A Putative *int* Domain for Mouse Mammary Tumor Virus on Mouse Chromosome 7 Is a 5' Extension of *int-2*

G. PETERS,^{1*} S. BROOKES,¹ M. PLACZEK,^{1†} M. SCHUERMANN,^{2‡} R. MICHALIDES,² AND C. DICKSON³

Imperial Cancer Research Fund Laboratories, St. Bartholomews Hospital, Dominion House, Bartholomew Close, London, EC1A 7BE,¹ and Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London, WC2A 3PX,³ United Kingdom, and Netherlands Cancer Institute, Antoni van Leeuwenhoekhuis, 1066 CX Amsterdam, The Netherlands²

Received 1 September 1988/Accepted 8 November 1988

We extended the physical map of the mouse *int-2* locus by demonstrating that the site of insertion for mouse mammary tumor virus DNA in plaque-type mammary tumors of GR mice is directly linked to *int-2*. An additional example of proviral integration is described in which a provirus in a presumed enhancer-insertion mode 15 kilobases upstream of the *int-2* promoters is capable of activating expression of the gene at levels typical of other virally induced mammary tumors.

One of the key steps in the induction of mammary tumors by mouse mammary tumor virus (MMTV) is the transcriptional activation of specific cellular genes by the intrusion of a provirus in the neighboring chromosomal DNA (6a, 9). Thus, most virally induced tumors are clonal proliferations of cells that have sustained a proviral insertion in one (or more) of a series of common integration domains, termed *int* loci (1, 2, 8, 11). To date, there are four bona fide *int* loci located on different mouse chromosomes and associated with the expression of tumor-specific RNAs from the respective genes (1-3, 6a, 7, 8, 10; R. Nusse, personal communication). In addition, there are indications of the existence of further *int* loci for which the presence of a cellular gene has not been rigorously established (4, 5, 12). One of these was described in two separate cases of plaque-type (P-type) mammary tumors characteristic of the GR strain of mice (12). For the purposes of this discussion, it will be convenient to refer to this chromosomal domain as *int-P*.

As part of the original characterization of *int-P*, two overlapping lambda clones, designated $\lambda 40/1$ and $\lambda 40/2$, were isolated and physically mapped (12). Unique sequence probes from these recombinants were used to extend the physical map of the region by Southern blotting of cellular DNA and to establish the chromosomal position of the locus. The derived physical map spanned over 35 kilobases (kb) of DNA and was clearly different from the published maps of *int-1*, *int-2*, and *int-3* (2, 7, 11, 12). However, only two examples of proviral insertion in this domain (tumors MT40 and MT42; see Fig. 1) had been detected among the panel of mammary tumors surveyed, and there were no indications that the region was expressed into RNA in the two positive cases. Here we report that *int-P* is a direct extension of the *int-2* locus, that proviral insertion in the region has occurred in at least one additional tumor from the BR6 strain of mouse, and that proviral insertion at *int-P* can activate the expression of the *int-2* gene.

The initial indication that *int-P* and *int-2* may be linked came from their coincidence on mouse chromosome 7 (10,

12). Although neither tumor MT40 nor MT42 appeared to express detectable *int-2* transcripts (12), such assays could have been compromised by the condition of the tumor RNA or the sensitivity of the probes, particularly as the gene is normally expressed at only low levels (1). The structures of the various *int-2* transcripts are now better understood and more sensitive probes and techniques are available (6, 13), but there is unfortunately no residual RNA from MT40 or MT42 with which to reassess gene expression.

The second indication stemmed from a direct comparison of the restriction maps of *int-P* and *int-2*. Significantly, the extreme 5' end of *int-P*, as originally drawn, bore similarities to the most upstream regions of the *int-2* domain, but in an inverted orientation. In particular, both shared a 7.2-kb *EcoRI* fragment in which there were similarly located *BamHI* and *BglII* sites (see Fig. 1). There were also some discrepancies, but given the fact that the extremities of the physical maps were based on Southern blotting data as opposed to detailed characterization of cloned DNAs, it was decided to examine the issue more rigorously.

The 7.2-kb *EcoRI* fragment was isolated as a recombinant DNA clone (designated $\lambda 779$) during the analysis of mammary tumor U5C from a BR6 mouse (Fig. 1) (M. Placzek and G. Peters, unpublished data). Several examples of proviral insertion within this *EcoRI* fragment were reported in an earlier publication (11). In a preliminary experiment, it was found that the 7.2-kb *EcoRI* fragment contained in $\lambda 779$ hybridized to DNA from $\lambda 40/1$ (data not shown). Moreover, the 1.0-kb *SacI* subfragment (designated S-S in Fig. 1) from $\lambda 779$ detected appropriate restriction fragments in digests of $\lambda 40/1$. This allowed us to construct a composite map of the *int-P* and *int-2* domains based on published information, essentially as depicted in Fig. 1. However, to exclude the unlikely possibility that this cross-reactivity was spurious and to resolve some apparent discrepancies in the published maps, we used several unique sequence DNA fragments from $\lambda 40/2$, $\lambda 40/1$, and $\lambda 779$ as probes on Southern blots of total mouse DNA. The examples shown in Fig. 2 again indicate that $\lambda 40/1$ and $\lambda 779$ overlap by approximately 4 kb and confirm that this section of the genomic map is as shown in Fig. 1. The only part of the composite map that was not established by the overlapping lambda clones shown here, but rather by Southern blotting of total mouse DNA, was the

* Corresponding author.

† Present address: College of Physicians and Surgeons, Columbia University, New York, NY 10032.

‡ Present address: Institut für Molekular Biologie und Tumorforschung, D-3500 Marburg, Federal Republic of Germany.

FIG. 1. Extended physical map of the mouse *int-2* locus. The figure represents a linear map (bold line) encompassing both *int-2* and *int-P* and indicating the sites of cleavage for the restriction endonucleases *Bam*HI (B), *Bgl*II (Bg), *Eco*RI (E), *Hind*III (H), *Kpn*I (K), *Sac*I (S), and *Xba*I (X). The coding exons of the *int-2* gene are shown as filled boxes. The arrowheads above the map depict the positions and orientations of MMTV proviruses that have been located within the *int-2* domain in different mouse mammary tumors. Those discussed in the text are identified by filled symbols together with their respective designations. The lines below the map show the extents and designations of the various lambda DNA clones from which the structure of the locus was deduced, with the lengths in kilobases reflected in the scale. Two unique sequence probes (E-H and S-S) used to establish linkage between *int-P* and *int-2* are represented as shaded boxes.

junction between the 7.2-kb *Eco*RI fragment (λ 779) and the adjacent 10-kb *Eco*RI fragment (contained in λ 418).

Having established a direct link between *int-P* and *int-2*, it was of interest to determine whether and at what level a provirus at *int-P* might activate the transcription of the *int-2* gene. In a previous survey of mammary tumors from BR6

mice, there were examples in which *int-2* RNAs were clearly expressed but in which the activating provirus could not be located (11). In one of these, designated A38, we detected an MMTV provirus very close to the integrations described in tumors MT40 and MT42 (Fig. 1). Novel virus-host junction fragments were observed (Fig. 3a) from which it was possible to deduce that the MMTV provirus is complete and in the opposite transcriptional orientation to that of the *int-2* gene

FIG. 2. *int-P* and *int-2* are physically linked in chromosomal DNA. Samples (10 μ g) of BALB/c mouse DNA were digested with the restriction enzymes *Bam*HI (B), *Bgl*II (Bg), *Eco*RI (E), *Sac*I (S), or *Xba*I (X) or combinations thereof and fractionated by electrophoresis through a 0.8% agarose gel. After transfer to nitrocellulose, the DNAs were hybridized sequentially with the 1.9-kb *Eco*RI-*Hind*III fragment from λ 40/1 (a) and the 1.0-kb *Sac*I-*Sac*I fragment from λ 779 (b) (Fig. 1). Conditions for electrophoresis, transfer, nick translation of probes, hybridization, and autoradiography were as described previously (1, 10, 11). Specific fragments hybridizing with both probes are indicated ($\blacktriangle$). The positions and sizes (kilobases) of the fragments of *Hind*III-digested lambda DNA used as standards are shown on the right of the figure.

FIG. 3. Activation of *int-2* by a provirus at *int-P*. (a) Southern blot analysis of tumor A38 DNA (Fig. 1) digested with the enzymes *Sac*I (S), *Kpn*I (K), *Hind*III (H), or the indicated combinations and hybridized to the *Eco*RI-*Hind*III probe from λ 40/1 (Fig. 1). Numbers on the right of the panel indicate the positions and sizes of *Hind*III-digested lambda DNA markers. Novel fragments generated by the insertion of MMTV proviral DNA at the position shown in Fig. 1 are indicated ($\triangleleft$). (b) Polyadenylated RNA (5 μ g) from tumors A38 and T108 was fractionated by electrophoresis through a formaldehyde-agarose gel as previously described (1, 11). After transfer to nitrocellulose, the RNA was hybridized to a 32 P-labeled DNA probe (*int-2f*) from the third exon of *int-2* (11). The sizes of the various classes of *int-2* RNA are indicated in kilobases.

(Fig. 1). Despite the distance between this provirus and the inferred promoter domains of *int-2*, tumor A38 expressed *int-2* RNAs at levels comparable to those seen in other *int-2*-positive tumors (11) (Fig. 3b). The pattern of transcription was also typical of tumors in which the gene is presumably activated by enhancers in the proviral DNA, as exemplified by tumor T108 (Fig. 1 and 3b). In this regard, it is interesting that there may be three quite dissimilar promoter regions from which mouse *int-2* expression can be controlled clustered within the 1 kb preceding the presumed initiation codon and that all three appear equally susceptible to *cis* activation by MMTV, even at distances of 15 kb (6, 13; C. Dickson et al., manuscript in preparation). However, the smaller *int-2* transcripts that terminate at the first poly(A) addition signal (6, 13) are generally under represented in tumors in which the gene is activated by proviral insertion in the upstream DNA (unpublished observations).

Now that DNA probes are available for over 60 kb surrounding the *int-2* gene, it would be interesting to survey further panels of tumors to establish the maximum limits of the *int-2* domain. In view of the distances over which MMTV can apparently act to enhance gene expression, it will be important to identify the cellular genes that lie adjacent to *int-2* in the chromosome and to determine whether the *cis*-acting effects of an MMTV provirus are limited to the most proximal gene. Such information may prove significant for several of the common integration regions for murine leukemia viruses in which, as for *int-P*, clusters of proviruses have been detected but have not yet been associated with expression of a neighboring cellular gene.

We thank Roger Watson and Gerard Evan for helpful comments on the manuscript.

LITERATURE CITED

- Dickson, C., R. Smith, S. Brookes, and G. Peters. 1984. Tumorigenesis by mouse mammary tumor virus: proviral activation of a cellular gene in the common integration region *int-2*. *Cell* 37:529-536.
- Gallahan, D., and R. Callahan. 1987. Mammary tumorigenesis in feral mice: identification of a new *int* locus in mouse mammary tumor virus (Czech II)-induced mammary tumors. *J. Virol.* 61:66-74.
- Gallahan, D., C. Kozak, and R. Callahan. 1987. A new common integration region (*int-3*) for mouse mammary tumor virus on mouse chromosome 17. *J. Virol.* 61:218-220.
- Garcia, M., R. Wellinger, A. Vessaz, and H. Diggelmann. 1986. A new site of integration for mouse mammary tumor virus proviral DNA common to BALB/cf(C3H) mammary and kidney adenocarcinomas. *EMBO J.* 5:127-134.
- Gray, D. A., C. M. McGrath, R. F. Jones, and V. L. Morris. 1986. A common mouse mammary tumor virus integration site in chemically induced precancerous mammary hyperplasias. *Virology* 148:360-368.
- Mansour, S. L., and G. R. Martin. 1988. Four classes of mRNA are expressed from the mouse *int-2* gene, a member of the FGF gene family. *EMBO J.* 7:2035-2041.
- Nusse, R. 1988. The activation of cellular oncogenes by proviral insertion in murine mammary cancer, p. 283-306. In M. E. Lippman and R. Dickson (ed.), *Breast cancer: cellular and molecular biology*. Martinus Nijhoff, Boston.
- Nusse, R., A. van Ooyen, D. Cox, Y. K. T. Fung, and H. E. Varmus. 1984. Mode of proviral activation of a putative mammary oncogene (*int-1*) on mouse chromosome 15. *Nature (London)* 307:131-136.
- Nusse, R., and H. E. Varmus. 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* 31:99-100.
- Peters, G., and C. Dickson. 1988. On the mechanism of carcinogenesis by mouse mammary tumor virus, p. 307-319. In D. Medina, W. Kidwell, G. Heppner, and E. Anderson (ed.), *Cellular and molecular biology of mammary cancer*. Plenum Publishing Corp., New York.
- Peters, G., C. Kozak, and C. Dickson. 1984. Mouse mammary tumor virus integration regions *int-1* and *int-2* map on different mouse chromosomes. *Mol. Cell. Biol.* 4:375-378.
- Peters, G., A. E. Lee, and C. Dickson. 1986. Concerted activation of two potential proto-oncogenes in carcinomas induced by mouse mammary tumour virus. *Nature (London)* 320:628-631.
- Schuermann, M., and R. Michalides. 1987. A rare common integration site of proviruses of the mouse mammary tumor virus in P-type mammary tumors of mouse strain GR. *Virology* 156:229-237.
- Smith, R., G. Peters, and C. Dickson. 1988. Multiple RNAs expressed from the *int-2* gene in mouse embryonal carcinoma cell lines encode a protein with homology to fibroblast growth factors. *EMBO J.* 7:1013-1022.