Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5759–5763, September 1981
Genetics

Polymorphic DNA region adjacent to the 5′ end of the human insulin gene

(human genetics/DNA polymorphism/diabetes mellitus)

GRAEME I. BELL*, JOHN H. KARAM†, AND WILLIAM J. RUTTER*

*Department of Biochemistry and Biophysics and †Department of Medicine, Metabolic Research Unit, University of California, San Francisco, California 94143

Communicated by Arno G. Motulsky, May 26, 1981

ABSTRACT **The length of a segment of DNA associated with the human insulin gene, which has been localized to the short arm of chromosome 11, is heterozygous in 63% of 52 individuals analyzed. This polymorphic region is approximately 500 base pairs from the nucleotide encoding the 5′ end of insulin mRNA. The polymorphism appears to be due to an insertion or deletion of DNA sequences so that DNA fragments of different length are generated when DNA from a heterozygous individual is digested with selected restriction endonucleases.**

An indication of the DNA sequence variation among individual humans is emerging from studies of the structure of allelic segments of chromosomes (1–5). A comparison of the nucleotide sequence of two allelic human insulin genes [*INS*, chromosome 11 (6, 7)] that included coding, intervening, and 5′ and 3′ flanking sequences indicated that only 4/1725 (0.23%) of the nucleotides varied (4). Two of these differences were in the region encoding the 3′ untranslated portion of the mRNA, and the others were within the two intervening sequences. A similar analysis of the coding and intervening sequences of two alleles of the embryonic β-globin gene, γ_A, demonstrated 15 single base changes and 3 gaps of 4, 4, and 18 base pairs (bp) in a segment of 1482 bp (1.0–1.2% variation) (5). With the exception of a base substitution in the region encoding the 5′ untranslated portion of the mRNA, the allelic differences were within the intervening sequences. Jeffreys (1) analyzed the entire β-globin gene region [*NAG*, chromosome 11 (8)] of 60 individuals by restriction endonuclease digestion and hybridization techniques and estimated that 1% of the nucleotides in this region varied between individuals. In these studies, no nucleotide differences were observed in those gene regions that encode protein sequences. Thus, comparisons of DNA sequence variation of two different chromosomal segments, insulin and γ_A-globin, have revealed that the sequences of the alleles are similar, that most of the variability is a consequence of base substitutions, and that the amount of sequence variation is not constant and varies from one gene to another. In addition, Wyman and White (9) have recently described a highly polymorphic locus not associated with any particular gene. Their analysis indicated that the polymorphism in restriction fragment length was a consequence of DNA rearrangement.

In our continuing analysis of the human insulin gene region of chromosome 11, we have examined the insulin gene and flanking regions of 52 unrelated individuals. In contrast to the small differences between allelic DNA segments described above, we have observed a region of length heterogeneity located approximately 500 bp before the start of insulin mRNA synthesis. The sizes of the insertions in this region are not uniform and appear to be primarily of two classes: 0–600 and 1600–2200 bp. As a consequence of this length heterogeneity, the two parental insulin genes can be distinguished in 63% of the individuals examined. The physiological consequences of this length instability are unknown, but the observation demonstrates that substantial differences in length as well as less dramatic differences in nucleotide sequence can identify chromosome homologues.

MATERIALS AND METHODS

DNA Preparation. Human DNA was isolated from lymphocyte nuclei by using a procedure obtained from L. Kunkel (10). Ten milliliters of blood collected in a heparinized tube and then kept at 4°C was mixed with 90 ml of 0.32 M sucrose/10 mM Tris·HCl (pH 7.5)/5 mM $MgCl_2$/1% Triton X-100 at 4°C to lyse all cells. The nuclei were collected by centrifugation at 1000 × *g* for 10 min. The nuclear pellet was suspended in 4.5 ml of 0.075 M NaCl/0.024 M EDTA (pH 8.0) with a Pasteur pipette. Then 0.5 ml of 5% sodium dodecyl sulfate and proteinase K at 2 mg/ml were added and the mixture was incubated for approximately 12 hr at 37°C. The digest was gently mixed with 5 ml of phenol saturated with 20 mM Tris·HCl (pH 8.0). Five milliliters of chloroform/isoamyl alcohol (24:1, vol/vol) was added and gentle mixing was continued. The phases were separated by centrifugation for 15 min at 1000 × *g*. The upper, aqueous, phase was removed and gently extracted with the chloroform/isoamyl alcohol mixture. After centrifugation, the aqueous phase was removed; 0.5 ml of 3 M sodium acetate and 11 ml of 100% ethanol (at room temperature) were added. The DNA was precipitated by inverting the tube several times and then removed with a Pasteur pipette and placed in 1 ml of 10 mM Tris·HCl (pH 7.5)/1 mM EDTA. The DNA was allowed to dissolve for several days at 4°C before being digested with restriction endonucleases. From 10 ml of blood, 200–500 μg of DNA was obtained. High molecular weight DNA could also be prepared from blood stored several days at 4°C.

Analysis of Insulin Gene and Flanking Sequences. Approximately 5 μg of DNA was digested in a volume of 100 μl for 12–20 hr with 5–20 units of restriction enzyme. The enzymes were obtained from New England BioLabs and Bethesda Research Laboratories (Rockville, MD) and used according to the manufacturer's specifications. Digestions always contained nuclease-free bovine serum albumin (Bethesda Research Laboratories) at 100 μg/ml to stabilize enzymes. The digestions were terminated by the addition of 10 μl of 3 M sodium acetate and 330 μl of 100% ethanol. After 10 min in a dry ice/ethanol bath, the precipitated DNA was collected by centrifugation for 5 min in an Eppendorf microcentrifuge (Brinkmann). The DNA pellet

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: bp, base pairs; kb, kilobase(s); NaCl/Cit, 0.15 M NaCl/0.015 M sodium citrate.

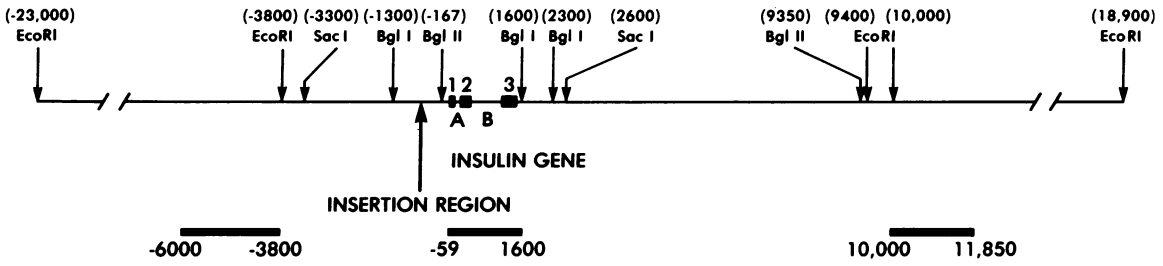

FIG. 1. Map of the human insulin gene and adjacent regions. The relevant restriction enzyme sites are indicated as well as their coordinates in bp relative to the 5′ end of insulin mRNA (position 1). The dark boxes (1–3) are the insulin mRNA coding regions and A and B are intervening sequences. The three regions from which DNA fragments were prepared for use as hybridization probes are indicated as heavy lines below the restriction map. The arrow labeled insertion region indicates the position of the insertion–deletion loop present in heteroduplexes of the cloned human insulin gene alleles λHi1 and λHi2 (3). The region from −21,000 to 25,800 has been isolated as a series of overlapping λCh4A–human DNA recombinants.

was rinsed with 70% (vol/vol) ethanol and then dried for 10 min under reduced pressure. The digested DNA was dissolved in 25 μl of 5% Ficoll 400 (Sigma)/5 mM EDTA (pH 8.0)/0.2% sodium dodecyl sulfate/0.01% bromphenol blue for at least 60 min at room temperature with frequent mixing. The samples were heated at 65°C for 10 min before electrophoresis. The DNA fragments were separated by electrophoresis in 0.85% agarose gels (0.3 × 14 × 10 cm), prepared in 0.040 M Tris acetate (pH 8.1)/0.025 M sodium acetate/0.002 M EDTA, at 25 mA constant current, generally for 10 hr. *Hin*dIII-digested λ phage DNA, *Hae* III-digested φX174 phage DNA, and fragments of the cloned human insulin gene were used as size markers. The separated DNA fragments were transferred from the agarose gel to nitrocellulose filters (Schleicher & Schuell, BA85) as described by Southern (11). The filter was dried under reduced pressure for 2 hr at 80°C and then prehybridized in 5× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate)/50% (vol/vol) formamide/50 mM sodium phosphate (pH 6.8)/sonicated and denatured salmon sperm DNA at 250 μg/ml/5× Denhardt's solution [1× Denhardt's solution is 0.02% bovine serum albumin/0.02% Ficoll 400/0.02% polyvinylpyrrolidone 360 (Sigma)]/0.1% sodium dodecyl sulfate in a sealed bag at 42°C. After 24 hr this solution was replaced with one containing 5× NaCl/Cit, 50% formamide, 10% dextran sulfate (11), 20 mM sodium phosphate (pH 6.8), sonicated and denatured salmon sperm DNA at 100 μg/ml, 5× Denhardt's solution, 0.1% sodium dodecyl sulfate, and 2.5 × 10^5 cpm/ml of nick-translated DNA (12, 13). After 24 hr at 42°C, the filters were washed briefly at room temperature in 2× NaCl/Cit and 0.1% sodium dodecyl sulfate and then for 1 hr at 50°C in 0.1×

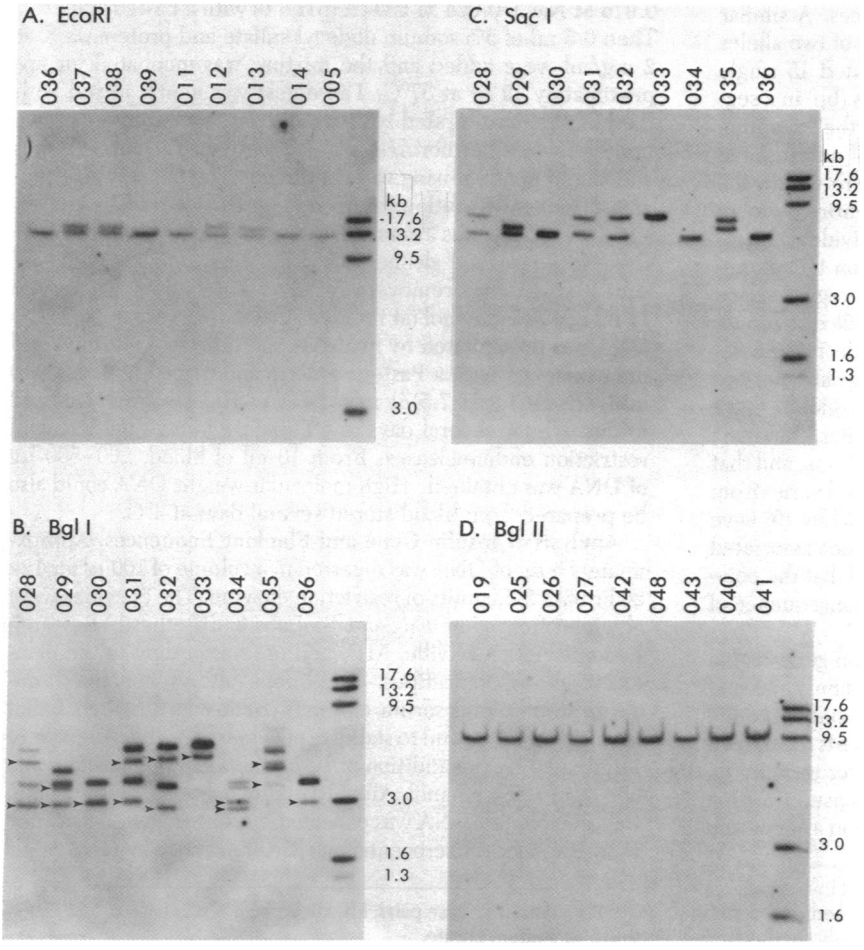

FIG. 2. Polymorphism of restriction fragment length in human DNA. The panels are representative autoradiograms of hybridization of a DNA fragment, which contains the human insulin gene and closely flanking regions (−59 to 1600, Fig. 1), to the indicated restriction endonuclease digests of DNA prepared from individuals (indicated by number). As discussed in the text, the *Bgl* I site at coordinate 1600 (Fig. 1) is partially modified in these DNA preparations. Therefore, two hybridizing fragments are generated from each chromosome. One, *Bgl* I fragment −1300 to 1600, indicated by the arrowheads in *B*, *Bgl* I digests, is produced by cleavage of the *Bgl* I sites at −1300 (or equivalent position) and 1600. When the *Bgl* I site at 1600 is modified, it is not cleaved by this enzyme. Cleavage of the next *Bgl* I site, position 2300, generates the other hybridizing fragment, which is 700 bp larger, *Bgl* I fragment −1300 to 2300. The right-hand column in each autoradiogram is the hybridization of fragments obtained by digestion of the 13.2-kb insulin gene-containing *Eco*RI fragment of λHi2 (2, 3) subcloned in pBR322 (pHi2) with *Sal* I (17.6 kb), *Eco*RI (13.2 kb), *Bgl* II (9.5 kb), *Bgl* I (3.0 kb), or *Pvu* II (1.6 and 1.3 kb). *Bgl* I and *Sac* I digests of DNA from the same individuals are presented and indicate that, although the two alleles may appear to have identical sizes after *Sac* I digestion (e.g., individual 034), they can often be resolved by digestion with *Bgl* I.

NaCl/Cit and 0.1% sodium dodecyl sulfate with constant gentle shaking. The washed filter was blotted dry, wrapped in Saran Wrap (Dow), and exposed to preflashed Kodak X-Omat R x-ray film (14) with a single Du Pont Lightning Plus intensifying screen for 5–7 days at −76°C. After exposure the filter-bound DNA could be hybridized with another labeled DNA probe. The previously hybridized filter was washed at 65°C for 5–6 hr in the prehybridization solution, then transferred to 42°C for the duration of the prehybridization step, before hybridization with another DNA probe.

RESULTS

Comparative restriction endonuclease and DNA heteroduplex analyses of two cloned human insulin gene alleles and their adjacent noncoding regions revealed a section of different length due to insertion-deletion of DNA in the 5′ flanking region of the gene located approximately 500 bp from the start of insulin mRNA transcription (3). This observation prompted us to examine this region in other individuals to determine whether this DNA length polymorphism was a general feature of the 5′ flanking region of the human insulin gene.

DNA was prepared from 52 unrelated individuals and digested with restriction endonucleases, and the size(s) of insulin gene-containing fragments were determined by hybridization. The probe was a 1650-bp fragment prepared from one of the cloned insulin gene alleles (λHi1) (2, 3) and in addition to the structural gene included 59 5′ flanking and approximately 120 3′ flanking nucleotides. Fig. 1 shows a map of approximately 40 kilobases (kb) of human DNA containing the insulin gene, including relevant restriction endonuclease sites and the locations of the DNA fragments used as probes in the hybridization analyses described here. *Eco*RI digestion generated insulin gene-containing DNA fragments of approximately 13 kb in 30 of the 52 individuals analyzed—i.e., similar in size to the *Eco*RI fragment in the clone λHi2 (see Fig. 2*A* for representative patterns); thus the insulin alleles were indistinguishable in these individuals. However in 19 individuals the two insulin gene alleles were not identical, one was 13 kb, the other 15 kb. In three other individuals, the two *Eco*RI-generated insulin gene fragments were equivalent in size but larger, 15 kb.

This restriction fragment length heterogeneity could be due to an *Eco*RI restriction site polymorphism or to a DNA rearrangement (i.e., insertion–deletion) as suggested by the analysis of the cloned alleles. These two possibilities were tested by determining the sizes of the insulin gene alleles obtained after digestion with another restriction endonuclease. If the restriction fragment length polymorphism observed with *Eco*RI were a consequence of a restriction site polymorphism, it should not be evident on digestion with another enzyme. However, if it were due to a DNA rearrangement it should be observed with other enzymes as well. The restriction fragment length heterogeneity was still present after *Bgl* I or *Sac* I digestion (Fig. 2 *B* and *C*), indicating that the length difference observed between the two alleles was due to insertion into or deletion of DNA from the chromosome near the insulin gene. Moreover, the differences in length of allelic DNA segments of each individual examined were similar whether determined by *Sac* I or *Bgl* I digestion. However, as discussed below, *Bgl* I digestion was more sensitive for resolving small differences in DNA fragment length, and it was often possible to distinguish the two allelic DNA segments generated upon *Bgl* I digestion, even though they may have been indistinguishable upon *Sac* I digestion. The *Bgl* I digestion pattern contains an additional and unexpected hybridizing fragment, 700 bp larger than the fragment expected from complete digestion. This additional fragment is resistant to further *Bgl* I digestion. Moreover, it is not observed when DNA prepared from placenta or cultured cells is digested with *Bgl* I. A comparison with the *Bgl* I restriction map in the vicinity of the insulin gene indicates that the *Bgl* I site at coordinate 1600 (Fig. 1) is resistant to digestion in some fraction of the DNA molecules, possibly due to a modification of the DNA. The DNA used in these studies was prepared from a mixed population of blood cells, so the *Bgl* I-1600 site may be altered in a subfraction of these cells.

Analysis of the *Bgl* I and *Sac* I digests indicated that the fragment lengths of the DNA in this region were not uniform and that no simple pattern existed. Whereas the insulin gene fragments could be distinguished in only 22 of 52 individuals after *Sac* I digestion, 33 of 52 individuals were shown to be heterozygous in this region by *Bgl* I digestion. This apparent increase in heterozygosity most likely reflects the ability to resolve small differences (100 bp) in fragment size when DNA is digested with *Bgl* I because this enzyme generates much smaller insulin-gene containing fragments (approximately 3.0 kb) than either *Sac* I (6.0 kb) or *Eco*RI (13.5 kb). Furthermore, the actual degree of

Table 1. Analysis of sizes of *Bgl* I fragments of human DNA that contain insulin gene sequences

Individual	Size, kb	Individual	Size, kb
001 (IDDM)	3.08	028 (NIDDM)	2.90, 4.45
002 (NIDDM)	3.10	029 (IDDM)	2.90, 3.53
003 (IDDM)	2.83, 3.05	030 (NIDDM)	2.98
004 (NIDDM)	2.80, 3.00	031 (NIDDM)	3.00, 4.55
005 (IDDM)	2.85	032 (IDDM)	2.80, 4.65
006 (IDDM)	2.93, 4.60	033 (NIDDM)	4.80
007 (IDDM)	2.83, 3.03	034 (IDDM)	2.78, 2.98
008 (NIDDM)	2.83, 2.98	035 (NIDDM)	3.55, 4.35
009	3.00, 4.60	036 (IDDM)	2.93
010 (IDDM)	3.03	037 (NIDDM)	3.00, 4.70
011	2.93	038 (NIDDM)	3.00, 4.70
012	2.80, 5.03	039 (IDDM)	2.90
013	2.80, 4.40	040	2.95
014 (NIDDM)	2.73, 2.90	041	2.75, 2.90
015	2.78, 2.98	042	4.50
016	2.93	043	2.93, 4.45
017	2.98, 4.55	044	2.93, 4.70
018	3.00, 4.40	045	2.83, 3.03
019	3.05, 5.30	046	2.90, 4.60
020	3.03, 3.25	047	3.00, 4.80
021	2.95, 3.15	048	4.60
022	2.93, 3.08	049	2.95
023	3.08	050 (IDDM)	2.80, 2.95
024	3.00	051	2.85
025	3.00, 4.80	052 (NIDDM)	2.95
026	2.95, 4.80		
027	3.03, 4.80		

Each *Bgl* I fragment corresponds to a segment of one of the two chromosomes 11 present per diploid nucleus. A single region of insulin-specific hybridization indicates that the *Bgl* I fragment from each chromosome is the same size. DNA samples from those individuals grouped together were analyzed on the same agarose gel. The fragment size is the mean of two determinations, and for those fragments of approximately 3 kb, the agreement between determinations was ±50 bp. It is possible to identify the two alleles if they differ by 100 bp—e.g., individual 021. A comparison of the sizes of the *Bgl* I fragments of individuals in the group 028–036 with the data in Fig. 2*B* shows that it is possible to observe small differences in *Bgl* I fragment size when two individuals are compared even when the differences are too small to be identified within an individual. Individuals with diabetes mellitus are indicated: IDDM, insulin-dependent diabetes mellitus; NIDDM, non-insulin-dependent diabetes mellitus. All other individuals were normal, with no pronounced family history of diabetes mellitus.

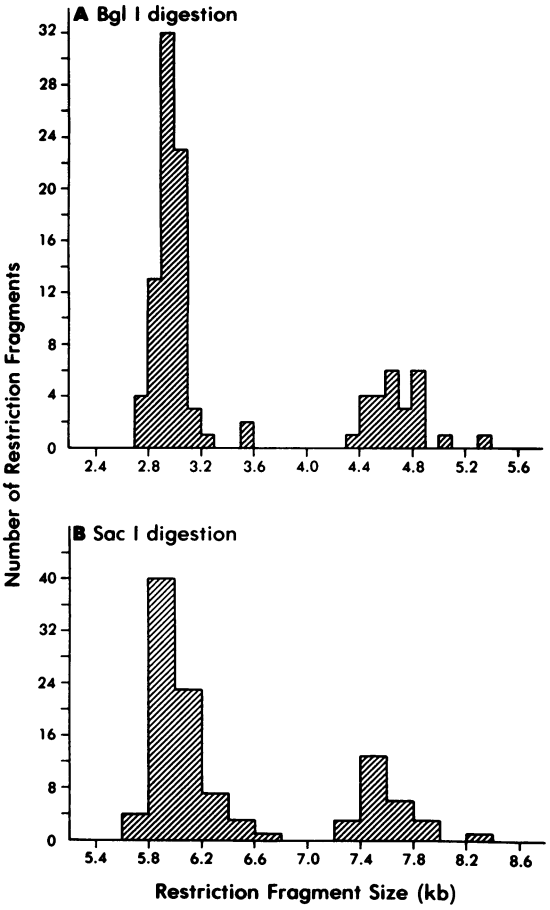

FIG. 3. Distribution of sizes of *Bgl* I-and *Sac* I-generated restriction fragments that hybridize with an insulin gene-specific sequence homologous to the region −59 to 1600. Data tabulated as indicated in Table 1 are presented graphically. Those individuals whose DNA shows only a single hybridizing region are assumed to have two insulin gene fragments of the same size.

polymorphism at this locus, 63% (33/52), may be even greater than demonstrated by *Bgl* I digestion, because length differences of less than 100 bp have not been resolved.

The comparison of two cloned insulin alleles localized the region of DNA length polymorphism to the 5′ flanking region of the insulin gene, 500 bp from the start of the mRNA (Fig. 1) (3). If this were the location in other alleles as well, digestion with *Bgl* II should generate fragments of identical length, 9.5 kb, on hybridization with the insulin gene probe used, because the fragments generated by this enzyme do not include the region of heterogeneous length (Fig. 1). The digestion with *Bgl* II of 20 DNA samples, which displayed length heterogeneity after *Bgl* I digestion, produced a uniquely hybridizing fragment of 9.5 kb (Fig. 2*D*). This suggests that location of the polymorphic region is most likely the same (see Fig. 1) in all the alleles analyzed.

An analysis of the sizes of the insulin gene fragments produced by *Bgl* I digestion indicates that they can differ by 150 to 2500 bp (Table 1). However, display as a histogram (Fig. 3*A*) reveals that the distribution of fragment sizes is not random but is bimodal, with the largest group being of alleles of 2.9–3.0 ± 0.3 kb and the other of 4.6–4.7 ± 0.3 kb. This distribution is confirmed by a similar analysis of the *Sac* I digests (Fig. 3*B*).

We have also analyzed adjacent regions more distant from the insulin gene for restriction fragment length polymorphisms. Hybridization to *Eco*RI digests of 52 individuals with a probe from coordinates −6000 to −3800 (Fig. 1) indicated a single region of hybridization of 19 kb (Fig. 4*A*) and the absence of large rearrangements in this region. Similarly, analysis of *Eco*RI digests of the DNA of 39 individuals with a probe from coordinates 10,000 to 11,850 (Fig. 1) revealed a single hybridizing region of 9 kb (Fig. 4*B*). Thus in the 42 kb of DNA flanking and including the insulin gene, there appears to be only a single highly polymorphic region.

We have analyzed this length polymorphism in normal and diabetic individuals (Table 1). This preliminary analysis did not reveal any direct correlation between diabetes mellitus and this polymorphism in normal, insulin-dependent diabetic, or non-insulin-dependent diabetic individuals.

DISCUSSION

We have identified a highly polymorphic DNA region in the 5′ flanking region of the insulin gene that we previously localized to the short arm of chromosome 11, in the region p13→pter (6, 7). This polymorphic region lies outside the mRNA coding

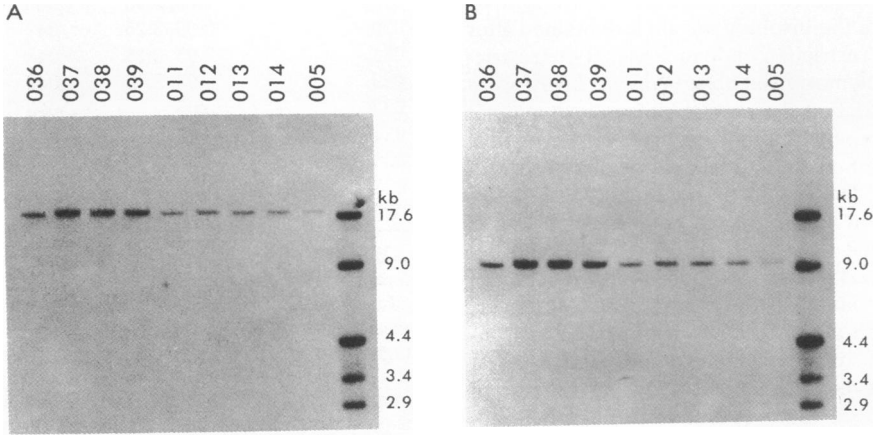

FIG. 4. Analysis of regions flanking the human insulin gene for other highly polymorphic DNA regions. DNA was digested with *Eco*RI, and the fragments were separated by agarose gel electrophoresis and transferred to a nitrocellulose filter. The filter-bound DNA was hybridized sequentially with *in vitro* labeled DNA after elution of previously hybridized DNA. The 1.85- and 2.2-kb *Eco*RI fragments of λHi2 (λCh4Aλ Hi14.1) (2, 3) subcloned in pBR322 were used as probes [the order of the 1.85- and 2.2-kb *Eco*RI fragments in λHi2 is incorrect in Bell *et al.* (2)]. (*A*) Autoradiogram of hybridization of DNA homologous to the region −6000 to −3800 (pBR322 −2.2). (*B*) Autoradiogram of hybridization of DNA homologous to the region 10,000 to 11,850 (pBR322 −1.85). The right-hand column is hybridization of the pBR322 sequences in the probes to the digests of pHi2 described in Fig. 2. Fig. 2*A* is an autoradiogram of hybridization to this same filter of DNA homologous to the region −59 to 1600.

portion of the insulin gene and the immediately flanking regulatory regions and therefore it is unknown whether it should be considered as part of the insulin gene or whether it is an intergenic spacer region or is part of another unidentified gene linked to insulin. In the absence of contradictory information and because of its proximity to the insulin gene, we consider it part of this gene. The rearrangements that generate the polymorphism are unlikely to be tissue-specific, because we have also observed them in DNA prepared from placenta. Also the two cloned insulin alleles, whose comparison provided the first evidence for this length heterogeneity, were from a library prepared from fetal liver DNA. Wyman and White (9) recently described another restriction fragment length polymorphism due to DNA rearrangement at an unidentified locus. The length differences they observed at this locus, 0.3–14 kb, were generally much larger than those in the region flanking the insulin gene, 100–2600 bp. Wyman and White also demonstrated that the polymorphism they were examining segregated as a Mendelian trait. Therefore, it is unlikely that these length polymorphisms are generated by somatic rearrangement.

A comparison of the sizes of restriction fragments that span the polymorphic region flanking the human insulin gene indicates that their distribution is nonrandom and that length differences are of two classes, 0–600 and 1600–2200 bp. The significance of this observation is unclear, but it suggests that there may be functional constraints operating on this section of DNA. For example, it is possible that due to the close proximity of this region to the insulin gene, approximately 500 bp, the DNA length variation could affect the structure of the insulin gene in chromatin. The molecular basis for this length heterogeneity is also undetermined. However, the 100-bp DNA segment present in only one of the two cloned insulin gene alleles, λHi2, is a single-copy sequence, because it hybridizes to a single size class of DNA fragments after restriction endonuclease digestion. Thus the smaller class of polymorphic fragments observed is probably not a consequence of mobilization of a transposable genetic element like the elements that have been shown to generate restriction fragment length polymorphisms in *Escherichia coli*, yeast, and *Drosophila* (15–17). Also, the sizes of the insertions, 1600 bp, in the other, larger class of polymorphic restriction fragments are greater than *Alu* family middle repetitive sequences, approximately 350 bp (3, 18), or the unit size, 170 bp, of several highly repetitive sequence elements in the human genome (19), assuming only a single element.

Besides this highly polymorphic rearrangement we have also previously identified a *Pst* I restriction site polymorphism in the region of the insulin gene encoding the 3′ untranslated portion of the mRNA (2). However, we have not determined the frequency of this polymorphism in populations. We have also identified, in a preliminary survey, a *Hin*cII site polymorphism in one of 20 individuals tested. Thus it should be possible to distinguish the two parental insulin genes in most individuals by determining the size(s) of restriction fragment(s) containing insulin gene sequences after *Bgl* I, *Hin*cII, or *Pst* I digestion.

The analysis of polymorphisms in DNA restriction fragment length in human populations has facilitated the molecular understanding and detection of abnormal hemoglobins in thalassemias and sickle cell anemia (20, 21). Moreover, as Botstein *et al.* (22) have suggested, a catalogue of highly polymorphic regions throughout the human genome could be the basis for a new aspect to human genetics and for increased understanding of some disease states. Diabetes mellitus is a poorly classified family of diseases, and the primary causes and underlying molecular bases are unknown. However in many instances there appears to be a family association with predisposition to the disease. The ability to distinguish the maternal and paternal insulin genes and flanking regions of most individuals now allows us to determine if an altered insulin gene or a closely linked locus can be involved in the disease process. Twenty-four individuals with diabetes mellitus were examined in this study, including 12 individuals with insulin-dependent (type I) diabetes and 12 with non-insulin-dependent (type II) diabetes. There was no correlation of any feature of the highly polymorphic DNA region that distinguished either group from normal, nondiabetic individuals (Table 1). However, the description of mutant insulins (23–25), including one that causes diabetes, suggests that diabetes in some individuals is no doubt a consequence of a mutant insulin gene and should segregate with this polymorphism. Thus an analysis of the segregation of DNA polymorphisms and diabetes in families with a pronounced hereditary predisposition to diabetes could provide a basis for identifying certain forms of the disease.

We thank Drs. Orgad Laub and Raymond Pictet for their material and valuable discussions during this project, and Ms. Leslie Spector for her assistance in the preparation of the manuscript. This work was supported by the Elliott P. Joslin Research and Development Award of the American Diabetes Association to G.I.B., and by grants from the National Institutes of Health (GM28520) and Eli Lilly to W.J.R.

1. Jeffreys, A. J. (1979) *Cell* **18**, 1–10.
2. Bell, G. I., Pictet, R. L., Rutter, W. J., Cordell, B., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **284**, 26–32.
3. Bell, G. I., Pictet, R. & Rutter, W. J. (1980) *Nucleic Acids Res.* **8**, 4091–4109.
4. Ullrich, A., Dull, T. J., Gray, A., Brosius, J. & Sures, I. (1980) *Science* **209**, 612–615.
5. Slightom, J. L., Blechl. A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
6. Owerbach, D., Bell, G. I., Rutter, W. J. & Shows, T. B. (1980) *Nature (London)* **286**, 82–84.
7. Owerbach, D., Bell, G. I., Rutter, W. J., Brown, J. A. & Shows, T. B. (1981) *Diabetes* **30**, 267–270.
8. McKusick, V. A. (1980) *Am. J. Med.* **69**, 267–276.
9. Wyman, A. R. & White, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6754–6758.
10. Kunkel, L. M., Smith, K. D., Boyer, S. H., Borgaonkar, D. S., Wachtel, S. S., Miller, O. J., Breg, W. R., Jones, H. W., Jr. & Rary, J. M. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 1245–1249.
11. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
12. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683–3687.
13. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
14. Laskey, R. A. (1980) *Methods Enzymol.* **65**, 363–371.
15. Calos, M. P. & Miller, J. H. (1980) *Cell* **20**, 579–595.
16. Cameron, J. R., Loh, E. Y. & Davis, R. W. (1979) *Cell* **16**, 739–751.
17. Strobel, E., Dunsmuir, P. & Rubin, G. M. (1979) *Cell* **17**, 429–439.
18. Rubin, C. M., Houck, C. M., Deininger, P. L., Friedman, T. & Schmid, C. W. (1980) *Nature (London)* **284**, 372–374.
19. Wu, J. C. & Manuelidis, L. (1980) *J. Mol. Biol.* **142**, 363–386.
20. Kan, Y. W. & Dozy, A. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5631–5635.
21. Maniatis, T., Fritsch, E. F., Lauer, J. & Lawn, R. M. (1980) *Annu. Rev. Genet.* **14**, 145–179.
22. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980) *Am. J. Hum. Genet.* **32**, 314–331.
23. Gabbay, K. H., Bergenstal, R. M., Wolff, J., Mako, M. E. & Rubenstein, A. H. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2881–2885.
24. Tager, H., Given, B., Baldwin, D., Mako, M., Markese, J., Rubenstein, A., Olefsky, J., Kobayashi, M., Kolterman, O. & Poucher, R. (1979) *Nature (London)* **281**, 122–125.
25. Gabbay, K. H. (1980) *N. Engl. J. Med.* **302**, 165–167.