DNA RESEARCH 5, 309–313 (1998) Short Communication

cDNA Cloning, Tissue Distribution and Chromosomal Localization of the Human ID4 Gene

Muriel RIGOLET,[1,2,*] Tina RICH,[2] Marie-Sylvie GROSS-MORAND,[1] Denise MOLINA-GOMES,[1] Evani VIEGAS-PEQUIGNOT,[1] and Claudine JUNIEN[1]

U383 INSERM, Clinique Maurice Lamy, Hôpital Necker, 149 rue de Sèvres, Université ParisV, 75015 Paris, France[1] *and Preuss Laboratory, Brain Tumor Research Center, Department of Neurological Surgery, San Francisco, CA 94143, USA*[2]

(Received 10 July 1998)

Abstract

A cDNA encoding the human Id4 protein has been isolated from an astrocytoma library. The predicted protein product shares 98% identity with the mouse Id4 protein and is markedly different from that already reported. By FISH analysis, the human ID4 gene was more precisely mapped to chromosome 6p22.3-p23. Northern blot analysis showed that ID4 is mainly expressed in thyroid, brain and fetal tissue and in some nervous system tumor cell lines.

Key words: ID gene; HLH protein; Brain tumors

The helix-loop-helix (HLH) proteins are transcriptional regulators which have been shown to play a major role in cell growth, differentiation and oncogenesis.[1–4] These proteins are involved in the development of several lineages, including β cells of the pancreas,[5] B lymphocytes[6] and cells from muscle and nervous system tissues.[7,8] They are subdivided into two groups depending on their ability to bind DNA.[9] The basic helix-loop-helix proteins (bHLH) bind as dimers to a specific DNA enhancer element known as the E-box (CANNTG) and induce gene expression in their target tissue. Members of the HLH protein family which contain no DNA-binding basic domain inactivate the transcriptional activity of bHLH proteins through the formation of inactive heterodimer complexes with bHLH partners. This property has led to their designation as inhibitors of DNA binding or Id-HLH proteins. Four genes of this group have been cloned thus far in mouse and human: Id1, Id2, Id3 and Id4.[2,10–17] Of Id genes, the role of Id1 has been the most extensively studied and to a lesser extent Id2. It has been reported that both proteins induce cell proliferation. Id1 antagonizes the action of the E proteins, and Id2 acts through binding to pRB (retinoblastoma) and related proteins p107 and p130.[1–4,18] Id genes are also involved in the regulation of cell differentiation. Several studies have shown that over-expression of Id1 inhibits the differentiation of many lineages.[10,11,19–21] Since dysregulation of developmental genes has often been found to be associated with malignancies, we were interested in cloning new ID family members expressed in the developing nervous system in order to evaluate the role of this class of genes in nervous system tumors. Here we report the isolation, the pattern of expression in normal tissues and tumor cell lines, and the chromosomal localization of the human ID4 gene.

Computer analysis of the protein products of the ID gene family revealed that the sequence corresponding to the HLH domain is highly conserved through evolution. To isolate novel ID family members preferentially expressed in the nervous system, we designed polymerase chain reaction (PCR) primers complementary to this conserved domain. The reverse primer comprised a degenerated sequence complementary to the helix II region (5′ ATRTARIATIACRTG), where I is for inosine and R is for A or G. The forward primer was complementary to a three-amino-acid motif (LVP), at the helix I-loop junction (5′ ACTGGTGCCC). Fetal brain RNAs were amplified by reverse transcriptase (RT)-PCR and the PCR products were cloned and sequenced. Several clones showed 100% identity at the amino acid level with the HLH domain of the mouse Id4 gene product. To isolate the full length human ID4 cDNA, one of these latter clones was used as a probe to screen the astrocytoma U251 λgt10 library. Screening of 10^6 individual clones resulted in the isolation of a 1147-bp cDNA which was found to include an open reading frame (ORF) of 161 amino acids (Fig. 1). The assigned ATG start codon was found to agree with the Kozak consensus for translation initiation.[22] Nucleotide sequence analysis showed that the human ID4 gene was 92.3% identical to the murine

Communicated by Yoshiyuki Sakaki

* To whom correspondence should be addressed. Tel. +33-44-49-44-86, Fax. +33-47-83-32-06, E-mail: rigolet@lovelace.infobiogen.fr

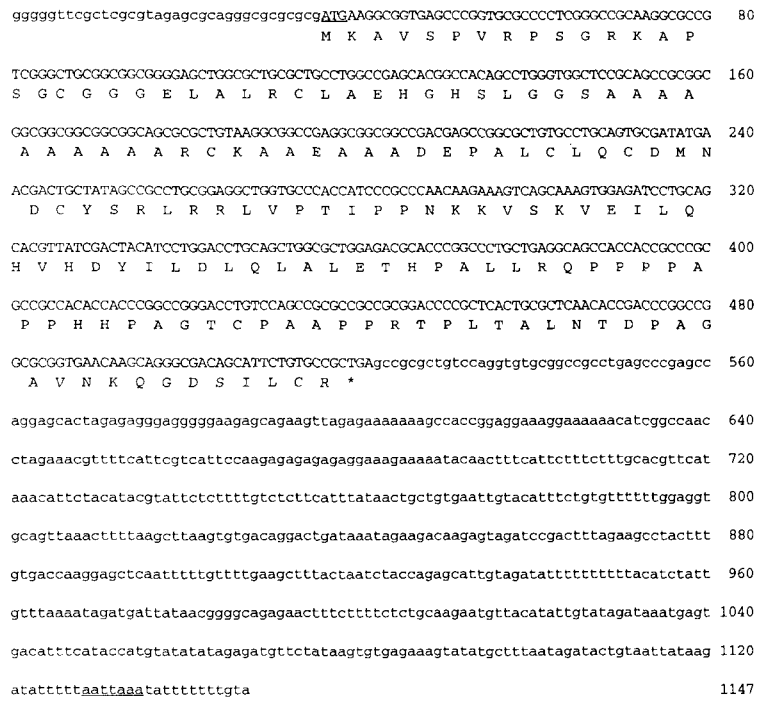

Figure 1. Nucleotide and deduced amino acid sequences of the human ID4 cDNA. Nucleotide numbers are indicated, at the end of each line. The ATG initiation codon and the polyadenylation site are underlined. The sequence has been deposited in EMBL Sequence Database under Accession No: Y07958.

Id4 gene across the coding region and 82.4% across the 3′ untranslated region. The highly GC-rich sequence of the ID4 cDNA contained a CpG-island covering the 5′ end and most of the ORF, which could be involved in the transcriptional regulation of the gene.[23]

During the course of this work, two other human ID4 cDNA sequences were deposited in Genbank. The first one, (accession number: U28368), covering nucleotide 1 to 1030, is 100% identical to our sequence. The other,[17] encompassing nucleotide 1 to 756, shows substential differences. Sequence alignment of the predicted human and mouse Id4 proteins is shown in Fig. 2. While the predicted protein products of the human (our data and accession number: U28368) and mouse Id4 c-DNAs share 98% identity and are absolutely identical in the N-terminal and in the highly conserved HLH domains, the previously reported human *ID4* protein by Pagliuca et al., shares 91% identity with that of the mouse, and most of the amino acid divergences (7 out of 9) lie in important functional domains: the N-terminal and the HLH domains. As suggested and then shown by Nagata et al.,[24] Id proteins are phosphoproteins and phosphorylation is assumed to regulate their function. The longest part of sequence divergence lies in the N-terminus of the protein, between amino acid positions 9 and 15, disrupting two phosphorylation sites for cAMP-dependent kinase and protein kinase C, which are as well conserved in the other Id proteins, Id1, Id2 and Id3 from rat, mouse, and human. In the HLH domain, two other amino acid

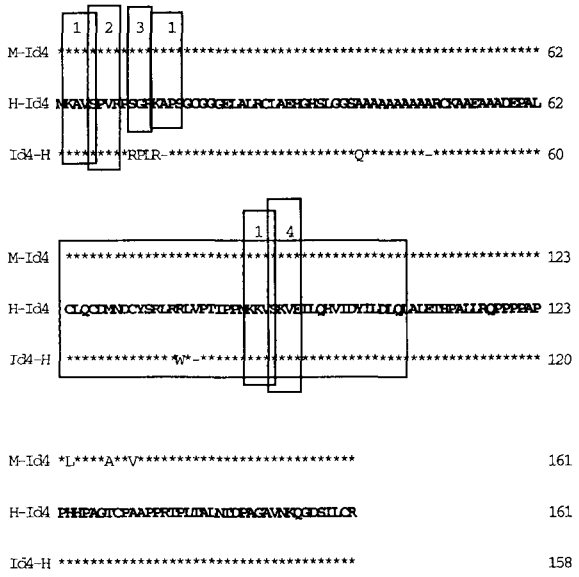

Figure 2. Amino acid sequence comparison of the mouse[13] and human Id4 gene product. H-Id4: our data and accession number U28368, Id4-H (17). Amino acids are indicated by single-letter amino acid letter abbreviations. Amino acids identical to those in H-Id4 are marked by asteriks. Hyphens represent gaps introduced to optimize the alignment. The HLH domain is boxed. The consensus sequences for phosphorylation sites are indicated by numbered boxes: 1 is for cAMP-dependent kinase; 2, 3, and 4 are for protein kinase C, casein kinase II and cdc 2 kinase, respectively. Amino acid numbers are indicated at the end of each line.

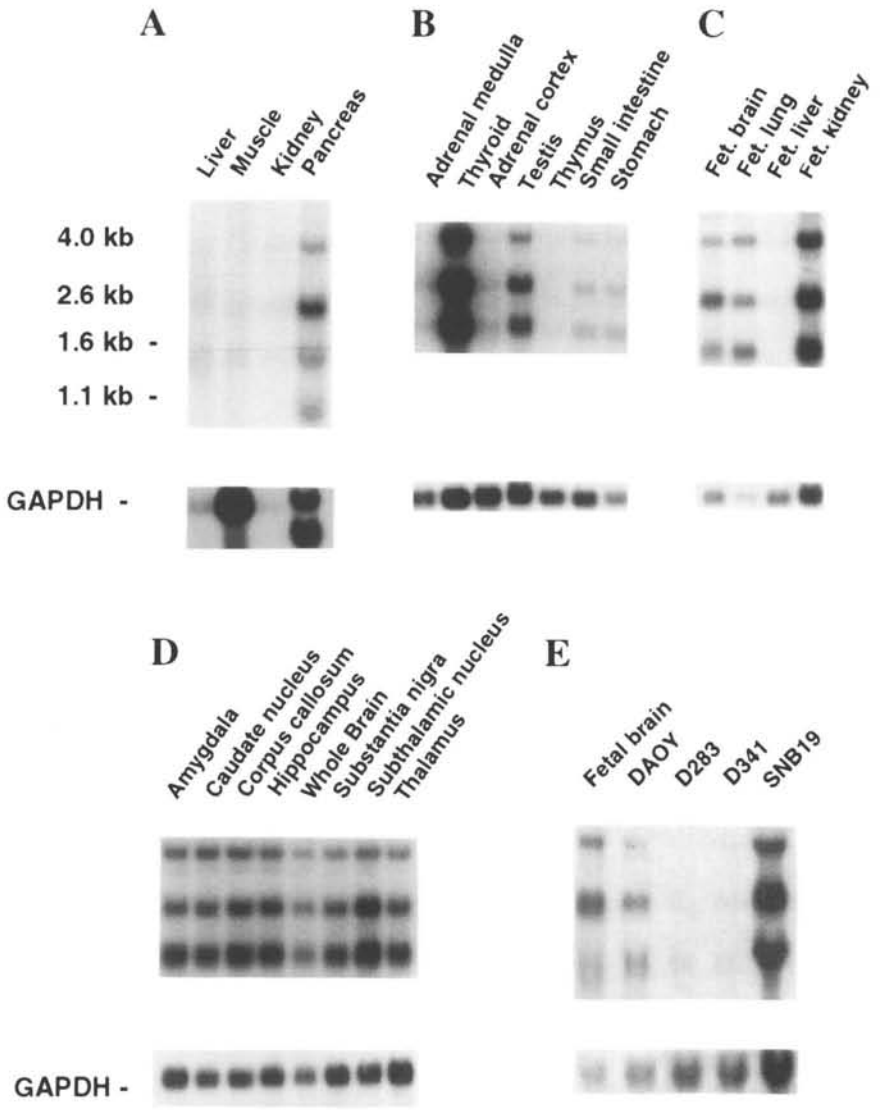

Figure 3. Northern blot analysis of ID4 expression. Human tissue blots were purchased from Clontech (A, B, C, D). Total RNA (15 μg) from fetal brain and each cell line were resolved on a 1.2% agarose gel containing formaldehyde and transferred to a nylon membrane (E). Blots were probed with the ID4 cDNA labeled by random priming. To normalize the RNA loading in each lane, the blots were stripped and rehybridized with a ^{32}P-labeled oligonucleotide complementary to the GAPDH message: 5′ ATGACAAGGTGCGGCTCCC (A, B, C, D). Control of the RNA loading in each lane is indicated by the amount of the 18 S RNA stained with ethidium bromide (E).

changes lie at positions 77 and 79. The HLH domain which is the dimerization domain, is highly conserved, in sequence and length, in all classes of HLH proteins (b-HLH, b-HLH-ZIP and Id proteins). At amino acid position 77 the charged amino acid (E, D or R) which is present in all the mammalian Id proteins as well as in the Drosophila Id-related protein emc, is replaced by a nonpolar amino acid: W. Moreover, the valine residue at position 79, at the junction helixI-loop, is missing. This valine along with the leucine at position 78 and the proline at position 80 are absolutely conserved through evolution, in all Id proteins from Drosophila, Xenopus, rodents and humans. Mutational analysis of the HLH domain has shown that the conserved hydrophobic and the charged hydrophilic residus are important for the function of the protein.[9] Consequently, we believe that our clone as well as the one deposited in GenBank under accession number U28368 are the human homologue of the mouse Id4 gene.

Figure 3 shows ID4 expression in human tissue and central nervous system tumor cell lines. As for the mouse, three transcripts of 1.6 kb, 2.6 kb and 4 kb were present in all tissues and cell lines examined. However, an additional 1.1-kb message was expressed in the adult pancreas. The latter transcript could arise from an alternative splicing, as has been described for the Id1 and

Id2 genes.[14] In adult tissues the highest expression was found mainly in thyroid and to a lesser extent in brain and testis (Figs. 3A, 3B, and 3D). In human fetal tissues ID4 is highly expressed in brain, lung and kidney and is silent in the liver (Fig. 3C). This pattern of expression in human fetal tissues is markedly different from that described in fetal mouse tissues. It is noteworth that ID4 is not expressed in mouse fetal lung and kidney,[25] whereas, high levels of ID4 transcripts are present in those two human tissues (Fig. 3C). Our analysis of ID4 expression in nervous system tumor cell lines showed that ID4 expression was found in one astrocytoma cell line SNB19 and in one out of three medulloblastoma cell lines, DAOY (Fig. 3E). Medulloblastomas are thought to arise from poorly differentiated neuroectodermal precursor cells. It has been recently suggested that ID2 and ID4 gene expression was found in tumor cell lines derived from tumors arising from cells of the astrocytic lineage.[26] The DAOY cell line expresses substantial levels of ID4 as well as ID2 transcripts (Fig. 3E and data not shown). This cell line could correspond to a precursor commited to a glial lineage.

To test the hypothesis of active downregulation of ID4 in the D341 and D283 medulloblastoma cell lines, we studied the methylation status of three methylation-sensitive sites in the ID4 CpG-island, since CpG-island methylation is a common inactivating feature of gene expression.[23] However, none of the methylation sensitive sites were found to be methylated, indicating that this mechanism does not contribute to the low level of ID4 expression in these cell lines (data not shown).

We determined the chromosomal assignment of ID4 on chromosome 6 by amplification of genomic DNA of a human/rodent cell hybrid panel (panel #2, Coriell, Camden, NJ; data not shown). Furthermore, using the full-length cDNA as a probe, we precisely mapped this gene, by FISH analysis, to chromosome 6p22.3-p23 (Fig. 4).[27] Interestingly, this part of the genome has been described to be rearranged during progression of adult astrocytomas.[28,29] Previous chromosomal localization of ID4 by Pagliuca et al. to 6p21.3-p22, included two chromosome bands.[17] We have now mapped ID4 more precisely in the distal part of 6p22, at the junction with 6p23.

Comparison of the ID4 sequence with those deposited in GenBank identified an EST mapped in a YAC contig (WI 7424), that showed that ID4 is located on the genetic map in a 1-cM interval between markers D6S1700 (distal border) and D6S422 (proximal border).[30,31]

HLH proteins constitute a complex regulatory network, in which the precise balance of these inhibitory and activating proteins determines different cellular fates. A dysregulation of any of these transcriptional reguators, leading to a relative imbalance of these factors, could participate in tumor development. Further studies will be required to test the contribution of the ID family members in tumors.

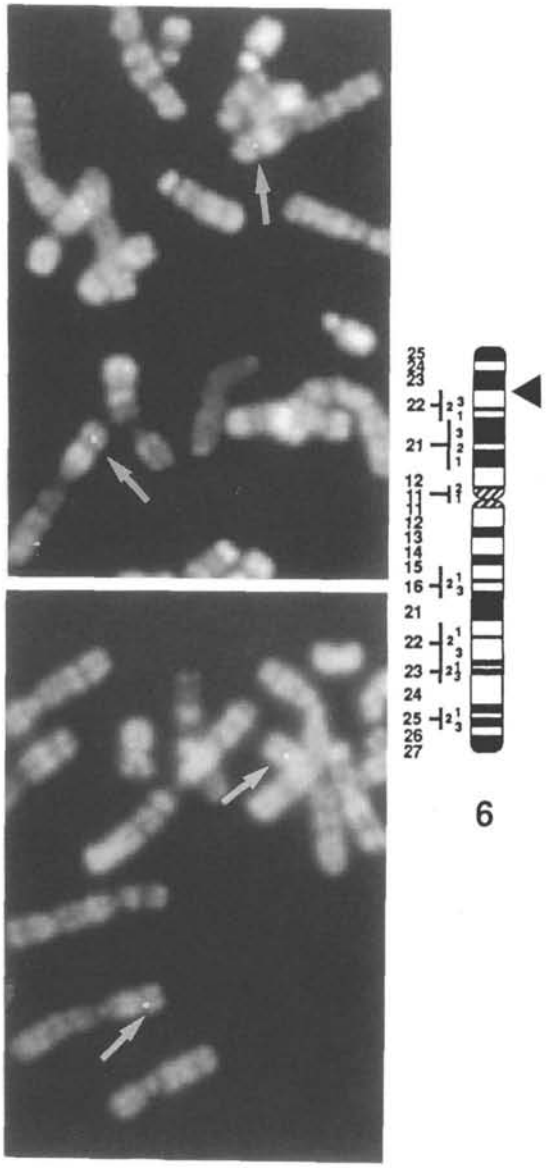

Figure 4. Chromosomal localization of ID4. In situ hybridization of the biotinylated ID4 cDNA probe on human chromosomes. Immunofluorecence detection and propidium iodide counterstaining. Arrows point to fluorescent spots on 6p22.3-p23.

References

1. Barone, M. V., Pepperkok, R., Peverali, F. A., and Philipson, L. 1994, Id proteins control growth induction in mammalian cells, *Proc. Natl. Acad. Sci. USA*, **91**, 4985–4988.
2. Hara, E., Yamagushi, T., Nojima, H., Ide, T., Campisi, J., Okayama, H., and Oda, K. 1994, Id-related genes encoding helix-loop-helix proteins are required for G1 progression and are repressed in senescent human fibroblasts, *J. Biol. Chem.*, **269**, 2139–2145.
3. Iavarone, A., Garg, P., Lasorella, A., Hsu, J., and

Israel, M. A. 1994, The Helix-Loop-Helix protein Id2 enhances cell proliferation and binds to the retinoblastoma protein, *Genes Dev.*, **8**, 1270–1284.
4. Lasorella, A., Iavarone, A., and Israel M. A. 1996, Id2 specifically alters regulation of the cell cycle by tumor suppressor proteins, *Mol. Cell. Biol.*, **16**, 2570–2578.
5. Naya, F. J., Stellrecht, M. M., and Tsai, M-J. 1995, Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor, *Genes Dev.*, **9**, 1009–1019.
6. Zhuang, Y., Soriano, P., and Weintraub, H. 1994, The helix-loop-helix E2A is required for B cell formation, *Cell*, **79**, 875–884.
7. Weintraub, H. 1993, The Myo D family and myogenesis: redundency, networks and thresholds, *Cell*, **75**, 1241–1244.
8. Kageyama, R., Sasai, Y., Akazawa, C. et al. 1995, Regulation of mammalian neural development by helix-loop-helix transcription factors, *Critical Review of Neurobiology*, **9**, 177–188.
9. Murre, C. and Baltimore, D. 1992, In: Yamamoto, K., and McKnight, S. (eds) Transcriptional Regulation, Vol. 2. Cold Spring Harbor, New York, pp. 861–879.
10. Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. 1990, The protein Id: a negative regulator of helix-loop-helix DNA binding proteins, *Cell*, **61**, 49–59.
11. Sun, X. H., Copeland, N. G., Jenkins, N. A., and Baltimore, D. 1991, Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins, *Mol. Cell. Biol.*, **11**, 5603–5611.
12. Christy, B. A., Sanders, L. K., Lau, L. F., Copeland, N. G., Jenkins, N. A., and Nathans, D. 1991, An Id-related helix-loop-helix protein encoded by a growth factor inducible gene, *Proc. Natl. Acad. Sci. USA*, **88**, 1815–1819.
13. Riechmann, V., van Crüchten, I., and Srablizki, F. 1994, The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3, *Nucl. Acids Res.*, **212**, 749–755.
14. Zhu, W., Dahmen, J., Bulfone, A., Rigolet, M., Hernandez, M-C., Kuo, W-L., Puelles, L., Rubenstein, J. L. R., and Israel, M. A. 1995, Id genes expression during development and molecular cloning of the human Id1 gene, *Mol. Brain Res.*, **30**, 312–326.
15. Biggs, J., Murphy, E. V., and Israel, M. A. 1992, A human Id-like helix-loop-helix xpressed during early development, *Proc. Natl. Acad. Sci. USA*, **89**, 1512–1516.
16. Deed, R. W., Bianchi, S. M., Atherton, G. T., Jhonston, D., Santibanez-Koref, M., Murphy, J. J., and Norton, J. D. 1993, An immediate early gene encodes an Id-like helix-loop-helix protein and is regulated by protein kinase C activation in diverse cell types, *Oncogene*, **8**, 599–607.
17. Pagliuca, A., Bartoli, P. C., Saccone, S., Della Valle, G., and Lania, L. 1995, Molecular cloning of ID4, a novel dominant negative helix-loop-helix human gene on chromosome 6p21.3-p22, *Genomics*, **27**, 200–203.
18. Peverali, F. A., Ramqvist, T., Saffrich, R., Pepperkok, R., Barone, M. V., and Philipson, L. 1994, Regulation of G1 progression by E2A and Id helix-loop-helix proteins, *EMBO J.*, **13**, 4291–4301.
19. Shoji, W., Yamamoto, T., and Obinata, M. 1994, The helix-loop-helix protein Id inhibits differentiation of murine erythroleukemia cells, *J. Biol. Chem.*, **269**, 5078–5084.
20. Kreider, B. L., Benezra, R., Rovera, G., and Kadesch, T. 1992, Inhibition of myeloid differentiation by the helix-loop-helix protein Id, *Science*, **255**, 1700–1702.
21. Desprez, P. Y., Hara, E., Bissell, M. J., and Campisi, J. 1995, Suppression of mammary epithelial cell differentiation by the helix-loop-helix protein Id1, *Mol. Cell. Biol.*, **1555**, 3398–3404.
22. Kozak, M. 1984, Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs, *Nucl. Acids Res.*, **12**, 857–872.
23. Baylin, B. S. 1997, Tying it all together: Epigenetics, Genetics, Cell cycle and Cancer, *Science*, **277**, 1948–1949.
24. Nagata, Y., Shoji, W., Obinata, M., and Todokoro, K. 1995, Phosphorylation of helix-loop-helix proteins ID1, ID2 and ID3, *Biochem. Biophys. Res. Commum.*, **207**, 916–926.
25. Reichmann, V. and Stablitzki, F. 1995, Mutually exclusive expression of two dominant negative helix-loop-helix (dnHLH) genes, Id4 and Id3, in the developing brain of the mouse suggests distinct regulatory roles of these dnHLH proteins during cellular proliferation and differentiation of the nervous system, *Cell Growth and Differentiation*, **6**, 837–843.
26. Andres-Barquin, P., Hernandez, M-C., Hayes, T. E., McKay, R. D. G., and Israel, M. A. 1997, Id genes encoding inhibitors of transcription are expressed during *in vitro* astrocyte differentiation and in cell lines derived from astrocytic tumors, *Cancer Res.*, **57**, 215–220.
27. Viegas-Péquignot, E. 1992, In: Wilkinson, D. G. (eds), In situ hybridization, Vol. 109, IRL Press, pp. 137–158.
28. Liang, B. C., Ross, D. A., Greenberg, H. S., Meltzer, P. S., and Trent, J. M. 1994, Evidence of allelelic imbalance of chromosome 6 in human astrocytomas, *Neurology*, **44**, 533–536.
29. Weber, R. G., Sabel, M., Reifenberger et al., 1996, Characterization of genomic alterations associated with glioma progression by comparative genomic hybridization, *Oncogene*, **13**, 983–994.
30. Dib, C., Faure, S., Fizames, C. et al. 1996, A comprehensive genetic map of the human genome based on 5,264 microsatellites, *Nature*, **380**, 152–154.
31. Hudson, T. J., Stein, L. D., Gerety, S. S. et al. 1996, An STS-based map of the human genome, *Science*, **274**, 547–562.