

22 using the hybrids described above (Fig. 1B) led to the assignment of the HIR gene to bin 15, a subregion of 22q13.1.

The somatic cell hybrid mapping results were independently confirmed by fluorescence *in situ* hybridization (FISH) using a cosmid specific for the HIR locus (Fig. 2). A gridded, chromosome 22-only cosmid library (LL22NC03) was screened by colony hybridization using the 1.9-kb HIR cDNA to obtain HIR-specific cosmids. Twelve positive clones were identified in the primary screen, and eight of these were confirmed in a secondary screen using PCR primers specific for the HIR gene. Two cosmids used for the FISH studies gave similar results, which agreed with FISH studies using the 1.9-kb cDNA probe. The results shown in Fig. 2A were performed using cosmid N61E10.

From physical mapping, bin 15 appears large, with over 50 markers mapping to this region (Budarf *et al.*, submitted). A YAC contig extending from bin 12 into bin 15 has been constructed (1). Colony hybridization screening of chromosome 22-specific YACs led to the identification of a CEPH mega-YAC, 849E1, positive for the HIR gene. This YAC was positive for four additional loci: D22S272, D22S292E, D22S412E, and D22S792. The latter three markers have also been uniquely assigned to bin 15. These results have allowed us to map the HIR gene finely and provide physical reagents for further molecular studies of the HIR genomic locus. In addition, the localization of the HIR gene will facilitate investigation of potential candidates for inherited cardiovascular and neurological disorders.

ACKNOWLEDGMENTS

We thank Dr. P. de Jong at the Lawrence Livermore National Laboratory for providing the chromosome 22 cosmid library, Terrence McDonald, Yvonne Tatsamura, Stacey Gavigan, and Carolyn Radeke for technical support, and Robert Burnham for help preparing the figures. This work was supported by National Institutes of Health Grants HL41656 (C.V.), CA39926 (M.B.), and HG00425 (M.B.) and the American Heart Association, California Affiliate (C.V.) B.L.B. was supported by NIH Training Grant HD07107, and F.P. by National Multiple Sclerosis Society postdoctoral fellowship.

REFERENCES

- Bell, C. J., Budarf, M. L., Nieuwenhuijsen, B. W., Barnoski, B. L., Buetow, K. H., Campbell, K., Colbert, A., Collins, J., Desjardins, P. R., DeZwaan, T., Eckman, B., Foote, S., Hart, K., Hiester, K., Van Het Hoog, M. J., Hopper, E., Kaufman, A., McDermid, H. E., Overton, G. C., Reeve, M. P., Searls, D. B., Stein, L., Valmiki, V. H., Watson, E., Williams, S., Winston, R., Nussbaum, R. L., Lander, E. S., Fischbeck, K., Emanuel, B. S., and Hudson, T. J. (1995). Integration of physical, breakpoint and genetic maps of chromosome 22: Localization of 587 yeast artificial chromosomes with 238 mapped markers. *Hum. Mol. Genet.* **4**: 59–69.
- Budarf, M., Sellinger, B., Griffin, C., and Emanuel, B. S. (1989). Comparative mapping of the constitutional and tumor-associated 11;22 translocations. *Am. J. Hum. Genet.* **45**: 128–139.
- Croce, C. M., Huebner, K., Isobe, M., Fainstain, E., Lifshitz, B., Shtivelman, E., and Canaani, E. (1987). Mapping of four distinct BCR-related loci to chromosome region 22q11: Order of BCR loci relative to chronic myelogenous leukemia and acute lymphoblastic leukemia breakpoints. *Proc. Natl. Acad. Sci. USA* **84**: 7174–7178.
- Delattre, O., Azambuja, C. J., Aurias, A., Zucman, J., Peter, M., Zhang, F., Hors-Cayla, M. C., Rouleau, G., and Thomas, G. (1991). Mapping of human chromosome 22 with a panel of somatic cell hybrids. *Genomics* **9**: 721–727.
- Frazer, K. A., Boehnke, M., Budarf, M. L., Wolff, R. K., Emanuel, B. S., Myers, R. M., and Cox, D. R. (1992). A radiation hybrid map of the region on human chromosome 22 containing the neurofibromatosis type 2 locus. *Genomics* **14**: 574–584.
- Geurts van Kessel, A. H. M., Westerveld, A., de Groot, P. G., Meera Khan, P., and Hagemeijer, A. (1980). Regional localization of the genes coding for human ACO2, ARSA, and NAGA on chromosome 22. *Cytogenet. Cell Genet.* **28**: 169–172.
- Lichter, P., Ledbetter, S. A., Ledbetter, D. H., and Ward, D. C. (1990). Fluorescence *in situ* hybridization with Alu and L1 polymerase chain reaction probes for rapid characterization of human chromosomes in hybrid cell lines. *Proc. Natl. Acad. Sci. USA* **87**: 6634–6638.
- Makhina, E. N., Kelly, A. J., Lopatin, A. N., Mercer, R. W., and Nichols, C. G. (1994). Cloning and expression of a novel human brain inward rectifier potassium channel. *J. Biol. Chem.* **269**: 20468–20474.
- Périer, F., Radeke, C. M., and Vandenberg, C. A. (1994). Primary structure and characterization of a small-conductance inwardly rectifying potassium channel from human hippocampus. *Proc. Natl. Acad. Sci. USA* **91**: 6240–6244.
- Tang, W., and Yang, X.-C. (1994). Cloning a novel human brain inward rectifier potassium channel and its functional expression in *Xenopus* oocytes. *FEBS Lett.* **348**: 239–243.
- Vandenberg, C. A. (1994). Inward rectifier potassium channels. In "Ion Channels in the Cardiovascular System: Function and Dysfunction" (P. M. Spooner, A. M. Brown, W. A. Catterall, G. J. Kaczorowski, and H. C. Strauss, Eds.), pp. 145–167, Futura Publ. Co., Armonk.

Human Protein Kinase C Iota Gene (PRKCI) Is Closely Linked to the BTK Gene in Xq21.3

Richard Mazzearella,* Alfredo Ciccodicola,†‡
Teresa Esposito,† Alessandro Arcucci,†
Carmela Migliaccio,† Carmela Jones,*
David Schlessinger,* Michele D'Urso,†
and Maurizio D'Esposito¹

*Washington University School of Medicine, St. Louis, Missouri;
†International Institute of Genetics and Biophysics, CNR, Naples,
Italy; and ‡Dipartimento Scienze Morfologiche e Medico-Legali,
Università degli Studi di Modena, Italy

Received June 20, 1994; accepted November 28, 1994

The human X chromosome contains many disease loci (10), but only a small number of X-linked genes have been cloned and characterized. One approach to finding genes in genomic

¹ To whom correspondence should be addressed.

DNA uses partial sequencing of random cDNAs to develop "expressed sequence tags" (ESTs) (1).

Many authors have recently reported chromosomal localization of such ESTs using hybrid panels (4, 8, 9). Twenty ESTs specific for the X chromosome have been localized to defined regions with somatic cell hybrids (6, 7), and 12 of them have been physically linked to markers that detect polymorphisms (6). One of these ESTs, EST02087 (also known as DXS1179E), was physically linked in a 650-kb contig to the GLA (α -galactosidase) gene involved in Fabry disease (2). A comparison of this contig with a 7.5-Mb YAC contig (13, 14) indicated that this gene is also within 250 kb of the src-like protein-tyrosine kinase BTK (X-linked agammaglobulinemia protein-tyrosine kinase) gene in Xq21.3 (Fig. 1A).

Using a PCR-based screening procedure (3) and an STS specific for EST02087, three positive pools in a human teratocarcinoma cDNA library (12) were identified. A λ phage clone of about 600 bp was identified from one of the pools and was subsequently used as a hybridization probe to identify the

clones from the two other pools. *EcoRI* digests of these two cDNA clones showed them to be about 1.75 and 2.3 kb. The largest cDNA clone was chosen for subsequent studies.

Southern analyses of the YACs that were shown to be positive for the EST02087 STS (6), using the EST02087 cDNA as a probe, all confirm the localization of this cDNA. A single band of 7.4 kb was detected by this cDNA probe in the five YACs known to contain EST02087 in total human DNA and in DNA from a cell line containing five X chromosomes. Non-overlapping YACs were negative for the probe, as expected (data not shown).

The complete sequence for EST02087 was deduced from overlapping fragments of the cDNA subcloned into pGEM-4Z vector (Promega Biotech). The nucleotide sequence was determined using a fluorescent-labeling modified method of the dideoxy chain reaction (ABi sequencing manual) with the vector sequencing primers. Sequencing reactions were performed according to a cycle-sequencing protocol (ABi sequencing manual) on the 9600 apparatus (Perkin-Elmer). All four reactions for each sample were pooled, precipitated with ethanol, and electrophoresed in an ABI 373A DNA sequencer. The sequence was assembled using the ABD INHERIT program.

The sequence of the EST02087 cDNA was determined to be 2261 bp. This sequence has been assigned GenBank Accession No. L33881, and the symbol assigned to this gene is PRKCI. This cDNA appears to encode a protein of 587 amino acids with $M_r = 67,258$. The first ATG is a poor initiation codon, whereas the second ATG, which is in the same frame, closely matches the consensus (5). N-terminal protein sequencing will be required to confirm the translation start site. This sequence is polyadenylated after the G at base 2261 and contains the putative poly(A) addition signal AATGAA 35 bases upstream from the site of poly(A) addition.

A GenBank search revealed that a similar human cDNA sequence was recently identified as an iota isoform of protein kinase C (11). A comparison of the two sequences shows that they are completely identical in coding region, although considerable differences are seen in the putative 5' and 3' untranslated region. Both cDNAs are also polyadenylated at their 3' ends. There are eight nucleotide differences in the 204-bp overlap upstream of the start codon, and the earlier sequence contains an additional 62 bp of 5' untranslated DNA. In the 3' untranslated region, 10 of the last 11 overlapping bases are different, and an additional 201 bp are seen 5' to the poly(A) tail in the sequence reported here. The differences probably do not result from a cocloning event, since genomic Southern analysis with the entire cDNA as a probe detects only a single 7.4-kb band. Rather, the two cDNA sequences may arise from a single gene by differential splicing in the different cells of origin (i.e., teratocarcinoma and insulin-secreting cells).

The relative expression levels of the cDNA were examined in a multiple tissue Northern analysis (Fig. 1B). After standardization for the amount of messenger RNA using an actin probe, the cDNA detected highly expressed RNA species in kidney, muscle, lung, and brain, moderately expressed RNAs in pancreas and placenta, and lower expression in heart tissue. The predominant form of hybridizing mRNA is about 6 kb; in addition, forms of 2.8 and 2.4 kb are also detected. The cDNA sequenced here presumably corresponds to the 2.4-kb species observed in placenta and brain. These observations suggest that a number of RNA species hybridize to the cDNA reported here. It may be an example of duplication and divergence, since there may be other members of the same isoform.

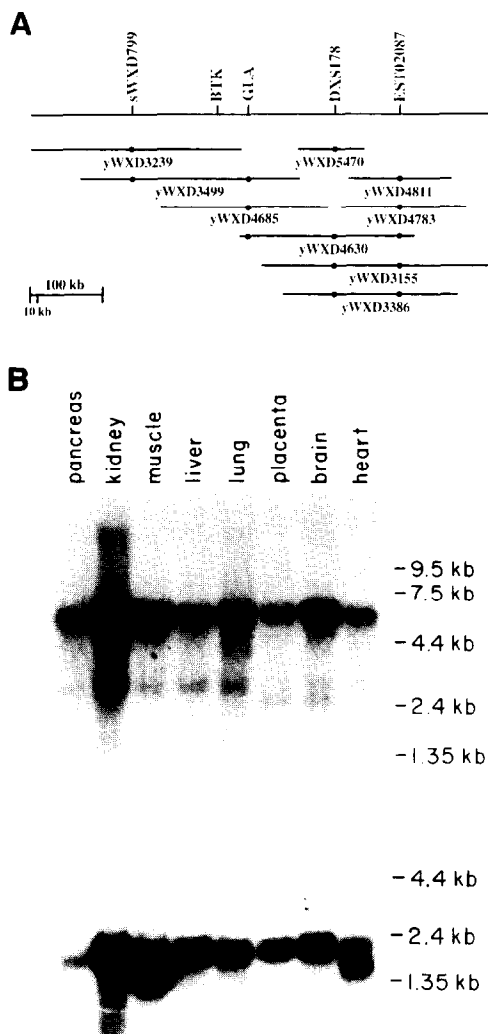


FIG. 1. EST02087 cDNA. (A) Physical distance of EST02087 to known X-chromosomal markers and loci in Xq21.3. (B) (Top) Northern analysis of EST02087. A multiple tissue mRNA blot (Clontech) was hybridized to the EST02087 cDNA probe. (Bottom) Actin probe. The blots were washed twice in $2\times$ SSC, 0.5% SDS at room temperature followed by two washes of $0.2\times$ SSC, 0.1% SDS at 50°C or, in the case of actin, at 65°C .

Protein kinases mediate the control of cellular responses to external stimuli through the phosphorylation of target proteins. A defect in the protein-tyrosine kinase BTK gene results in the immunodeficient syndrome agammaglobulinemia. The BTK cDNA and the cDNA reported here have some functional regions in common but belong to different protein kinase families. Thus, the two gene products may have no functional relation, although their relative proximity on the X chromosome is intriguing.

ACKNOWLEDGMENTS

We acknowledge M. Terracciano for her helpful technical assistance. This work was supported by NIH Grant HG00247 to D. Schlessinger and funds from CNR Progetto Finalizzato "Ingegneria Genetica" and Telethon Italia (Grant 417) and the Italian Association for Cancer Research (AIRC) to M. D'Urso.

REFERENCES

- Adams, M. D., Dubnick, M., Kerlavage, A., Moreno, R., Kelley, J., Utterback, T., Nagle, J., Fields, C., and Venter, J. C. (1992). Sequence identification of 2,375 human genes. *Nature* **335**: 632–634.
- Bernstein, H. S., Bishop, D. F., Astrin, K. H., Kornreich, R., Eng, C. M., Sakuraba, H., and Desnick, R. J. (1989). Fabry disease: Six gene rearrangements and an exonic point mutation in the α -galactosidase A gene. *J. Clin. Invest.* **83**: 1390–1399.
- D'Esposito, M., Mazzarella, R., Pengue, G., Jones, C., D'Urso, M., and Schlessinger, D. (1994). PCR-based immortalization and screening of hierarchical pools of cDNAs. *Nucleic Acids Res.* **22**: 4806–4809.
- Durkin, A. S., Maglott, D. R., and Niermann, W. C. (1992). Chromosomal assignment of 38 human brain expressed sequence tags (ESTs) by analyzing fluorescently labeled PCR products from hybrid cell panels. *Genomics* **14**: 808–810.
- Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **47**: 1–45.
- Mazzarella, R., and Srivastava, A. K. (1994). Physical linkage of expressed sequence tags (ESTs) to polymorphic markers on the X chromosome. *Hum. Mol. Genet.* **3**: 1095–1101.
- Parrish, J., and Nelson, D. (1993). Regional assignment of 19 X-linked ESTs. *Hum. Mol. Genet.* **2**: 1901–1905.
- Polymeropoulos, M. H., Xiao, H., Glodek, A., Gorski, M., Adams, M. D., Moreno, R. F., Fitzgerald, M. G., Venter, J. C., and Merrill, C. R. (1992). Chromosomal assignment of 46 brain cDNAs. *Genomics* **12**: 492–496.
- Polymeropoulos, M. H., Xiao, H., Sikela, J. M., Adams, M., Venter, J. C., and Merrill, C. R. (1993). Chromosomal distribution of 320 genes from a brain cDNA library. *Nature Genet.* **4**: 381–386.
- Schlessinger, D., Mandel, J.-L., Monaco, A. P., Nelson, D. L., and Willard, H. F. (1993). Report of the fourth international workshop on human X chromosome mapping 1993. *Cytogenet. Cell Genet.* **64**: 147–194.
- Selbie, L. A., Schmitz-Peiffer, C., Sheng, Y., and Biden, T. J. (1993). Molecular cloning and characterization of PKC ϵ , an atypical isoform of protein kinase C derived from insulin-secreting cells. *J. Biol. Chem.* **268**: 24296–24302.
- Skowronski, J., Fanning, T. G., and Singer, M. F. (1988). Unit-length line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol.* **8**: 1385–1397.
- Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C. I. E., and Bentley, D. R. (1993). The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**: 226–233.
- Vetrie, D., Kendall, E., Coffey, A., Hassock, S., Collins, J., Todd, C., Lehrach, H., Bobrow, M., Bentley, D. R., and Harris, A. (1994). A 6.5-Mb yeast artificial chromosome contig incorporating 33 DNA markers on the human X chromosome at Xq22. *Genomics* **19**: 42–47.

Human Sp4 Transcription Factor Gene (SP4) Maps to Chromosome 7p15

Martha Kalff-Suske,^{*1} Jürgen Kunz,^{*}
Karl-Heinz Grzeschik,^{*} and Guntram Suske[†]

^{*}Medizinisches Zentrum für Humangenetik, Philipps-Universität Marburg, Bahnhofstrasse 7, 35037 Marburg, Germany; and [†]Institut für Molekularbiologie und Tumorforschung, Philipps-Universität Marburg, Emil-Mannkopff-Strasse 2, 35037 Marburg, Germany

Received November 2, 1994; accepted January 31, 1995

Transcriptional regulation is exerted through protein factors that bind to distinct DNA elements in the vicinity of target genes. Many control elements are recognized by different members of families of transcription factors opening combinatorial possibilities of the regulatory interactions. Among the most widely distributed control elements are GC boxes and related motifs bound by Sp1 (6). Until recently, Sp1 was thought to be the only factor acting through these motifs. The cloning of cDNAs encoding the DNA-binding proteins Sp3 (4, 8) and Sp4 (4), however, uncovered the existence of a small multigene family. Its members exhibit very similar structural features. Like Sp1, Sp3 and Sp4 contain DNA-binding domains composed of three zinc fingers of the Cys₂/His₂ type near the C-terminus that recognize the GC box and another related motif with identical affinities (4, 5). The N-termini of all three proteins are composed of long serine/threonine- and glutamine-rich regions that have been identified in Sp1 as transactivation domains (2, 7). Similarity in function suggested by the high degree of structural conservation was recently shown for Sp4. Sp4 acts as a positive regulator by stimulating gene expression from Sp1-responsive promoters (5). The frequency of Sp binding sites in the vicinity of genes as well as the restricted expression pattern of Sp4 transcripts *in vivo*, in contrast to the ubiquitous Sp1 and Sp3 mRNA expression (4), prompt the question of the particular regulatory roles of these transcription factors that bind to the same DNA elements.

¹To whom correspondence should be addressed. Telephone: +49 (6421) 283584. Fax: +49 (6421) 288920.

GENOMICS **26**, 631–633 (1995)
0888-7543/95 \$6.00

Copyright © 1995 by Academic Press, Inc.
All rights of reproduction in any form reserved.