Brief Genetics Report

Human Insulin Receptor Substrate-2

Gene Organization and Promoter Characterization

Lothar Vaßen, Wojciech Wegrzyn, and Ludger Klein-Hitpass

Insulin receptor substrate-2 (IRS-2) belongs to a family of cytoplasmic adaptor proteins, which link insulin, IGF-1, and cytokine receptor tyrosine kinases to signaling pathways regulating metabolism, growth, and differentiation (1–3). IRS-2–deficient mice display all characteristics of type 2 diabetes, suggesting that dysfunction of the *IRS-2* gene may contribute to the pathogenesis of human type 2 diabetes (4). Based on its progesterone inducibility, we have recently cloned and sequenced a full-length human *IRS-2* cDNA containing an open reading frame (ORF) of 4,014 bp and 5′- and 3′-untranslated regions (UTRs) of 516 and 2,466 bp (5). Although the *IRS-2* gene has previously been thought to lack introns within the coding region (6,7), the amino acid sequence predicted from our cDNA sequence differed at its very COOH-terminal end from an IRS-2 protein sequence derived from genomic *IRS-2* sequences. Therefore, we carefully analyzed the genomic structure of the *IRS-2* gene and found that the IRS-2 gene contains an intron that disrupts the ORF. Characterization of promoter and 5′-flanking regions of *IRS-2* by sequencing, reporter gene assays, and chromatin structure analysis suggests that elements conferring progesterone in-ducibility are not located immediately upstream of the gene promoter. *Diabetes* 48:1877–1880, 1999

Two phages containing genomic inserts of 11 kb (λ6.1) and 8.5 kb (λ1.1) in size were isolated from a λL47.1 library (Fig. 1). Both phages contained the entire 5′-UTR (516 bp) and nearly the complete coding region of the cDNA, followed by sequences that were not represented in the cDNA. Since this suggested the presence of an intron, we isolated additional genomic clones (BACH-8o4, BACH-222E8, BACH-228d12) from a human bacterial artificial chromosome (BAC) library. Southern hybridization confirmed that all three BACs were colinear with λ6.1 (Fig. 1). Two BAC clones (BACH-8o4 and BACH-222E8) additionally contained the entire 3′-UTR of the cDNA. Restriction fragments covering the putative splice donor and acceptor sites were subcloned and partially sequenced. Comparison of genomic and cDNA sequences revealed that the human (h)*IRS-2* gene consists of two exons separated by an intron spanning at least 4.2 kb (Fig. 1). Exon 1 (4.5 kb) contains the 5′-UTR and the ORF, except for the final two nucleotides of the last codon of the IRS-2 protein sequence. Exon 2 (2.4 kb) harbors the missing two nucleotides of the ORF, the stop codon, and the entire 3′-UTR of the message in a single exon. As indicated in Fig. 2, the intron-exon boundaries of *IRS-2* comply perfectly with the consensus sequences derived for splice donor and acceptor sites. The identification of an intron that disrupts the ORF of *IRS-2* explains, at least in part, the diversity of published *IRS-2* sequences, since one published *IRS-2* sequence contains a COOH-terminal extension of 17 amino acids due to the presence of an intron sequence (7). Inspection of a mouse *IRS-2* expressed sequence tag (GenBank accession number AA560776) and the corresponding genomic sequence reveals that the mouse *IRS-2* gene, which has previously been thought to be intronless (6), also contains a splice junction within the last codon, suggesting a similar organization of human and mouse *IRS-2* genes. Thus, future studies addressing the role of *IRS-2* mutations in diseases have to include analysis of the exon-intron borders.

The 5′-end of the *IRS-2* mRNA was determined by primer extension analysis on poly(A)$^+$ mRNA from HeLa3B2 cells (8). As shown in Fig. 3*A*, a single extension product of ~155 nucleotides in size defined a major start site at –488. Ribonuclease protection analysis with a probe covering sequences from –363 to –614 resulted in specific protected bands of 95–145 bp (Fig. 3*B*, *lanes 4* and *5*), suggesting that transcription initiation occurs at multiple sites between –457 and –502 (Fig. 4). Because our cDNA extended up to position –516 (5), there may be additional minor transcription start sites further upstream. The discrepancy between primer extension and protection analysis might be due to the higher sensitivity of the RNase protection assay. Alternatively, primer extension might be affected by extensive secondary structure of the G/C-rich sequence. While the 5′-ends of *IRS-2* transcripts were clearly detectable in RNA from R5020 stimulated cells, *IRS-2* mRNA was barely detectable in RNA from untreated cells (Fig. 3*B*). This result confirms our previous analysis with antisense probes directed against the 3′-end of the mRNA (5) and shows that R5020 regulation of *IRS-2* mRNA occurs at the transcriptional level.

From the Institut für Zellbiologie (Tumorforschung), Universitätsklinikum, Essen, Germany.

Address correspondence and reprint requests to Ludger Klein-Hitpass, Institut für Zellbiologie (Tumorforschung), Universitätsklinikum, Virchowstr. 173, D-45122, Essen, Germany. E-mail: ludger.klein-hitpass@uni-essen.de.

Received for publication 13 January 1999 and accepted in revised form 19 May 1999.

BAC, bacterial artificial chromosome; GRE, glucocorticoid response element; h, human; IRS, insulin receptor substrate; ORF, open reading frame; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; PRE, progesterone response element; SSC, sodium chloride–sodium citrate; tk, thymidine kinase; UTR, untranslated region.

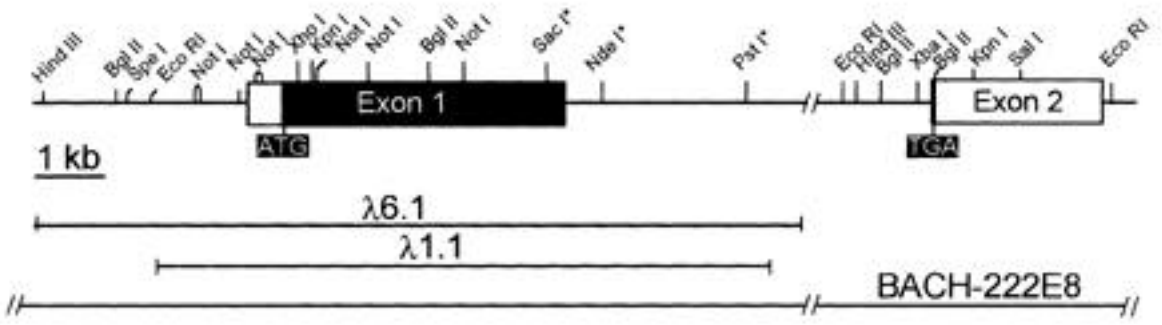

FIG. 1. Structural organization of the *hIRS-2* gene. Boxes represent exons; filled regions correspond to the coding region. The figure also illustrates the origin of the two genomic DNA fragments (λ6.1, BACH-222E8) used to derive the genomic structure. Relevant restriction sites are indicated. Restriction sites marked by an asterisk were not completely mapped.

Sequencing of the *IRS-2* promoter region revealed multiple binding sites for ubiquitously expressed transcription factors, such as Sp1, AP2, and CCAAT-box binding factor, but no TATA-box or initiator element (Fig. 4). Consistent with the observation that the *IRS-2* gene is expressed in most organs and cell types analyzed (5–7), these features suggest that *hIRS-2* is a gene with a housekeeping function. Sequences with similarity to the glucocorticoid/progesterone response element (GRE/PRE) were not found within the promoter region. Transient transfection of a luciferase reporter gene construct containing a 675-bp fragment extending from –1052 to –378 and including the transcription initiation region into two progesterone

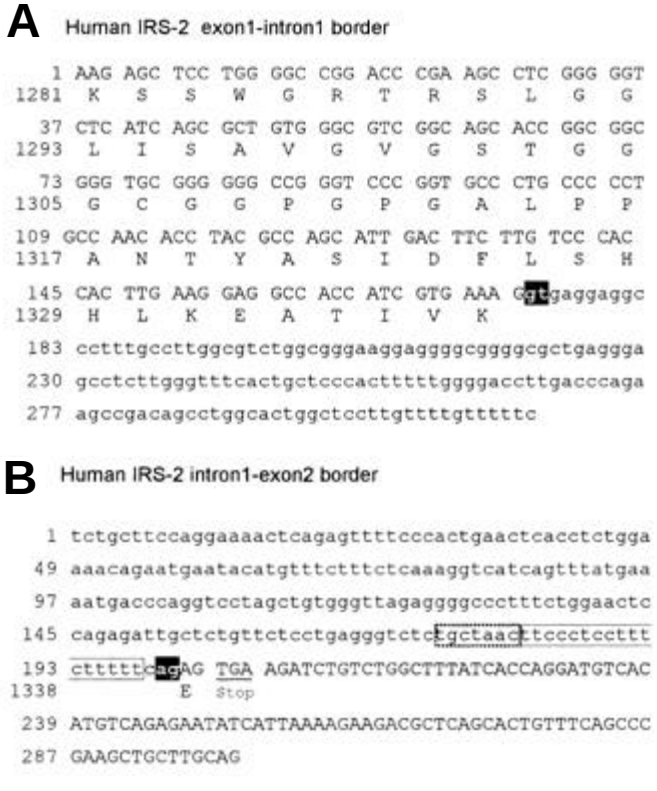

FIG. 2. DNA sequences encompassing the 5′ and 3′ boundaries of intron 1 of the *hIRS-2* gene. Numbering of amino acids refers to the cDNA sequence deposited under accession number AF073310. Exon sequences are given in capital letters, whereas intron sequences are represented in lower case letters. The deduced amino acid translations are given in one letter code below the sequences. *A*: Sequence of the exon 1–intron 1 boundary. The conserved GT dinucleotide sequence at the exon-intron border is highlighted as white characters on a black background. The nucleotide sequence has been deposited in the GenBank database with accession number AF074850. *B*: Sequence of the intron 1–exon 2 boundary. The pyrimidine-rich region and the branch point sequence are indicated by solid and dashed boxes, respectively, and the conserved AG dinucleotide is highlighted as white characters on a black background. The nucleotide sequence has been deposited in the GenBank database with accession number AF074851.

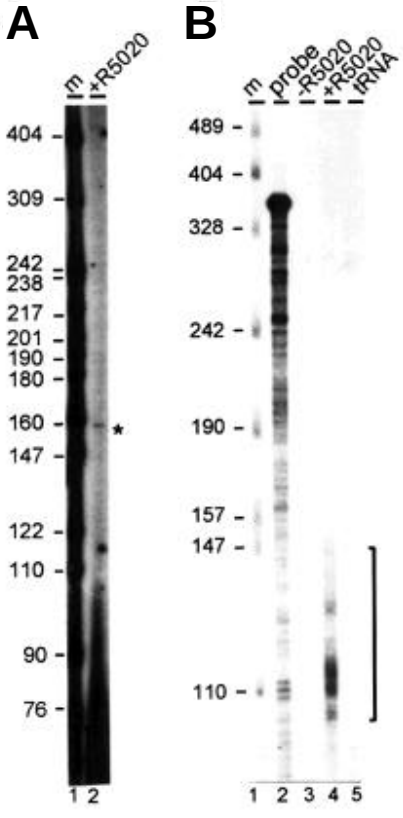

FIG. 3. Mapping of the *IRS-2* promoter. *A*: Primer extension analysis of the *hIRS-2* gene. An oligonucleotide complementary to the 5′-region was end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase. About 2 µg of poly(A)$^+$ mRNA isolated from progestin-stimulated HeLa3B2 cells was mixed with the labeled primer and analyzed as described in METHODS. *Lane m* shows the pBR322 *Hpa*II molecular weight marker. The asterisk at the right indicates the primer extension product. *B*: RNase protection analysis. RNA from untreated (*lane 3*) or R5020-treated HeLa3B2 cells (*lane 4*) was analyzed by RNase protection with a radiolabeled antisense RNA probe complementary to positions –614 to –363 of the *IRS-2* gene. The bracket at the right indicates specific signals. *Lanes 2* and *5* show the undigested probe and a control reaction with tRNA, respectively. *Lane m* shows the pBluescript *Hpa*II molecular weight marker.

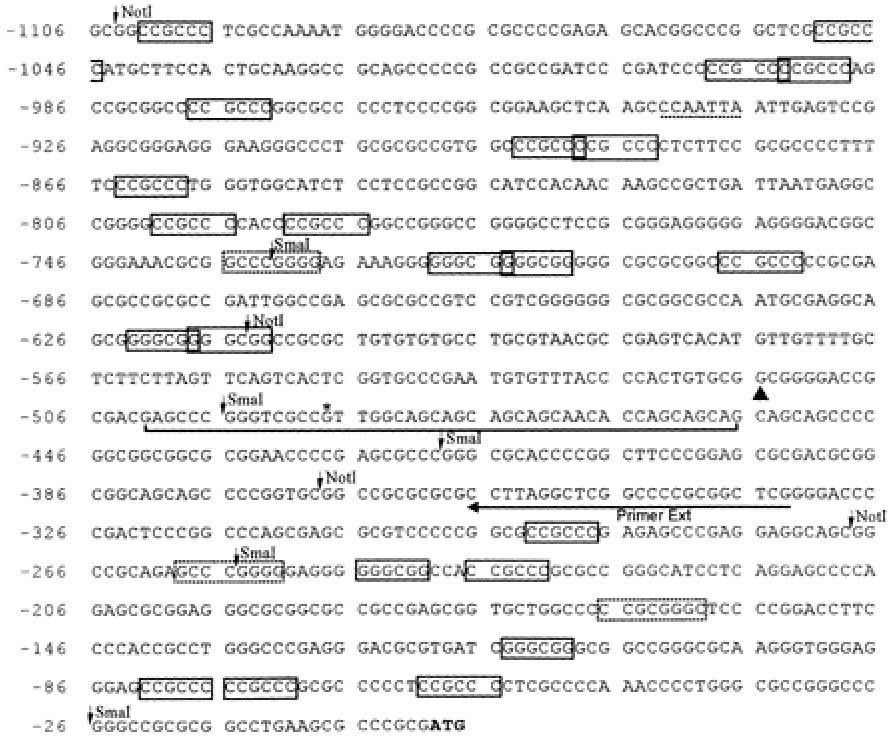

FIG. 4. Nucleotide sequence of the *hIRS-2* promoter region. The numbers at the left are the nucleotide numbers with +1 corresponding to the A of the ATG codon. Representative restriction enzyme sites are indicated by arrows. The initiator methionine codon (ATG) is indicated in bold. The 5′-end of the cloned human cDNA is marked by a filled triangle. Transcriptional start sites as mapped by primer extension and RNase protection analysis are indicated by an asterisk above and a bracket below the sequence, respectively. The primer used for primer extension (Ext) is shown by an arrow. Potential binding sites for the transcription factors SP1 (GGGCGG) and AP2 (CCSCRGGC) on both strands of the DNA are shown in solid and dashed boxes, respectively. A potential CCAAT-box (CCAATNA) is indicated by a dashed line. The nucleotide sequence has been deposited in the GenBank database with accession number AF080254.

receptor–expressing cell lines (HeLa3B2 and T47D) confirmed that these sequences exhibit promoter activity comparable to the SV40 basic promoter (Fig. 5*A*). However, treatment with the progestin R5020 did not significantly stimulate luciferase expression, while a control construct containing a single consensus GRE/PRE in front of the thymidine kinase (tk) promoter exhibited clear inducibility in HeLa3B2 cells (Fig. 5*A*). Furthermore, various restriction fragments representing 5′-flanking sequences from –3.5 to –0.7 kb proved to be unable to confer progesterone inducibility to the tk promoter (data not shown). In agreement with these functional analyses, determination of DNaseI hypersensitive sites within a fragment spanning from –2600 to +2066 revealed a number of sites between positions –1900 and –500, but treatment with R5020 did not result in the appearance of additional sites or enhancement of constitutive sites (Fig. 5*B*). Together, these data suggest that regulatory elements mediating progesterone inducibility may be located in a more distant enhancer of the gene.

Since initial studies provided no evidence that human type 2 diabetes is associated with mutations within the ORF of the *IRS-2* gene (7,9), it has been suggested that reduced expression of IRS-2 may contribute to type 2 diabetes (7). Potential mechanisms resulting in reduced expression of IRS-2 include mutations within the promoter or enhancer of the gene reducing the transcription initiation rate, but also mutations resulting in decreased stability or translation efficiency of the mRNA. Thus, our characterization of the *IRS-2* promoter and the determination of the correct gene structure, together with the cloning of the full-length cDNA, provides a framework for future studies addressing the role of *IRS-2* in the pathogenesis of type 2 diabetes.

RESEARCH DESIGN AND METHODS

Genomic DNA cloning. A human genomic library in λL47.1 was screened using a probe representing nucleotides 3601–3877 of *hIRS-2* cDNA (GenBank accession number AF073310). A human genomic library in pBeloBAC11 (GenomeSystems, St. Louis, MO) was screened by polymerase chain reaction (PCR) with forward primer 5′-TTTTGTTGAGTTTTCTTCGTTGCG-3′ and reverse primer 5′-ATCGCGTTTGGGAGGTAATAGGA-3′. Restriction fragments identified by Southern hybridization were subcloned into pBluescript II SK+. Sequencing with T3 or T7 primers and primers derived from partially determined sequences was performed using an ALF automated sequencer (Amersham-Pharmacia Biotech, Uppsala, Sweden).

Primer extension analysis. Poly(A)$^+$ mRNA was isolated using the Oligotex Direct mRNA kit (Qiagen, Hilden, Germany) from progesterone receptor–positive HeLa3B2 cells (8) treated for 6 h with progestin R5020 to increase *hIRS-2* mRNA. An oligonucleotide complementary to nucleotides –357 to –334 of *IRS-2* (Fig. 4) was end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP. Some 2 µg of poly(A)$^+$ mRNA were mixed with labeled primer (500,000 counts/min), heated to 72°C for 10 min, annealed at 55°C for 30 min, and then reverse transcribed using AMV reverse transcriptase (Boehringer, Mannheim, Germany). After purification, extension products were separated in a 6% urea-acrylamide gel in 1 × Tris boric acid EDTA buffer.

RNase protection analysis. Total RNA was isolated from untreated and R5020-treated (10 nmol/l, 4 h) HeLa3B2 cells using Trizol reagent (Gibco BRL, Rockville, MD). Radiolabeled *hIRS-2* antisense RNA was generated in vitro with T7 RNA polymerase from a linearized pBluescript II SK+ plasmid containing the 246 bp *NotI* fragment corresponding to nucleotide positions –614 to –363 upstream of the ATG initiation codon of *hIRS-2* (Fig. 4). Approximately 150,000 counts/min of the probe were hybridized with 60 µg of total RNA at 42°C in buffer containing 80% formamide, digested with RNases A and T1, and analyzed on sequencing gels as described by Kahmann et al. (8).

Reporter genes, transient transfection, and luciferase assays. To construct the *IRS-2* promoter reporter plasmid, the *hIRS-2* promoter sequence from position –1052 to –378 was amplified by PCR with primers containing extensions with *Kpn*I and *Bgl*II restriction sites. After *Kpn*I and *Bgl*II digestion, the fragment was ligated into the promoterless luciferase reporter plasmid pGL3-Basic (Clontech, Palo Alto, CA). The activity of the *IRS-2* promoter was compared with that of pGL3-Promoter plasmid (Clontech) containing the SV40 basic promoter and with PRE_1-TK-luc, which contains one GRE/PRE (AGAACAGCCTGTTCT) in front of the tk promoter (–81/+52) and the luciferase gene. HeLa3B2 and T47D cells were cultured as described (5,8). Cells were transiently transfected with 0.4 µg of reporter DNA plus 0.4 µg of pBluescript DNA per well of a six-well plate by using the diethylaminoethyl-dextran technique (10). At 6 h after the addition of DNA, cells were treated with 10 nmol/l R5020 for 16 h. Controls received an appropriate amount of vehicle. Harvesting of cells, extract preparation, and luciferase analysis were performed as described (8).

Isolation of nuclei and mapping of DNaseI hypersensitive sites. HeLa3B2 cells (~50% confluent) were treated for 1.5 h with 10 nmol/l R5020 or vehicle. For

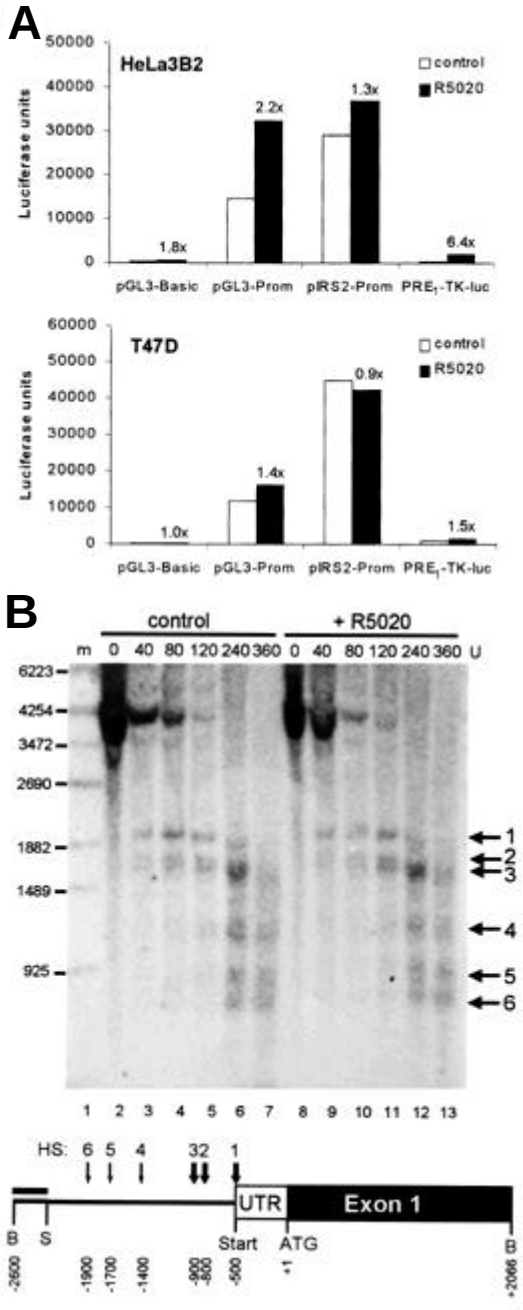

FIG. 5. *IRS-2* promoter and 5′-flanking sequences are unable to mediate progesterone inducibility. *A*: Promoter reporter analysis in HeLa3B2 and T47D. PR-positive HeLa3B2 (*upper panel*) and T47D cells (*lower panel*) were transiently transfected with the indicated luciferase reporter constructs and incubated for 16 h with vehicle (□) or 10 nmol/l R5020 (■). The luciferase activities measured in cell extracts represent the mean value of three independent determinations. Fold inductions generated by R5020 are given above the bars. *B:* Mapping of DNaseI hypersensitive sites. Nuclei were isolated from HeLa3B2 cells cultured in the absence (*lanes 2–7*) or presence (*lanes 8–13*) of 10 nmol/l R5020 and digested with the indicated amount of DNaseI. After *Bgl*II digestion, the DNA was separated by 0.8% agarose gel electrophoresis, blotted, and indirectly end-labeled at the *Bgl*II site at position –2600 of the *IRS-2* gene. *Lane m* contains end-labeled λ*Sty*I marker fragments whose sizes in base pairs are indicated. Fragments indicating hypersensitive sites 1–6 are indicated by arrows. In the schematic representation below the autoradiogram, the positions of the hypersensitive sites of the *hIRS-2* gene are given. The hybridization probe is indicated by a horizontal bar. B, *Bgl*II; S, *Spe*I.

the isolation of nuclei, ~5×10^7 cells were washed two times with ice-cold phosphate-buffered saline (PBS) and scraped into PBS. After centrifugation, the cells were resuspended in 15 ml of homogenization buffer (0.3 mol/l sucrose, 60 ml KCl, 15 ml NaCl, 15 ml Tris-HCl [pH 7.5], 5 mmol/l $MgCl_2$, 0.5 mmol/l EGTA, 0.2 mmol/l 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride [AEBSF]) per 10-cm dish and lysed by addition of an equal volume of homogenization buffer containing 1% NP40. After incubation for 5 min on ice, nuclei were collected by centrifugation at 3,000 rpm for 5 min, washed twice with homogenization buffer without NP40, and finally resuspended in 2 ml of DNaseI buffer (60 mmol/l KCl, 15 ml NaCl, 15 ml Tris-HCl [pH 7.5], 5 mmol/l $MgCl_2$, 0.2 mmol/l EGTA, 0.2 mmol/l AEBSF). Some 300 µl of nuclei each were digested with 0, 40, 80, 120, 240, and 360 U DNaseI (Grade II; Boehringer) for 15 min on ice. Reactions were stopped by addition of 1.7 ml of buffer containing 60 mmol/l Tris-HCl (pH 8.0), 25 mmol/l EDTA, 1.2% SDS, and 400 µg proteinase K and incubated at 37°C for 16 h. After phenol extraction and precipitation, DNA was dissolved in 400 µl Tris-EDTA (TE) per sample and treated with RNase A (20 µg) for 30 min at 37°C, precipitated again, and dissolved in 200 µl of TE (10 mmol/l Tris-HCl [pH 7.5], 1 mmol/l EDTA). Some 20 µg of DNA were digested for 6 h with 30 U of *Bgl*II in a final volume of 200 µl, precipitated, dissolved in 20 µl of TE, and resolved by electrophoresis on an agarose gel. Southern blots were hybridized with a radioactively labeled 332-bp *Bgl*II/*Spe*I fragment, which indirectly end-labels a 4.7-kb *Bgl*II fragment at position –2600. Labeling of the hybridization probe was performed using the Rediprime kit and [α-^{32}P]dCTP (Amersham-Pharmacia Biotech). Hybridization was performed for 16 h at 60°C in 0.25 mol/l Na_2HPO_4 (pH 7.2), 1 mmol/l EDTA, 20% SDS, and 0.5% blocking reagent. Blots were washed consecutively with 2 × sodium chloride–sodium citrate (SSC), 1 × SSC, and 0.5 × SSC in 0.1% SDS at 60°C and exposed to BIOMAX-MS films (Kodak, Rochester, NY) at –70°C.

ACKNOWLEDGMENTS

This work was supported by Bundesministerium fuer Bildung und Forschung Grant 0310780 (to L.K.-H.). L.V. is a postdoctoral fellow of the Schering Research Foundation (Berlin). GenBank accession nos. included AF074850, AF074851, and AF080254.

We are grateful to B. Horsthemke (Essen, Germany) for providing a phage library, Ulla Schmücker for sequencing, and Diemo Michels for a critical reading of the manuscript.

REFERENCES

1. White MF: The insulin signalling system and the IRS proteins. *Diabetologia* 40:S2–S17, 1997
2. Waters SB, Pessin JE: Insulin receptor substrate 1 and 2 (IRS1 and IRS 2): what a tangled web we weave. *Trends Cell Biol* 6:1–4, 1997
3. Sun XJ, Wang LM, Zhang Y, Yenush L, Myers MG Jr, Glasheen E, Lane WS, Pierce JH, White MF: Role of IRS-2 in insulin and cytokine signaling. *Nature* 377:173–177, 1995
4. Withers DJ, Gutierrez JS, Towery H, Burks DJ, Ren JM, Previs S, Zhang Y, Bernal D, Pons S, Shulman GI, Bonner-Weir S, White MF: Disruption of IRS-2 causes type 2 diabetes in mice. *Nature* 391:900–904, 1998
5. Vaßen L, Wegrzyn W, Klein-Hitpass L: Human insulin receptor substrate-2 (IRS-2) is a primary progesterone response gene. *Mol Endocrinol* 13: 485–494, 1999
6. Sun XJ, Pons S, Wang LM, Zhang Y, Yenush L, Burks D, Myers MG Jr, Glasheen E, Copeland NG, Jenkins NA, Pierce JH, White MF: The IRS-2 gene on murine chromosome 8 encodes a unique signaling adapter for insulin and cytokine action. *Mol Endocrinol* 11:251–262, 1997
7. Bernal D, Almind K, Yenush L, Ayoub M, Zhang Y, Rosshani L, Larsson C, Pedersen O, White MF: Insulin receptor substrate-2 amino acid polymorphisms are not associated with random type 2 diabetes among Caucasians. *Diabetes* 47:976–979, 1998
8. Kahmann S, Vaßen L, Klein-Hitpass L: Synergistic enhancement of PRB-mediated RU486 and R5020 agonist activities through cyclic adenosine 3′,5′-monophosphate represents a delayed primary response. *Mol Endocrinol* 12:278–289, 1998
9. Kalidas K, Wasson J, Glaser B, Meyer JM, Duprat LJ, White MF, Permutt MA: Mapping of the human insulin receptor substrate-2 gene, identification of a linked polymorphic marker and linkage analysis in families with type II diabetes: no evidence for a major susceptibility role. *Diabetologia* 41:1389–1391, 1998
10. Cato ACB, Miksicek R, Schütz G, Arnemann J, Beato M: The hormone regulatory element of mouse mammary tumour virus mediates progesterone induction. *EMBO J* 5:2237–2240, 1986