MOLECULAR AND CELLULAR BIOLOGY, Aug. 1988, p. 3251–3259
0270-7306/88/083251-09$02.00/0

Vol. 8, No. 8

Functional Cloning of Mouse Chromosomal Loci Specifically Active in Embryonal Carcinoma Stem Cells

KRISHNAMOORTHI BHAT,[1] MICHAEL W. McBURNEY,[2] AND HIROSHI HAMADA[1]†*

Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland A1B 3V6,[1] and Department of Medicine, University of Ottawa, Ottawa, Ontario K1H 8M5,[2] Canada

Received 12 February 1988/Accepted 17 May 1988

Chromosomal loci that are specifically active in embryonal carcinoma stem cells were cloned from the mouse genome by functional selection. P19 cells, a pluripotent embryonal carcinoma cell line, were transfected with an enhancer trap (a plasmid containing an enhancerless inactive *neo* gene), and NEO$^+$ transformants were isolated. All of the NEO$^+$ cell lines retained pluripotency and expressed the *neo* gene. When the cells were induced to differentiate, most of the cell lines continued to express the *neo* gene, while the *neo* gene in some cell lines became repressed. From the latter group of cell lines, we have cloned the integrated *neo* gene plus the flanking cellular DNA sequences. Three of the six cloned DNAs possessed a high NEO$^+$-transforming activity in undifferentiated P19 cells. Among these three, two (015 and 052) were inactive in differentiated P19 cells and NIH 3T3 cells, while the remaining one was active in these differentiated cells. Deletion analysis suggested that both 015 and 052 contain two regulatory elements (promoter and enhancer) of cellular DNA origin. The putative enhancer and promoter are separated by at least 6 kilobases in 015 and 1 kilobase in 052. Therefore, 015 and 052 cloned fragments contain regulatory DNA elements that are specifically active in the embryonal carcinoma stem cells.

In the early stage of mouse embryo development, the first overt differentiation of the totipotent embryonic stem cells takes place on day 3.5 postcoitum, which results in the formation of trophectoderm. At 1 day later, the pluripotent cells of the inner cell mass from primitive endoderm; the cells further differentiate to visceral and parietal endoderm (5). Understanding of the first stage of embryogenesis at a molecular level and of the mechanism that regulates gene expression during the early development requires (i) the isolation of genes that are differentially expressed during the formation of the early tissues and (ii) identification of *cis*-acting DNA elements and *trans*-acting factors that control their expression. A number of differentially expressed genes have been isolated and studied, for example, laminin and type IV collagen for parietal endoderm and alpha-fetoprotein for visceral endoderm. However, those genes are expressed in a relatively late stage of embryo development (2, 3). In contrast, no genes or regulatory elements have been isolated that are specifically active in stem cells of an earlier stage. Isolation of such genes would provide an ideal tool for studying the mechanism of gene regulation during the very early stage of embryogenesis. It is also possible that some of the stem-cell-specific genes may play an important role in the embryogenesis; for example, a stem-cell-specific gene(s) that is necessary to maintain the undifferentiated phenotype of embryonic stem cells has been postulated (15, 22).

We wish to isolate embryonic stem-cell-specific genes and to study their regulation and possible role in early embryogenesis. As the first step of such studies, we attempted to isolate, by a functional selection, regulatory DNA elements (such as promoters and enhancers) that are specifically active in the stem cells. Genes associated with such regulatory elements (presumably embryonic stem-cell-specific genes) can be subsequently identified. The selection procedure is based on our previous observation that an inactive enhancerless gene can be activated if it is integrated near an endogenous enhancer (7). We previously applied this method to HeLa cells and successfully isolated two enhancer elements from the HeLa genome (8, 20).

In the present study, the same selection method was applied to a murine embryonal carcinoma (EC) cell line. EC cells are undifferentiated stem cells that can differentiate in vitro and in vivo into a variety of somatic cell types and can form somatic tissues of chimeric mice when injected into blastocysts. Therefore, EC cells are thought to be equivalent to the stem cells of the inner cell mass. By applying the functional selection procedure to an EC cell line, we cloned mouse chromosomal loci that were specifically active in these stem cells. In two of the clones, we identified regulatory DNA elements specific to the EC stem cells.

MATERIALS AND METHODS

Cells and plasmids. Undifferentiated P19 cells were maintained in alpha-MEM medium containing 10% bovine calf serum as described previously (18). pA10neo was constructed by ligating the *Hin*dIII-*Bam*HI 2-kilobase-pair (kb) fragment from pSV2neo (19) to the *Hin*dIII-*Bam*HI 4-kb fragment from pA10cat (13). λEMBL12 was kindly provided by E. Natt and G. Scherer (17).

Isolation of NEO$^+$ transformants. P19 cells (5×10^5) plated in a dish were transfected with pA10neo by the calcium phosphate method. At 36 h after the transfection, nonselective medium was replaced by the selective medium (200 μg of G418 per ml). P19 cells seem to be very sensitive to G418; therefore, the transfected cells were not replated. At 10 to 12 days later, 80 NEO$^+$ colonies were isolated from 25 dishes and maintained in the selective medium; 40 of them were established as stable cell lines.

Induction of differentiation. P19 cells lines were induced to differentiate as described previously (4, 11, 18). In brief, undifferentiated cells were trypsinized and seeded in a bacterial dish and incubated for 4 days in the presence of 300

* Corresponding author.
† Present address: Department of Biochemistry, Faculty of Medicine, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan.

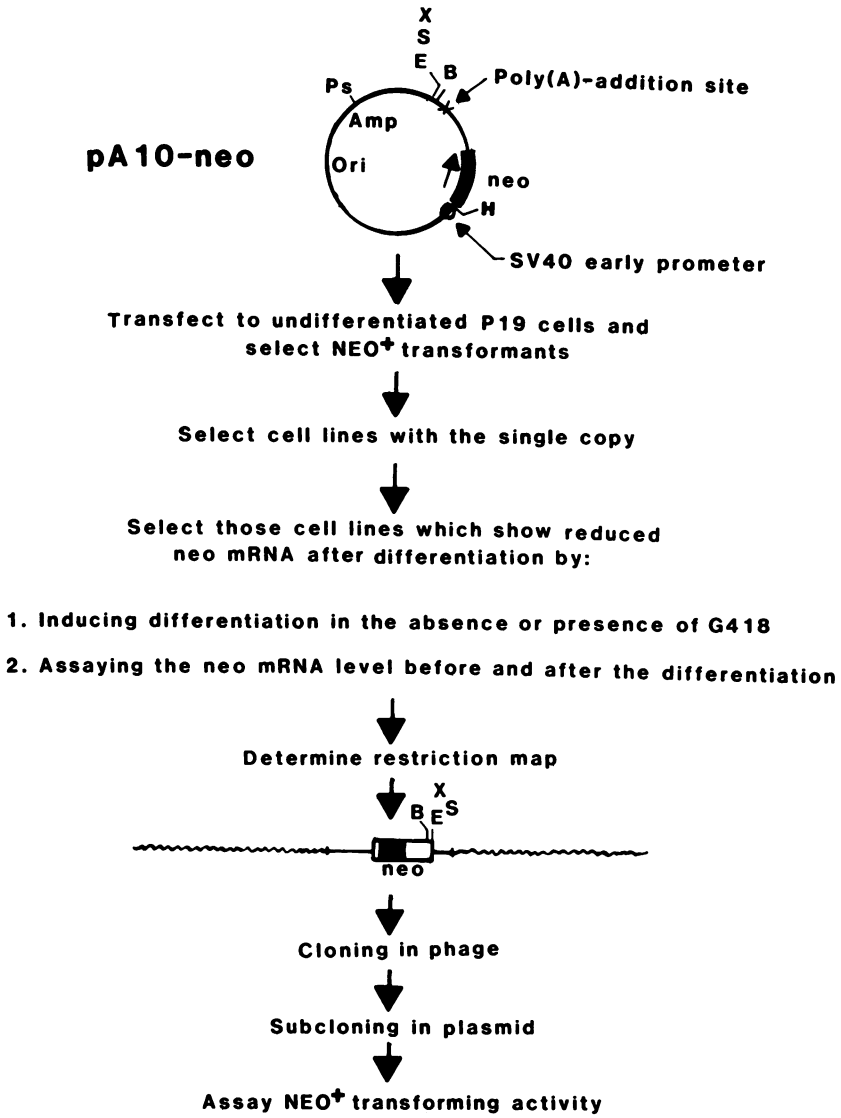

FIG. 1. Strategy for functional cloning of the EC stem-cell-specific chromosomal loci. The functional map and restriction sites of the enhancer trap are shown at the top. B, *Bam*HI; H, *Hin*dIII; E, *Eco*RI; Ps, *Pst*I; S, *Sst*I, X, *Xba*I. The SV40 early promoter (including a TATA box and GC boxes) is contained within a 200-base-pair region upstream of the *Hin*dIII site (see text for details).

nM retinoic acid or 1% dimethyl sulfoxide (DMSO). The cells formed aggregates during this time. The aggregates were subsequently plated on a tissue culture dish and incubated for an additional 4 days in the absence of the chemical inducer. Differentiated cell types such as neurons and muscle cells become obvious at day 7 or 8 after the induction (4, 11). Phase-contrast pictures were taken at day 8. When cells were induced to differentiate in the presence of G418, 200 μg of G418 per ml was present throughout.

Cloning of the flanking mouse DNA. Genomic DNA from each cell line was digested with an appropriate restriction enzyme and fractionated on a gel. DNA fragments of an appropriate size were partially purified and ligated to EMBL3A or EMBL12 arms. The ligated DNA was subjected to in vitro packaging by using Giga-pack Gold (Stratagene). The resulting bacteriophages were screened with a *neo*-specific probe (the *Hin*dIII-*Bam*HI 2-kb fragment of pA10neo).

Stable transformation assay. Cells (5×10^5) plated on a dish were transfected with 2.5 pmol of plasmid (e.g., 10 μg for pA10neo). The cells were incubated in nonselective medium for 36 h; the medium was then replaced by the selective medium (200 μg of G418 per ml for P19 cells and 400 μg of G418 per ml for NIH 3T3 cells). In the case of P19 cells, the transfected cells were not replated. The transfected NIH 3T3 cells were trypsinized 36 h after the transfection, and 10^5 cells were replated. Cells were incubated for an additional 12 days and stained with Giemsa stain. Colonies consisting of more than 100 cells were counted.

RESULTS

Strategy for selection of embryonic-stem-cell-specific loci. We have previously designed a procedure to select enhancer elements from mammalian genomes (8). The procedure is based on our previous observation that an enhancerless gene, which is transcriptionally inactive by itself, can be activated upon transfection if the gene is stably integrated near an endogenous enhancer (7). The chromosomal loci containing the endogenous enhancer can be isolated by

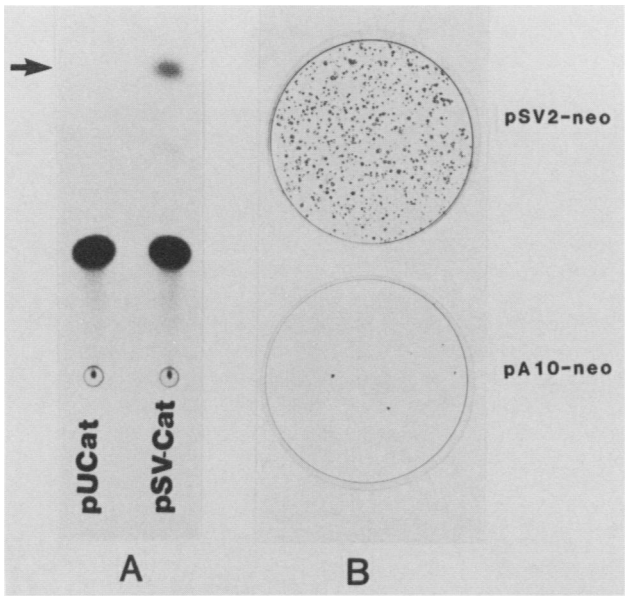

FIG. 2. Enhancer dependence of SV40 early promoter in P19 cells. (A) Transient expression assay. pUcat (10 μg) (a cat gene is linked to the enhancerless SV40 early promoter) or pSV2-cat (an enhancer plus counterpart) was transfected to P19 cells. At 48 h after the transfection, the cells were harvested, and the expression of the cat gene was monitored by enzymatic conversion of chloramphenicol to its acetylated forms as described previously (7). The acetylated chloramphenicol is indicated by the arrow. (B) Stable transformation assay. pA10neo (10 μg) (enhancerless) or pSV2neo (enhancer plus) was transfected to P19 cells, and NEO^+ colonies were selected as described in Materials and Methods.

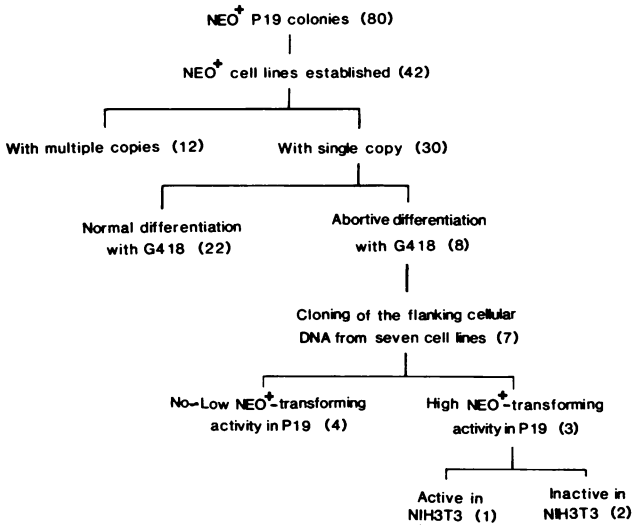

FIG. 3. Summary of isolation and selection of cell lines. The number of cell lines obtained is shown in parentheses (see text for details).

molecular cloning. The procedure was previously applied to HeLa cells, and two distinct enhancers were successfully isolated from the HeLa genome (8, 20).

In this study, the same procedure was applied to P19 cells, a murine EC cell line (16). P19 cells are pluripotent; they can be induced to differentiate into a variety of cell types, such as neurons, astroglia, fibroblasts, and muscle cells (4, 11). When the cells are injected into a blastocyst, they form somatic tissues of the resulting chimeric mouse. Therefore, P19 cells seem to represent embryonic stem cells of early mouse development (probably inner cell mass cells).

The strategy we used for selecting chromosomal loci specifically active in the EC stem cells is summarized in Fig. 1. First, P19 cells were transfected with pA10neo, an enhancer trap containing a *neo* gene linked to the enhancerless simian virus 40 (SV40) early promoter. We selected NEO^+ cells and established a number of transformants. Each NEO^+ cell line was subjected to the following assays: (i) copy number, (ii) ability to differentiate, and (iii) kinetics of the *neo* gene expression during differentiation. We then selected pluripotent cell lines that contained a single copy of the *neo* gene and that showed a greatly reduced expression of the *neo* gene after differentiation. In such cell lines, the *neo* gene is presumably activated by a stem-cell-specific element. Chromosomal sequences flanking the integrated *neo* gene were cloned from those cell lines and searched for stem-cell-specific regulatory elements.

To apply the selection procedure to a given cell line, it is essential that expression of a marker gene be enhancer dependent in the cells. We were concerned whether P19 cells satisfy this requirement, particularly because the SV40 early promoter has been shown to be enhancer independent in F9 cells, another EC cell line (6). However, our results indicate that the SV40 early promoter is enhancer dependent in P19 cells (Fig. 2). When the activity of the enhancerless and SV40-enhancer-linked promoters were examined by transfection, both transient expression and stable transformation assays showed that the enhancerless promoter was much less active than the enhancer-linked promoter. Therefore, P19 cells appear to satisfy this critical requirement.

Characterization of NEO^+ cell lines. Isolation and selection of cell lines are summarized in Fig. 3. When P19 cells were transfected with pA10neo, NEO^+ colonies appeared with a low frequency (Fig. 2B). A total of 80 NEO^+ colonies were recovered, and 40 of them were established as stable cell lines. By dot hybridization analysis using a *neo*-specific probe, 30 cell lines each containing a single copy of the *neo* gene were selected. When each of the 30 cell lines were induced to differentiate by retinoic acid, all were able to differentiate into neurons and glial cells like their parental cell line (data not shown), indicating that they retain pluripotency.

Next, we wished to select cell lines in which the *neo* gene is activated by a stem-cell-specific regulatory element. This was achieved by two assays. First, each cell line was induced to differentiate in the presence of G418. If the *neo* gene is under the control of a stem-cell-specific element, such a cell line would be unable to complete normal differentiation since the neo gene would be repressed during the differentiation. To test the efficiency of this assay, the parental P19 cells and 11 NEO^+ cell lines transformed with pSV2neo were induced to differentiate in the presence of G418 (the SV40 enhancer is probably active throughout the differentiation). While all of the pSV2neo-transformed cell lines were able to complete normal differentiation, P19 cells died before forming the differentiated colonies.

When each of the 30 NEO^+ cell lines transformed with pA10neo was induced to differentiate in the presence of G418, the 30 cell lines fell into two groups (Fig. 4). The first group (group I, 22 of 30 cell lines) differentiated normally regardless of the presence or absence of G418; the retinoic-acid-induced aggregates formed large colonies containing numerous neurons and glial cells. On the other hand, the second group (group II, 8 of 30 cell lines) showed an

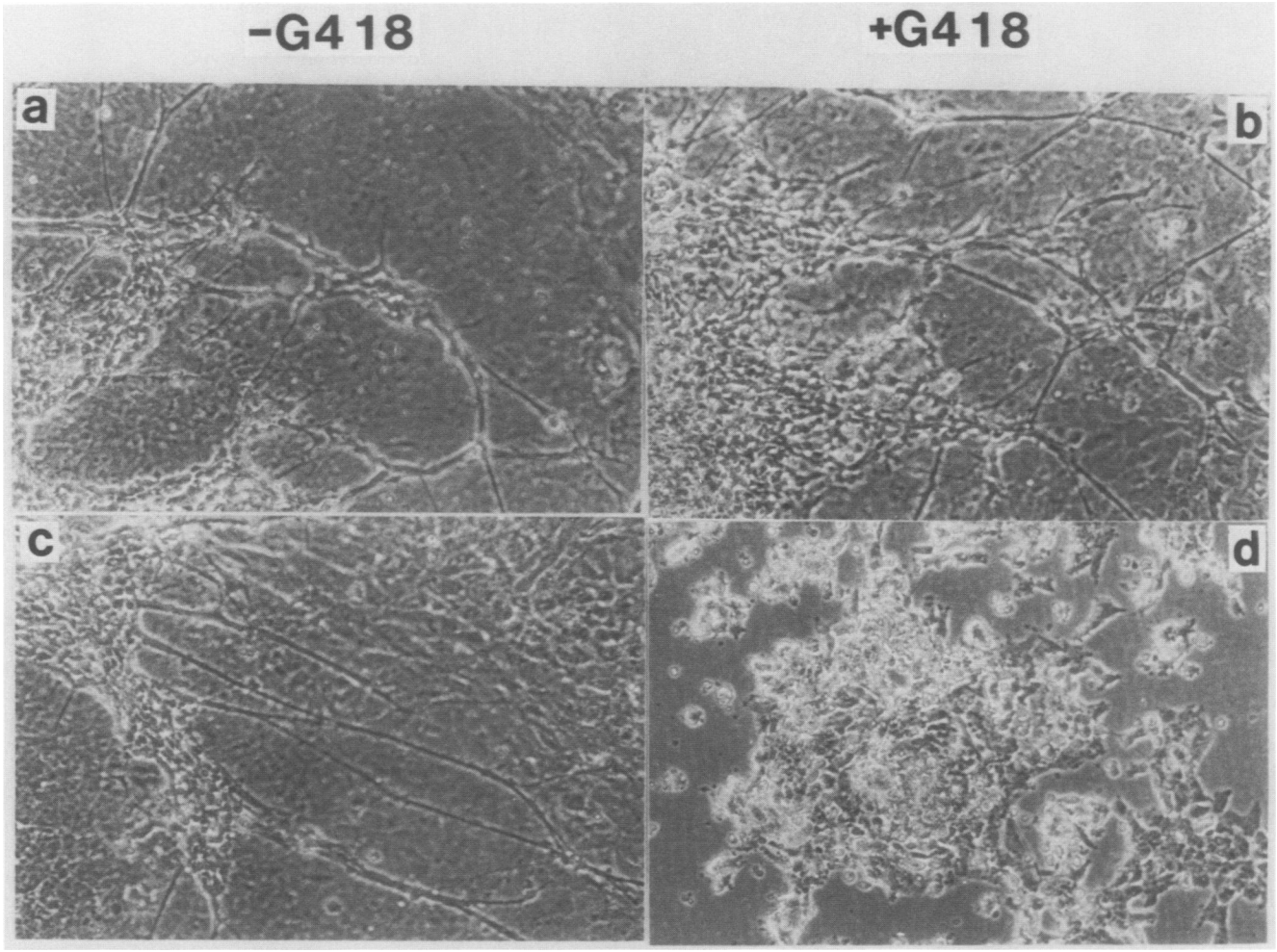

FIG. 4. Retinoic-acid-induced differentiation in the absence (−G418) or presence (+G418) of G418. Each NEO$^+$ cell line was induced by retinoic acid to differentiate in the absence (a and c) or presence (b and d) of G418. A total of 30 cell lines were classified into two groups (group I, a and b; group II, c and d) according to their response to G418.

abnormal differentiation pattern in the presence of G418. The aggregates did not grow well after being plated on a tissue culture dish, the replicated glial cells were much less frequent than normal, and neurons were rarely detected (Fig. 4).

To confirm the above results, we next examined the expression of the *neo* gene more directly, before and after differentiation. Poly(A)$^+$ RNA was prepared from undifferentiated or differentiated cells, and the *neo* mRNA level was determined by Northern (RNA) blotting (Fig. 5). Two pSV2neo-transformed P19 cell lines (cell lines 101 and 102) and a group I cell line (044) showed similar levels of *neo* mRNA before and after differentiation. On the other hand, in most of the group II cell lines (015, 017, 023, 042, 052, and 060), the *neo* mRNA level was greatly reduced following cell differentiation. In particular, the decrease in *neo* mRNA was remarkable in two cell lines, 023 and 052. Cell lines 024 and 060 showed only a modest decrease in *neo* mRNA level.

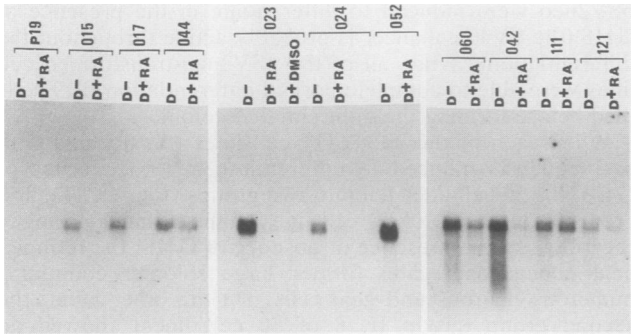

FIG. 5. The level of *neo* mRNA before and after differentiation. Poly(A)$^+$ RNA was extracted from undifferentiated (D$^-$) cells that had grown in the medium lacking G418 for 3 to 4 days. Poly(A)$^+$ RNA was also prepared from differentiated (D$^+$) cells. Differentiation was induced by retinoic acid (RA) or by DMSO in one case. G418 was not used during differentiation. Approximately 10 μg of poly(A)$^+$ RNA was loaded on a formamide-agarose gel, blotted, and hybridized to a *neo*-specific probe (the *Hin*dIII-*Bam*HI 2-kb fragment of pA10neo). The names of the cell lines are indicated at the top. P19 is the untransfected parental cell line. Cell lines 111 and 121 are transformed with pSV2neo. Cell line 044 belongs to group I, while cell lines 015, 017, 023, 024, 042, 052, and 060 belong to group II. Cell line 034 (group II) also showed a reduced *neo* mRNA after differentiation (data not shown).

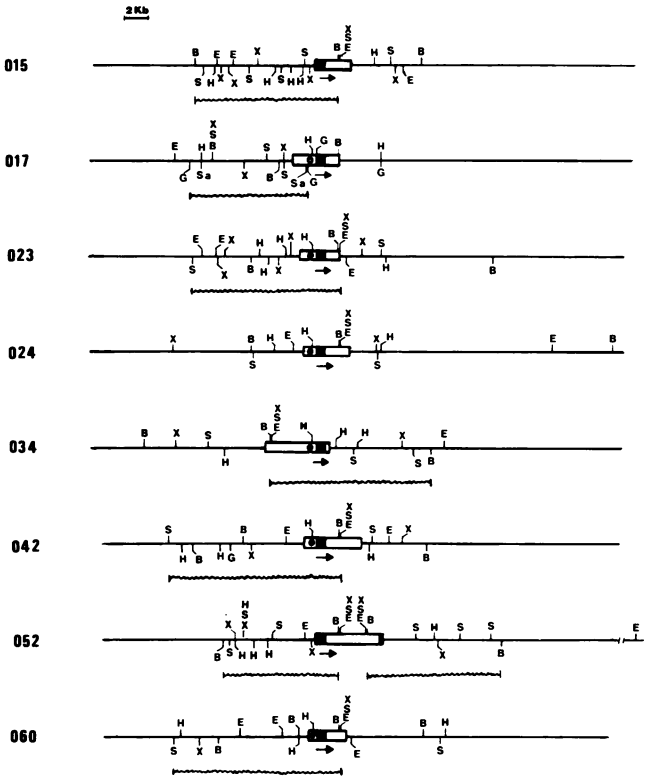

FIG. 6. Cloning of mouse DNA sequences flanking the integrated *neo* gene. Restriction maps obtained with genomic DNA and cloned DNA are summarized. The names of cell lines are indicated on the left. Symbols: ∿∿∿, eight cloned DNA fragments; ▬, mouse DNA; ▭, pA10neo sequences; ■, *neo* coding sequences; ○, the SV40 early promoter derived from pA10neo; →, direction of *neo* gene transcription. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; G, *Bgl*II; S, *Sst*I; X, *Xba*I.

These results suggest that in most of the group II cell lines, the *neo* gene is active when cells are undifferentiated but the gene is repressed when cells are induced to differentiate. Presumably, in those cell lines, the *neo* gene is activated by a regulatory element that is active in the stem cells but inactive in the differentiated cells. Therefore, the seven group II cell lines were subjected to subsequent molecular cloning.

Cloning of mouse DNA sequences flanking the *neo* gene. To clone the mouse DNA sequences flanking the *neo* gene, we first determined the restriction map of the integration site (Fig. 6) by genomic Southern blotting. As predicted by the dot hybridization analysis, all the group II cell lines contained a single copy of the *neo* gene, except that 052 had one intact *neo* gene plus a partial (nonfunctional) *neo* gene sequence. The portion of the enhancer trap sequence remaining in the cell lines varied; however, the *Hin*dIII-*Bam*HI 2-kb region that includes the *neo* coding sequence and the poly(A) additional signal sequence remained intact in five cell lines (017, 023, 024, 042, and 060). The *Hin*dIII site was deleted in two cell lines, 015 and 052. Since the SV40 early promoter is located approximately 100 base pairs upstream of the *Hin*dIII site, the SV40 early promoter must be deleted and replaced by an endogenous promoter in these two cell lines. Consistent with this is the observation that the *neo* mRNA expressed in 052 was slightly smaller than that detected in other cell lines (Fig. 5).

TABLE 1. NEO$^+$ transformation activity in P19 cells and NIH 3T3 cells[a]

DNA	Transformation frequency in cells:	
	P19	NIH 3T3
pOneo	4	5
pA10neo	6	18
015pBam12	480	12
052-1pBam9.5	240	15
034pBam14	620	320
023pSst13	5	ND
042pSst14	20	ND
060pSst13	8	ND
pSV2neo	250	580

[a] Various plasmids were transfected to P19 cells or NIH 3T3 cells, and their NEO$^+$-transforming activities were determined as described in Materials and Methods. The transformation frequency is expressed as the number of G418-resistant colonies per 5×10^5 cells. ND, Not determined.

By using a standard cloning procedure, mouse DNA sequences flanking the *neo* gene were cloned from seven group II cell lines (Fig. 6). Only one side of the integration site was cloned from cell lines 015, 017, 023, 034, 042, and 060, while both sides were cloned from cell line 052. The cloned phages were accordingly designated λ015, λ017, λ023, λ034, λ042, λ052-1, λ052-2, and λ060 (λ052-1 and λ052-2 contain the left and right sides of the *neo* integration site, respectively). The restriction maps of the cloned DNA were consistent with those determined by genomic Southern blotting, indicating that the structure of the DNA sequences was preserved throughout the cloning process. It should be noted that, as predicted, the SV40 early promoter was maintained in λ023, λ034, λ042, and λ060, while it was deleted in λ015 and λ052.

The eight cloned fragments (Fig. 6) showed distinct restriction maps. It appears that the integration sites of the enhancer trap in those cell lines are different from each other, although it is possible that some cloned fragments are derived from the same chromosomal locus.

Cloned DNA fragments are transcriptionally active in the stem cells. If the integrated *neo* gene is activated by a regulatory element such as an enhancer, then the cloned DNA should possess a high transcriptional activity in the stem cells. The DNA fragments cloned in λ015, λ023, λ034, λ042, λ052-1, and λ060 conveniently contain an intact *neo* coding sequence. Therefore, these cloned fragments were subcloned in a plasmid and transfected to P19 cells, and expression of the *neo* gene was monitored by a stable transformation assay (Fig. 7 and Table 1). pA10neo (an enhancerless construct) and p0neo (an enhancerless and promoterless construct), which served as negative controls, showed low transformation efficiencies. We observed a few false NEO$^+$ colonies with pUC12 DNA. pSV2neo, which served as a positive control, showed a high transformation efficiency. Among six cloned fragments tested, three (015pBam12, 034pBam14, and 052-1pBam9.5) showed very high transformation frequencies. Their activities were comparable with or even higher than that of pSV2neo. Thus, these three cloned fragments must contain a regulatory element that can activate the expression of the *neo* gene. On the other hand, the remaining three fragments (023pSst12, 042pSst13, and 060pSst14) showed efficiencies as low as that of pA10neo.

To locate the regulatory element responsible for *neo* gene activation, various deletion mutants were constructed from 015pBam12, 034pBam14, and 052-1pBam9.5, and their ac-

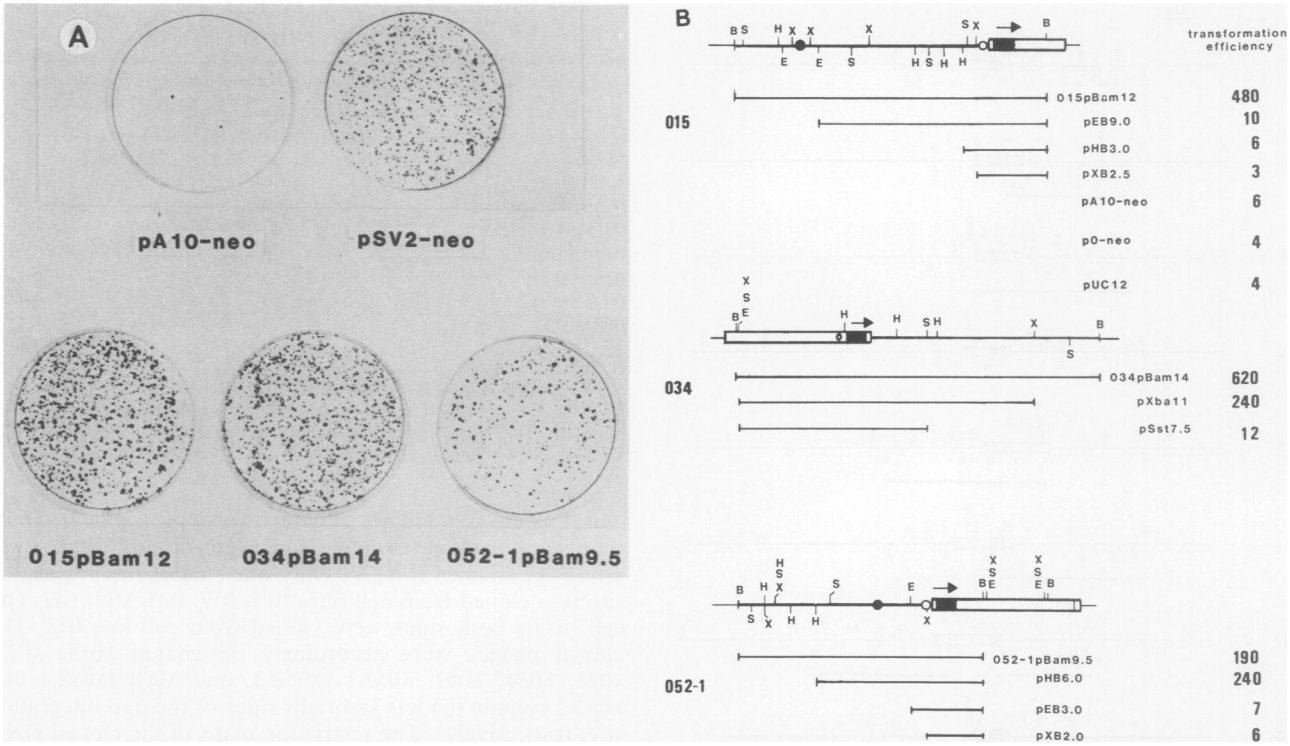

FIG. 7. Presence of activating elements in three cloned DNA fragments. (A) The entire lengths of the inserts cloned in phages (shown by the wavy lines in Fig. 6) were subcloned at the corresponding sites of pUC12: the BamHI 12-kb fragment from λ015, the *Sst* 13-kb fragment from λ023, the *Sst*I 14-kb fragment from λ042, the *Bam*HI 9.5-kb fragment from λ052-1 (the clone containing the left side of the integration [Fig. 6]), and the *Sst*I 13-kb fragment from λ060. The *Bam*HI 14-kb fragment cloned in λ034 was self-ligated and cloned as a plasmid. The resulting plasmids were transfected to P19 cells, and NEO^+ transformation efficiencies were determined as described in Materials and Methods. Representative Giemsa-stained plates are shown here. pA10neo and pSV2neo are negative and positive controls, respectively. Note that three cloned DNA fragments (015pBam12, 034pBam14, and 052pBam9.5) showed a high NEO^+-transforming activity. (B) The restriction maps of the 015, 034, and 052 cloned fragments are shown above each section. Various deletion fragments (indicated by the solid lines under the maps) were subcloned in a plasmid, and the resulting mutants were tested for NEO^+-transforming activity. The transformation activity is expressed as the number of NEO^+ colonies per 5×10^5 P19 cells. Symbols: ●, approximate locations of activating elements determined in this study; ○, SV40 early promoter in the 034 cloned DNA and a putative endogenous promoter found in the 015 and 052 cloned DNA. For restriction site abbreviations, see legend to Fig. 6.

tivities were determined by the stable transformation assay (Fig. 7B). In the case of the 015 constructs, the activity was lost in pEB9.0, pHB3.0, and pXB2.5, indicating that the activating element is located in the left 3-kb region (Fig. 7B). As described previously, the original SV40 early promoter is deleted in 015pBam12. Since the size of the *neo* mRNA expressed in cell line 015 was similar to that expressed in pSV2neo-transformed cells (Fig. 5), the *neo* gene in 015pBam12 should be transcribed from an endogenous promoter that is located near the integration site (Fig. 7B). However, results of experiments with the deletion mutants suggest that the endogenous promoter alone is not enough to confer the high level of expression of the *neo* gene. The other element located in the left 3-kb region is probably an enhancerlike element, since it functions from a distance of at least 6 kb.

A similar conclusion can be drawn for 052-1pBam9.5. The SV40 early promoter was deleted in 052pBam9.5. Since the *neo* mRNA detected in cell line 052 was slightly smaller than that expressed in pSV2neo-transformed cells (Fig. 5), the *neo* gene in 052-1pBam9.5 must be transcribed from an endogenous promoter (Fig. 7B) located very near the integration site. In the deletion analysis, pHB6.0 retained full activity. However, pEB3.0, which contains 0.7 kb of upstream mouse sequences (therefore, it should include the endogenous promoter) lost the entire activity. These results indicate that another element (probably an enhancer) is located in the *Hin*dIII-*Eco*RI 3-kb region (Fig. 7B).

In contrast to 015pBam12 and 052-1pBam9.5, 034pBam14 retained the original SV40 early promoter. This observation and the size of the *neo* mRNA expressed in cell line 034 (Fig. 5) suggest that the *neo* gene is transcribed from the SV40 early promoter. In the deletion analysis, 034pXba11 still retained 40% of the activity of pBam14, while pSst7.5 completely lost the activity. These results suggest that one or multiple enhancers are located in the *Sst*I-*Bam*HI 6.5-kb region (Fig. 7B).

015 and 052 cloned DNA fragments are inactive in differentiated cell types. We next asked whether the activities of the 015, 034, and 052 clones are specific to the stem cells. The following two experiments were performed to test this. First, P19 cells transformed with 015pBam12, 034pBam14, or 052-1pBam9.5 were induced to differentiate in the absence or presence of G418 (Fig. 8). Differentiation was induced by either retinoic acid or DMSO. In the absence of G418, 015pBam12-transformed and 052-1pBam9.5-transformed cells could differentiate normally; they differentiated to neurons and glial cells with retinoic acid (Fig. 8a and b), whereas they differentiated to beating muscle cells with DMSO (Fig. 8d and e). However, in the presence of G418,

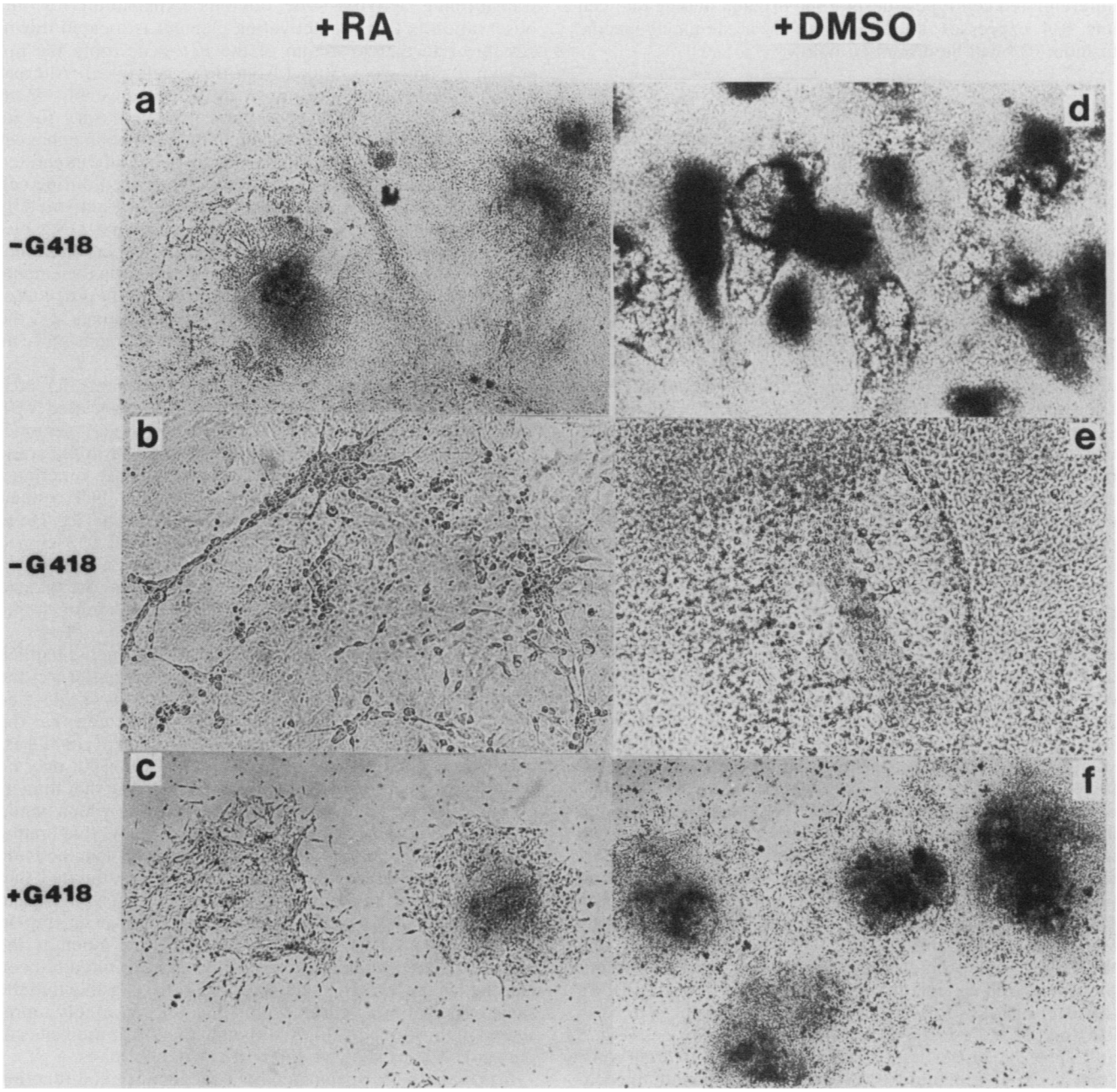

FIG. 8. 015 and 052 regulatory elements are inactivated during differentiation. NEO^+ colonies (200 to 500) transformed with 015pBam12, 034pBam14, or 052pBam9.5 were pooled and maintained in the selective medium for 4 days. Each pool was then induced to differentiate by either retinoic acid (a, b, and c) or DMSO (d, e, and f) in the absence (a, b, d, and e) or presence (c and f) of G418. Phase-contrast photographs taken 8 days after the induction are shown. Higher magnification of panels a and d are shown in b and e, respectively. The characteristics of 015pBam12-transformed cells are shown. The cells transformed with 052pBam9.5 also showed a similar pattern. On the other hand, the 034pBam14-transformed cells differentiated normally regardless of the absence or presence of G418 (data not shown).

both transformants were unable to differentiate normally; the aggregates did not grow large, and no obvious neurons (Fig. 8c) or beating muscle cells (Fig. 8f) were detected. This observation suggests that the expression of the *neo* gene in 015pBam12 and 052-1pBam9.5 is inactivated after the induction of differentiation. In contrast, the 034pBam14-transformed cells differentiated normally regardless of the presence or absence of G418 (data not shown).

Second, we determined the activity of the three cloned DNAs in NIH 3T3 cells, a mouse fibroblast cell line (Table 1). pSV2neo (a positive control) showed a high NEO^+-transforming activity, while pA10neo and pOneo (negative controls) showed much lower activities. As predicted from the previous observations (Fig. 8), 015pBam12 and 052-1pBam9.5 were inactive in NIH 3T3 cells, whereas 034p Bam14 was active.

These two lines of evidence indicate that the 015 and 052 cloned DNAs were inactive in differentiated cell types and suggests that their activities are restricted to embryonic stem cells. On the other hand, 034pBam14 was active in both

differentiated cell types, despite the original finding that cell line 034 expresses the *neo* gene in a stem-cell-specific fashion. 034 will be discussed below.

DISCUSSION

We have described functional cloning of mouse chromosomal loci and of transcriptional regulatory elements that are specifically active in EC stem cells. This is clearly demonstrated by the two cloned fragments (015 and 052), which were active in undifferentiated P19 cells but were inactive in the differentiated P19 cells and NIH 3T3 cells.

It was fortunate for us but rather unexpected that the enhancerless SV40 promoter is enhancer dependent in P19 cells, since it has been shown by others (6) that the same promoter is enhancer independent in other EC cells, such as F9 and PCC4. This discrepancy could well be due to the different cell lines used. Some EC cell lines including F9 are known to possess adenovirus E1a-like activity (9, 10). As suggested by Gorman et al. (6), it may be the same factor that renders the SV40 early promoter enhancer independent in F9 cells. If this is the case, the E1a-like activity should be absent in P19 cells. Alternatively, the discrepancy could be due to the difference in the construction of the enhancerless promoter (deletion of the SV40 enhancer is more extensive in our pA10neo than in their construct).

In 015 and 052 cloned fragments, the original SV40 early promoter was deleted and seems to be replaced by an endogenous promoter. To confirm the presence of an endogenous promoter near the integration site, we have recently determined the transcription initiation site of the *neo* gene in cell lines 015 and 052 by primer extension and S1 nuclease mapping. The results indicate that the transcription indeed starts at the mouse DNA sequences very near the integration site (200 and 160 base pairs upstream of the integration site for 015 and 052, respectively [H. Hamada, unpublished data). Furthermore, when the SV40 enhancer was inserted into 015pHB3.0 and 052pEB3.0 (the deletion mutants shown in Fig. 7B), such constructs regained a high NEO^+-transforming activity (unpublished data). From these observations, we conclude that the *neo* gene is driven by an endogenous promoter which is located near the integration site. The deletion analysis (Fig. 7B), however, shows that the endogenous promoter alone could not function without the enhancerlike element that is located at a distance of 1 kb (for 052) or 6 kb (for 015). The promoter enhancer units in 015 and 052 were inactive in differentiated cell types (Fig. 8 and Table 1), suggesting that they are specific to embryonic stem cells. It remains to be seen which element, the enhancer or the promoter, determines the specificity.

Cell line 034 was initially selected for its inability to differentiate in the presence of G418. Northern blotting confirmed that the *neo* gene in cell line 034 was repressed after the induction of differentiation (data not shown). However, the *Bam*HI 14-kb fragment cloned from cell line 034 was unexpectedly active in the differentiated P19 cells and NIH 3T3 cells (Table 1). One possibility that would reconcile these results is that there is, in addition to the activating element detected in the *Bam*HI 14-kb region, a negatively controlling element near the *neo* integration site but the latter element is not included in the BamHI 14-kb fragment. This negatively regulating element would not function in the stem cells but would become functional when cells are induced to differentiate; thus, it would repress the *neo* gene.

Three of the DNA fragments cloned from group II cell lines (023, 042, and 060) did not possess a high NEO^+-transforming activity. One obvious explanation for this observation is that the activating element is located downstream or further upstream of the *neo* gene (only the upstream regions were cloned from those cell lines). Alternatively, if an enhancer is required for the initial establishment of an active status of a gene but is not necessary for its maintenance, as suggested for the immunoglobulin enhancer (12, 23), it is possible that the *neo* gene is initially integrated near an enhancer which is subsequently deleted during cell division. Finally, the inactive *neo* gene could be activated by a mechanism which does not involve *cis*-acting DNA elements such as enhancers. In any case, the observation that three of six cloned DNAs contained activating elements again verifies the high efficiency of our selection procedure. Given the fact that the two most interesting clones lost the original SV40 promoter, a promoterless gene might serve as a better trap in future screenings.

It seems most likely that the regulatory elements detected in the 015 and 052 cloned fragments are associated with embryonic-stem-cell-specific genes in the mouse genome, although such genes are yet to be identified. One could argue that our selection method may select not only functional elements linked to genes but also nonfunctional cellular sequences which happen to possess enhancer activity. However, we should emphasize that the 015 and 052 cloned DNAs contained two regulatory elements, enhancer and promoter. The linkage of those elements in the cloned fragments makes it unlikely that they are nonfunctional sequences.

From the 30 NEO^+ transformants, we obtained eight cell lines that showed an abnormal differentiation pattern in the presence of G418 (Fig. 3 and 4). Further analysis of *neo* gene expression by Northern blotting (Fig. 5) indicated that the *neo* gene was integrated into stem-cell-specific loci in at least six of the eight cell lines. This frequency (6 of 30) may be higher than what one would expect, considering that numerous genes are active in stem cells, only a few of which would be stem cell specific. However, it is conceivable that among numerous active loci only a small number can activate nearby integrated *neo* gene. We previously estimated that there are approximately 50 different loci in the human genome that can activate an integrated enhancer trap in HeLa cells (8). This implies that specific loci (such as the ones associated with a strong enhancer) have already been enriched in the 30 NEO^+ cell lines. It is also possible that the same or neighboring chromosomal loci are repeatedly represented in those six cell lines, although the integration sites of the six cell lines showed different restriction maps.

It is tempting to speculate about the identity and function of the genes that are associated with the regulatory elements found in the 015 and 052 cloned DNAs. They may be marker genes which do not play a regulatory role in embryogenesis. Alternatively, they may be regulatory genes that play a key role in the early stage of embryogenesis. For example, there may be a gene(s) that is necessary to maintain pluripotency of the stem cells and whose inactivation triggers subsequent differentiation. The presence of such regulatory genes has been postulated by somatic cell hybridization studies (15, 22). We speculate that a candidate for the regulatory genes would be one of the stem-cell-specific genes that are expressed in EC stem cells but are rapidly repressed when EC cells are induced to differentiate. In this regard, it is interesting that the expression of the *neo* gene in most of the group II cell lines was greatly reduced after differentiation. We are currently cloning mouse chromosomal loci corresponding to the *neo* gene integration sites. Stem-cell-specific

genes will be identified in the cloned loci by searching them for mRNA coding sequences expressed in EC stem cells.

There have been several attempts by others to isolate stem-cell-specific genes (e.g., 14). Their methods included screening of a cDNA library made from undifferentiated EC cells by differential hybridization. However, the previous attempts failed to identify genes that are strictly specific to undifferentiated stem cells. If such genes exist, as suggested by the present study, it is not certain why they were previously overlooked. Perhaps the corresponding mRNAs exist in small amounts in EC stem cells and are easily overlooked because of much more abundant repeated sequences whose expression is developmentally controlled. Another approach taken by others was to use a marker gene (such as the *neo* gene) linked to the murine leukemia virus long terminal repeat LTR (1, 21). Since the murine leukemia virus LTR is repressed in EC cells (6), such a gene can be used as a trap, like pA10neo in this study, for selecting elements that can reactivate the LTR. However, the major disadvantage of this method is that it can not select stem-cell-specific elements, since the LTR is active in differentiated cells. In fact, the enhancerlike elements selected by this method were able to reactivate the function of the murine leukemia virus LTR in EC stem cells, but the specificity was not determined (1, 21). In this regard, it should be mentioned that mouse DNA fragments selected by this method (1, 21) and those obtained in this study show different restriction maps.

ACKNOWLEDGMENTS

We thank E. Natt and G. Scherer for EMBL12, F. Swift for technical assistance, and B. H. Younghusband for improving the manuscript.

This work was supported by grants from the Medical Research Council of Canada and the National Cancer Institute of Canada. K. B. was supported by a predoctoral fellowship from the Cancer Society of Canada. M.W.M. is a National Cancer Institute of Canada Terry Fox Cancer Research Scientist. H.H. was supported by a development grant from the Medical Research Council of Canada.

LITERATURE CITED

1. **Barklis, E., R. C. Mulligan, and R. Jaenisch.** 1986. Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell **47**:391–399.
2. **Cooper, A. R., A. Taylor, and B. L. M. Hogan.** 1983. Changes in the ratio of lamin and entactin synthesis in F9 embryonal carcinoma cells treated with retinoic acid and cyclic AMP. Dev. Biol. **99**:510–514.
3. **Dziadek, M., and E. Adamson.** 1978. Localization and synthesis of alpha fetoprotein in post-implantation mouse embryos. J. Embryol. Exp. Morphol. **43**:289–295.
4. **Edwards, M. K. S., J. F. Harris, and M. W. McBurney.** 1983. Induced muscle differentiation in an embryonal carcinoma cell line. Mol. Cell. Biol. **3**:2280–2286.
5. **Gardner, R. L., and J. Rossant.** 1979. Investigation of the fate of the 4–5 day post-coitum mouse inner cell mass cells by blastocyst injection. J. Embryol. Exp. Morphol. **52**:141–152.
6. **Gorman, C. M., P. W. Rigby, and D. P. Lane.** 1985. Negative regulation of viral enhancers in undifferentiated embryonic stem cells. Cell **42**:519–552.
7. **Hamada, H.** 1986. Activation of an enhancerless gene by chromosomal integration. Mol. Cell. Biol. **6**:4179–4184.
8. **Hamada, H.** 1986. Random isolation of gene activator elements from the human genome. Mol. Cell. Biol. **6**:4185–4194.
9. **Hen, R., E. Borrelli, C. Fromental, P. Sassone-Corsi, and P. Chambon.** 1986. A mutated polyoma virus enhancer which is active in undifferentiated embryonal carcinoma cells is not repressed by adenovirus-2 E1a products. Nature (London) **321**: 249–251.
10. **Imperiale, M. J., H. Kao, L. T. Feldman, J. R. Nevins, and S. Strickland.** 1984. Common control of the heat shock gene and early adenovirus genes: evidence for a cellular E1a-like activity. Mol. Cell. Biol. **4**:867–874.
11. **Jones-Villeneuve, E. M. V., M. A. Rudnicki, J. F. Harris, and M. W. McBurney.** 1983. Retinoic acid-induced neural differentiation of embryonal carcinoma cells. Mol. Cell. Biol. **3**:2271–2279.
12. **Klein, S., T. Gerster, D. Picard, A Radbruch, and W. Schaffner.** 1985. Evidence for transient requirement of the IgH enhancer. Nucleic Acids Res. **13**:8901–8912.
13. **Laimins, L. A., P. Gruss, R. Pozzatti, and G. Khoury.** 1984. Characterization of enhancer elements in the long terminal repeat of Moloney murine sarcoma virus. J. Virol. **49**:183–189.
14. **Levine, R. A., G. J. LaRosa, and L. J. Gudas.** 1984. Isolation of cDNA clones for genes exhibiting reduced expression after differentiation of murine teratocarcinoma stem cells. Mol. Cell. Biol. **4**:2142–2150.
15. **Littlefield, J. W., and J. S. Felix.** 1982. Rescue of terminally differentiating teratocarcinoma cells by fusing to undifferentiated parental cells. Somatic Cell Genet. **8**:743–757.
16. **McBurney, M. W., and B. J. Rogers.** 1982. Isolation of male embryonal carcinoma cells and their chromosome replication patterns. Dev. Biol. **89**:503–508.
17. **Natt, E., and G. Scherer.** 1986. EMBL12, a new lambda replacement vector with sites for SalI, XbaI, BamHI, SstI and EcoRI. Nucleic Acids Res. **14**:7128.
18. **Rudnicki, M. A., and M. W. McBurney.** 1987. Cell culture methods and induction of differentiation of embryonal carcinoma cell lines, p. 19–49. *In* E. J. Robertson (ed.), Teratocarcinomas and embryonic stem cells: a practical approach. IRL Press, Oxford.
19. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early promoter. J. Mol. Appl. Genet. **1**:327–341.
20. **Swift, V. F., K. Bhat, B. H. Younghusband, and H. Hamada.** 1987. Characterization of a cell-type specific enhancer found in human papillomavirus type 18 genome. EMBO J. **5**:1339–1344.
21. **Taketo, M., and M. Tanaka.** 1987. A cellular enhancer of retrovirus gene expression in embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA **84**:3748–3752.
22. **Wray, L. K., and H. K. Jaeckle.** 1983. Protein synthesis by rat hepatoma x mouse teratocarcinoma somatic cell hybrids. Somatic Cell Genet. **9**:111–119.
23. **Zaller, D. M., and L. A. Eckhardt.** 1985. Deletion of a B-cell specific enhancer affects transfected, but not endogenous, immunoglobulin heavy-chain gene expression. Proc. Natl. Acad. Sci. USA **82**:5088–5092.