
<td>7 (5)</td>
<td>37 (37)</td>
<td>4 (5)</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>R1SI</td>
<td>6 (5)</td>
<td>33 (34)</td>
<td>2 (5)</td>
<td>39 (34)</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>R2BI*</td>
<td>3 (9)</td>
<td>59 (45)</td>
<td>7 (3)</td>
<td>2 (13)</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>R2SI</td>
<td>12 (8)</td>
<td>59 (58)</td>
<td>1 (2)</td>
<td>15 (17)</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>B1SI</td>
<td>11 (13)</td>
<td>0 (2)</td>
<td>68 (58)</td>
<td>11 (8)</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>B1SI</td>
<td>0 (2)</td>
<td>14 (14)</td>
<td>9 (10)</td>
<td>69 (65)</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>S1SI</td>
<td>12 (11)</td>
<td>79 (82)</td>
<td>0 (2)</td>
<td>15 (11)</td>
<td>106</td>
<td></td>
</tr>
</tbody>
</table>

Observed haplotypes were constructed for each pair of RFLPs. Expected numbers (in parentheses) were calculated from the frequencies in Table 1, and the number of haplotypes available for each pair, n. Deviations from random association (linkage equilibrium) were tested. For R2BI, the disequilibrium statistic, Δ = −0.70. *P < 0.001.

The 10 observed haplotypes are shown in Table 3. Since not all families were studied with Bgl II, not all potential haplotypes at this site were available. Nonetheless, the data confirm the pairwise analysis. Since many of the expected haplotypes are observed. The frequencies do not differ significantly from those predicted assuming completely random association with the exception of those haplotypes that include the R2 “+” and B1 “+” alleles, which occur more often than expected. This occurrence is consistent with the linkage disequilibrium between these sites noted by pairwise analysis. If the sites described here were in completely random association, the predicted heterozygosities would be 0.85 for Caucasians, 0.91 for American Blacks, and 0.82 for Pima Indians. The heterozygosities actually observed were 0.91 (43/47), 0.87 (20/23), and 0.75 (18/24), respectively. Despite the completely random association of markers R1, R2, SI, and S2 in the population, in no family was recombination between sites observed.

**Association with Non-Insulin-Dependent Diabetes.** While the primary goal of these studies was to identify RFLPs suitable for linkage analysis, we also examined the distribution of each RFLP between diabetic and nondiabetic groups for each of the three populations. Because the number of individuals when analyzed in this manner was small (10–20 nondiabetics and 10–15 diabetics for each racial group), only marked associations with NIDDM would be noted. No such association of any RFLP was found in any racial group.

**DISCUSSION**

The availability of the insulin receptor cDNA permits the application of many new techniques to the study of insulin action at the receptor. Unfortunately, the direct analysis of potential genetic defects by cloning and sequencing this large gene would be a formidable task. The RFLPs described here provide a powerful alternative for screening for defects. Because most families will be informative at this locus, only families in which insulin-resistance, hyperinsulinemia, or diabetes cosegregate with a defined receptor haplotype need be studied further at this locus. Additionally, where genomic DNA is to be cloned, these RFLPs potentially provide the means to identify the defective allele. Fortunately, this approach is independent of any secondary changes in receptor properties, and thus eliminates the possibility of attributing defects in regulatory proteins to the insulin receptor.

Although several of the RFLPs defined here have minor allele frequencies of less than 0.20, the completely random association between four of these sites makes this locus highly informative nonetheless. Indeed, to our knowledge this is the most informative locus currently identified for...
mapping and linkage studies of the distal short arm of chromosome 19 (9).

The random associations between most of the markers reported here are surprising, especially since three of the five RFLPs are racially specific, and thus are presumably of relatively recent origin. Only the pairs SI–BamH1, and R2–B1 show significant linkage disequilibrium, although lesser degrees of nonrandom association might be detected for other pairs with a larger sample size. The exact size of the genomic DNA spanned by the insulin receptor cDNA is currently unknown. We have estimated the distance between our RFLPs to be less than 50 kb, although these estimates are crude. By the uniform recombination hypothesis these would be small map distances (19). In those genes in which linkage disequilibrium has been studied, the degree of disequilibrium has correlated poorly with physical distance (20). Nonetheless, the small distance and relatively recent propagation of these RFLPs would suggest increased recombination at this locus (18). Interestingly, those pairs showing significant nonrandom association are present only in Caucasian and American Black populations and thus may be of more recent origin. The hypothesis of increased recombination will be more formally testable if recombinants can be identified in pedigrees and when the distance between RFLPs is accurately determined. We have found no such instance of recombination to date, but even significant increases in recombination frequency would be detected only by the evaluation of large numbers of families. The presence of increased recombination could be of functional significance in promoting mutations or deletions resulting in defective genes in a fashion analogous to the deletion between Alu repeats of the low-density lipoprotein receptor gene (21).

Although our intent in identifying RFLPs for the insulin receptor was to permit linkage analysis, an alternative and complementary method of evaluating candidate genes is the association of RFLPs with the disease. Such studies have been successful for HLA associations with insulin-dependent diabetes mellitus (22) and for an apoprotein A-1 polymorphism with hypertriglyceridemia (23). We could identify no such associations with NIDDM when each racial group was evaluated separately, but our study is limited by small sample size and by the possibility that the diabetic population included individuals with diverse etiologies. The high degree of linkage equilibrium between markers suggests that associations, which depend on linkage disequilibrium between a marker and the actual defect, could be difficult to demonstrate. We would argue that linkage analysis is a more powerful tool for investigating the role of this gene in diabetes. Analysis has been informative for linkage with diabetes in nearly all families attempted (unpublished data), thus demonstrating the feasibility of such studies.

Further studies in additional diabetic families will be necessary to better evaluate the role of the insulin receptor in this disease. Hopefully these studies will identify families in which receptor defects are highly probable. This locus will also be very useful as a means to evaluate this extensively studied gene in a number of other human diseases and should especially permit the rapid screening of those families with type A insulin resistance and acanthosis nigricans, lipodystrophic diabetes, and leprechaunism for the identification of families with true insulin receptor defects. Equally important, this locus will be invaluable for general linkage studies of chromosome 19p.

We thank Bill Knowler for Pima Indian samples, Mark Skolnick for Utah Mormon pedigrees, Steve Fajans for several diabetic pedigrees, and Ken Ward for a large diabetic pedigree. Special thanks to Susan Starbuck for help in preparing the manuscript. This work was supported by Grants AM-31866 and AM-16746 from the National Institutes of Health. S.C.E. is a recipient of National Research Service Award AM-07120 from the National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases.