SHORT COMMUNICATION

The Chromosome Location of the Human Homolog of the Mouse Mammary Tumor-Associated Gene *INT6* and Its Status in Human Breast Carcinomas

Shukichi Miyazaki,* Akira Imatani,* Linda Ballard,† Antonio Marchetti,‡ Fiamma Buttitta,‡ Hans Albertsen,† Heli A. Nevanlinna,§ Daniel Gallahan,* and Robert Callahan*,[1]

*Laboratory of Tumor Immunology and Biology, National Cancer Institute, Bethesda, Maryland 20892; †Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah 84112; ‡Molecular Pathology Section, Department of Oncology, University of Pisa, Pisa, Italy; §Helsinki University Central Hospital, Haartmaninkatu 2, 00290 Helsinki, Finland

Received April 22, 1997; accepted August 25, 1997

The *INT6* gene is a common integration site for the mouse mammary tumor virus in mouse mammary tumors. We have determined that the human homolog of *INT6* is located on chromosome region 8q22–q23. A processed *INT6* pseudogene is located on chromosome 6q. *INT6* is composed of 13 exons that span 45 kb of genomic DNA. The deduced amino acid sequence of the gene product is identical to the mouse protein and contains three potential translation start signals. We have examined 100 primary breast carcinoma DNAs for evidence of genetic alteration affecting *INT6*. Loss of heterozyosity (LOH) was detected in 11 of 39 (28%) of the tumor samples informative for a polymorphic sequence in intron 7 of *INT6*. Since single-strand conformation and hybrid mismatch analysis of the remaining allele in these tumor DNAs failed to detect any mutations, we conclude that the target gene for LOH must be closely linked to *INT6*.

INT6 was identified as a common integration site for the mouse mammary tumor virus (MMTV) in a transplantable hyperplastic outgrowth (HOG) line derived from a preneoplastic hyperplastic alveolar nodule and in two independent mammary tumors from unrelated mice (7). In each case the proviral genome was integrated within a different intron of the gene and was in the opposite transcriptional orientation to that of *INT6*. One consequence of these integration events was the expression of a truncated *INT6* RNA species terminating within the long terminal repeat (LTR) element of the integrated MMTV proviral genome. Since no point mutations were detected in the non-rearranged allele, we hypothesized that MMTV integration into INT6 either caused the expression of a biologically activated protein or corresponded to a dominant-negative mutation. Expression of *INT6* can be detected as early as Day 8 of embryonic development and has been detected in all adult tissues that have been tested (5, 7). *INT6* has the potential to encode a 54-kDa protein whose amino acid sequence is unrelated to any known protein in the GenBank. Although the function of INT6 is unknown, it is primarily a cytoplasmic protein that localizes to the Golgi apparatus (5). In this report we present the complete nucleotide sequence of the human homologue of *INT6*, its chromosomal location, and a survey of 100 primary breast tumor DNAs to determine whether it is a target for mutation during tumor development.

We have previously shown by Southern blot analysis that *INT6* has been highly conserved in mammals and in more distantly related species such as *Drosophila* and *Caenorhabitis elegans*. (7). A recombinant clone of the 1.5-kb *INT6* RNA transcript was obtained from a human lung cDNA library (Clontech). Its nucleotide sequence (GenBank Accession No. U94175) is 90% similar to mouse *INT6*, and the deduced amino acid sequence of the gene products are identical. There are three potential translation start signals at 21, 168, and 183 bp, respectively. In a reticulocyte *in vitro* transcription/translation system, the primary start signal used is at 21 bp (5). The major INT6 protein species detected *in vivo* is approximately 43 kDa in size and comigrates with the *in vitro* product initiated at 183 bp (5). During the preparation of this article, Desbois *et al.* (4) reported the nucleotide sequence of a partial cDNA clone of human *INT6* missing 42 bases at the 5′ end including the first translation start signal.

The 1.5-kb *INT6* RNA transcript is encoded by 13 exons that span 45-kb of overlapping recombinant λ and P1 bacterial phage clones of human genomic DNA (Fig. 1A). Southern blot analysis of *Eco*RI-digested human genomic DNA resolves four *INT6*-related fragments above 2 kb in size (Fig. 1B, Lane 3). The 4.6-kb (exons 2 and 3) and 3.7-kb (exons 6 and 7) fragments were also detected in cellular DNA from a Chinese hamster/human somatic cell hybrid whose human com-

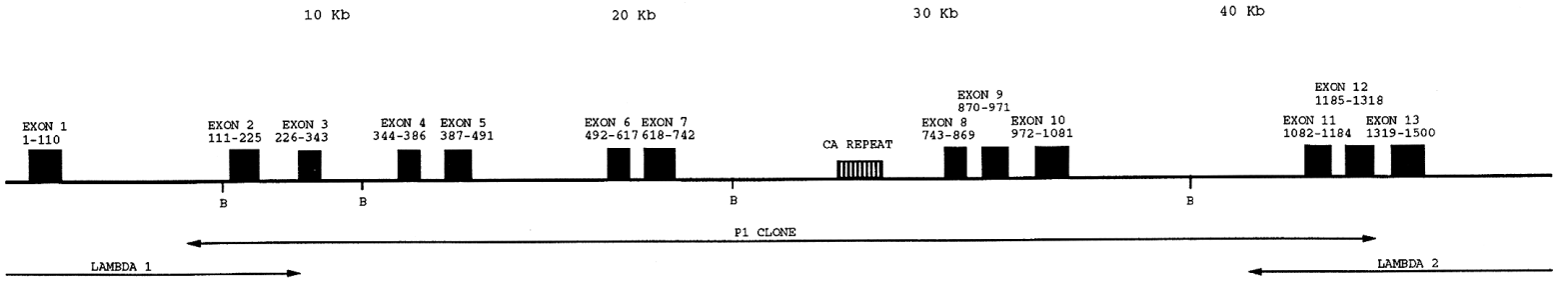
Sequence data from this article have deposited with the EMBL/GenBank Data Libraries under Accession Nos. U94162–U94174.

[1] To whom correspondence should be addressed at the National Cancer Institute, Building 10, Room 5B50, Bethesda, MD 20892. Telephone: (301) 496-9871. Fax: (301) 402-0711.

GENOMICS **46,** 155–158 (1997)
ARTICLE NO. GE974996
0888-7543/97 $25.00

A.

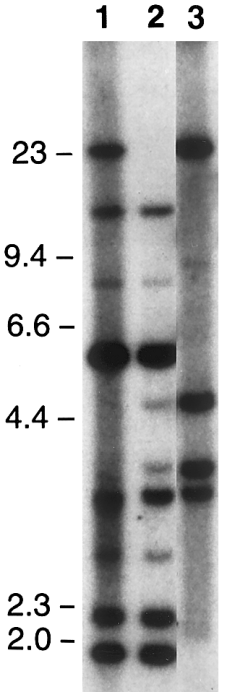
HUMAN INT6 GENOMIC MAP



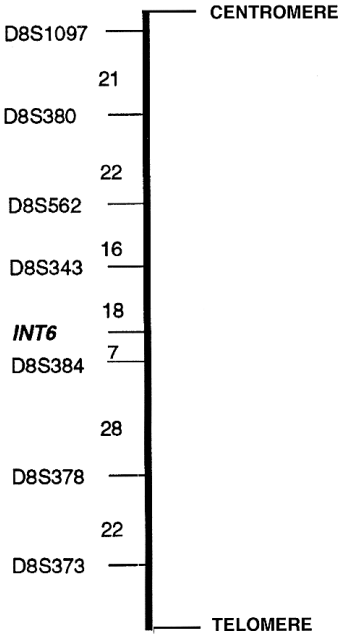
PRIMERS TO DETECT THE CA-REPEAT IN INTRON 7

5' GTGAAAATGACATGAAATTTCAG 40 bp $\binom{GT}{CA}$18 138 bp GCCCATATTGTCACACTGCA 3'
CACTTTTACTGTACTTTAAAGTC CGGGTATAACAGTGTGACGT

B.



C.



plement is only chromosome 8 (Fig. 1B, Lane 2). We have shown previously that mouse *INT6* is located on chromosome 15, 16 cM centromeric of the *Myc* protooncogene (8). This region of mouse chromosome 15 is conserved on human chromosome 8q.

The mouse genome also contains several processed *INT6* pseudogenes (8). The *INT6*-related 23-kb *Eco*RI fragment is detected in a Chinese hamster/human hybrid containing only chromosome 6 as a part of its human complement of chromosomes (Fig. 1B, Lane 1). It has been reported that *INT6* sequences (D6S1876) are located on chromosome 6q13 (NCBI, The Human Gene Map). The nucleotide sequence of the fragment detected on chromosome 6 is identical to 1136–1360 bp of our *INT6* cDNA nucleotide sequence. The primers used to detect the 224-bp fragment by PCR analysis are located in exons 10 and 12, respectively (Fig. 1A). It seems likely that the chromosome 6 *INT6*-related sequences correspond to a processed *INT6* pseudogene, since the genomic distance separating exons 10 and 12 on chromosome 8 exceeds 10 kb.

During the process of defining the exon/intron structure of *INT6*, we identified a polymorphic dinucleotide-repeat sequence in intron 7 (Fig. 1A). In a survey of 125 unrelated individuals, four distinct alleles were observed, and a total of 67 (54%) individuals were found to be heterozygous (data not shown). Based on genetic typing of *INT6* in four CEPH (K884, K1331, K1332, and K1362) (6, 12), we performed a genome-wide linkage scan and observed significant two-point lod scores between INT6 and several genetic markers previously mapped to the long arm of chromosome 8. This initial localization led to the genotyping of an extended set of CEPH pedigrees (K1463, K1447, K1346, K1451, K13293, K1375, and K1334) by multipoint linkage analysis to localize precisely *INT6* relative to seven reference markers on the long arm of chromosome 8. The most likely location of *INT6* (720 times more likely than the next best order) was determined by the REGION MAP algorithm (13) to be within the 25 cM interval between D8S343 and D8S384 on chromosome region 8q22–q23 (Fig. 1C). It may be pertinent that D8S384 has been shown to be in the middle of the minimal region of overlap (MRO) of the Langer–Giedion syndrome (LGS) locus (2).

The potential involvement of *INT6* in mouse preneoplastic mammary epithelial lesions and mammary tumors provided the rationale for determining whether this gene is mutated in primary human breast tumors. We and others (reviewed in (1)) have shown that loss of heterozygosity (LOH) represents the most common type of mutation in primary human breast carcinomas, affecting at least 20 different regions of the cellular genome. We have surveyed a panel of 89 primary human breast carcinomas (primarily invasive ductal carcinomas) and matching lymphocyte DNAs (3) for LOH using the *INT6* dinucleotide-repeat polymorphism. Previously, chromosome region 8q22–q23 has not been identified as a frequent target for LOH. In our study, of the 39 informative patients, 11 (28%) had LOH of *INT6* (Fig. 2). We have used the hybrid-mismatch (9, 14) and single-strand conformation polymorphism (SSCP) (10) methodologies to determine whether the remaining *INT6* allele in the affected tumors contains point mutations. The nucleotide sequences of the primer pairs used in these analyses have been assigned GenBank Accession Nos. U94162–U94174. Examination of the coding region of each *INT6* exon in the affected tumor genomic DNAs failed to reveal any evidence of mutation (data not shown). Chromosome region 8p21.3–p22 has been reported to be frequently affected by LOH in primary human breast carcinomas (1, 11). LOH at *INT6* therefore may reflect those tumors that are monosomic for chromosome 8. We have preliminary evidence from a different panel of primary human breast tumors that 37% have a loss of *INT6* RNA expression (Marchetti *et al.*, unpublished data). An alternative explanation for LOH at *INT6* is that the promoter in the remaining allele has been rendered

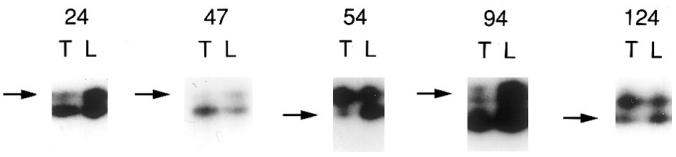


FIG. 2. Loss of heterozygosity at *INT6*. The nucleotide sequences shown in Fig. 1 A were used as primers in a PCR assay (30 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 2 min; after the last cycle, 2 min at 72°C, the products were electrophoresed on 6% nondenaturing polyacrylamde gels) to survey matched pairs of breast tumor/lymphocyte DNAs for LOH. Five examples are shown, the arrow indicates the lost allele.

FIG. 1. The organization of the *INT6* gene and identification of its chromosomal location by somatic cell hybrid and linkage analysis. (**A**) The organization and location of *INT6* exons in two λ phage clones (origin) that overlap with the ends of a human DNA containing P1 phage clone (Genomes Systems Inc., St Louis, MO) were determined by a combination of restriction enzyme and nucleotide sequence analyses. The genomic clones are indicated by arrows below the bar. *Bam*HI restriction sites are indicated by the letter **B.** Exons are depicted as solid boxes, and the corresponding portions of the *INT6* cDNA nucleotide sequence are indicated above the boxes (given in basepairs). The hatched box indicates the location of the *INT6* polymorphic CA-repeat sequence. The nucleotide sequence of the primers used to detect it by PCR analysis are given below the bar. (**B**) The chromosomes containing *INT6*-related sequences were determined by Southern blot analysis of *Eco*RI-digested DNA from Chinese hamster/human hybrids (NIGMS Human Genetic Mutant Cell Repository, Coriell Cell Institute for Medical Research,Camden, NJ) containing one human chromosome. Lanes **1** and **2** contain DNA from hybrids GM/NA10629 and GM/NA10156B, containing chromosomes 6 and 8, respectively. The 23-, 4.6-, and 3.7-kb *Eco*RI fragments contain human *INT6*-related sequences. Lane **3** contains *Eco*RI-digested human placental DNA. (**C**) The location of *INT6* CA-repeat polymorphism on chromosome 8 was determined by linkage analysis of DNA from 11 (4 + 7) CEPH families using 7 reference markers (see text). The numbers between the horizontal bars indicate the distance in centimorgans between markers.

nonfunctional by mutation or is inactive as consequence of genetic imprinting. We are currently testing these possibilities.

REFERENCES

1. Bieche, I., and Lidereau, R. (1995). Genetic alterations in breast cancer. *Genes Chromosomes Cancer* **14:** 227–251.
2. Chen, W., Hou, J., Wagner, M. J., and Wells, D. E. (1996). An integrated physical map covering 25 cM of human chromosome 8. *Genomics* **32:** 117–120.
3. Cropp, C. S., Nevanlinna, H. A., Pyrhonen, S., Stenman, U.-h., Salmikangas, P., Albertsen, H., White, R., and Callahan, R. (1994). Evidence for involvement of BRCA1 in sporadic breast carcinomas. *Cancer Res.* **54:** 2548–2551.
4. Desbois, C., Rousset, R., Bantignies, F., and Jalinot, P. (1996). Exclusion of Int-6 from PML nuclear bodies by binding to the HTLV-1 Tax oncoprotein. *Science* **273:** 951–953.
5. Diella, F., Levi, G., and Callahan, R. (1997). Characterization of the INT6 mammary tumor gene product. *DNA Cell Biol.* **16:** 839–847.
6. The Utah Marker Development Group (1995). A collection of ordered tetranucleotide-repeat markers from the human genome. *Am. J. Hum. Genet.* **57:** 619–628.
7. Marchetti, A., Butitta, F., Miyazaki, S., Gallahan, D., Smith, G. H., and Callahan, R. (1995). Int-6, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. *J. Virol.* **69:** 1932–1938.
8. Miyazaki, S., Kozak, C. A., Marchetti, A., Butitta, F., Gallahan, D., and Callahan, R. (1995). The chromosomal location of the mouse mammary tumor gene Int6 and related pseudogenes in the mouse genome. *Genomics* **27:** 420–424.
9. Myers, R. M., Larin, Z., and Maniatis, T. (1985). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science* **230:** 1242–1246.
10. Osbourne, R. J., Merlo, G. R., Mitudomi, T., Venesio, T., Liscia, D. S., Cappa, A. P. M., Chiba, I., Takahashi, T., Nau, M. M., Callahan, R., and Minna, J. D. (1991). Mutations in the p53 gene in primary human breast cancer. *Cancer Res.* **51:** 6194–6198.
11. Patel, U., Grundfest-Broniatowski, S., Gupta, M., and Banerjee, S. (1994). Microsatellite instabilities at five chromosomes in primary breast tumors. *Oncogene* **9:** 3695–3700.
12. Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G., and Laterop, M. (1992). A second-generation linkage map of the human genome. *Nature* **359:** 794–801.
13. White, R. L., Lalouel, J.-M., Nakamura, Y., Donis-Keller, H., Green, P., Bowden, D. W., and Mathew, C. G. P. (1990). The CEPH consortium primary linkage map of human chromosome 10. *Genomics* **6:** 393–412.
14. Winter, E., Yamamoto, F., Almohuera, C., and Perucho, M. (1985). A method to detect and characterize point mutations in transcribed genes. *Proc. Natl. Acad. Sci. USA* **82:** 7575–7579.