A Large-Scale Gene-Trap Screen for Insertional Mutations in Developmentally Regulated Genes in Mice

Wolfgang Wurst,* Janet Rossant,* Valerie Prideaux,* Malgosia Kownacka,* Alexandra Joyner,*,† David P. Hill,* Francois Guillemot,* Stephan Gasca,*,† Dragana Cado,*,[1] Anna Auerbach* and Siew-Lan Ang*,[2]

** Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada M5G 1X5 and † Department of Molecular and Medical Genetics, University of Toronto, Toronto, Canada*

Manuscript received May 5, 1994
Accepted for publication October 11, 1994

ABSTRACT

We have used a gene-trap vector and mouse embryonic stem (ES) cells to screen for insertional mutations in genes developmentally regulated at 8.5 days of embryogenesis (dpc). From 38,730 cell lines with vector insertions, 393 clonal integrations had disrupted active transcription units, as assayed by β-galactosidase reporter gene expression. From these lines, 290 clones were recovered and injected into blastocysts to assay for reporter gene expression in 8.5-dpc chimeric mouse embryos. Of these, 279 clones provided a sufficient number of chimeric embryos for analysis. Thirty-six (13%) showed restricted patterns of reporter-gene expression, 88 (32%) showed widespread expression and 155 (55%) failed to show detectable levels of expression. Further analysis showed that approximately one-third of the clones that did not express detectable levels of the reporter gene at 8.5 dpc displayed reporter gene activity at 12.5 dpc. Thus, a large proportion of the genes that are expressed in ES cells are either temporally or spatially regulated during embryogenesis. These results indicate that gene-trap mutageneses in embryonic stem cells provide an effective approach for isolating mutations in a large number of developmentally regulated genes.

THE ability to carry out large scale screens for developmental mutations has proven essential in unravelling the genetic programs underlying embryogenesis in such organisms as *Drosophila melanogaster* and *Caenorhabditis elegans*. These types of screens in mammals are made difficult by the large genome size and the development of the embryo inside the mother's uterus. Furthermore, the cost and space required to house large numbers of animals and the relatively long breeding period have limited the undertaking of large scale screens.

ES cell technology has permitted investigators with modest-sized animal facilities to enter the field of mammalian genetics by allowing the bulk of genetic manipulation to occur *in vitro*. To date, the most popular approach using this technology involves targetted mutagenesis of genes via homologous recombination in ES cells (CAPECCHI 1989; KOLLER and SMITHIE, 1992). This has allowed mutational analysis of the function of molecularly identified genes that are predicted to be developmentally important. Most often, such genes are identified by homology either to genes that have been shown to be developmentally important in other species or to genes that contain functionally conserved protein domains of interest. Although this method of selecting candidate genes has proven very successful in identifying and mutating important developmental genes, it does not provide access to genes that have not been characterized at the molecular level.

We and others have described an efficient means of identifying and mutating novel genes in ES cells by the introduction of vectors into ES cells that drive β-galactosidase reporter-gene expression from endogenous cellular promoters (reviewed in GOSSLER and ZACHGO 1993; HILL and WURST 1993a). In this screen, we used the type of vector that serves as an artificial exon after insertion into an endogenous transcription unit (GOSSLER *et al.* 1989; FRIEDRICH and SORIANO 1991). When cell lines containing this vectors are used to make chimeric embryos, the localization of β-galactosidase activity generated by the fusion protein is very similar to the expression pattern of the endogenous gene found at the site of insertion (SKARNES *et al.* 1992). Generation of fusion transcripts also makes it possible to directly clone the transcribed region of the disrupted host gene using the rapid amplification of cDNA ends protocol (RACE, SKARNES *et al.* 1992).

A large number of genes that are critical for murine

Corresponding author: David Hill, Herman B. Wells Center for Pediatric Research, Riley Hospital for Children, Room 2663, 702 Barnhill Rd., Indianapolis, IN 46202.

[1] *Present address:* 447 Life Science Addition, University of California, Berkeley, CA 94720-0001.

[2] All authors contributed equally to the project.

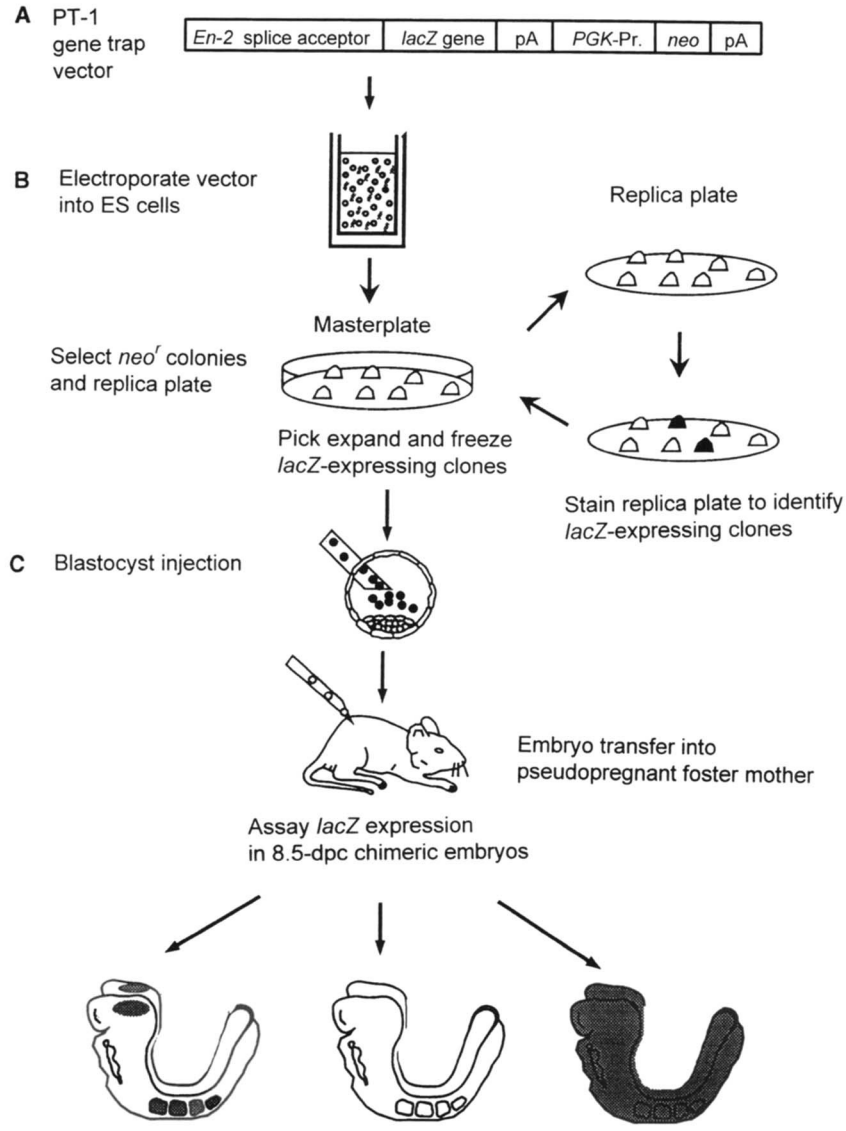

FIGURE 1.—Schematic representation of the gene-trap vector and the method used to analyze β-galactosidase expression patterns in 8.5-dpc ES cell chimeras.

embryonic development display specific spatial and temporal regulation during embryogenesis (*e.g.*, Hox genes, reviewed in KRUMLAUF 1993). Using the rationale that developmentally regulated genes may code for developmentally important molecules, we conducted a large scale screen for insertional mutations into genes that are developmentally regulated during mouse embryogenesis. We reasoned that the results of such a screen would reveal information on the fundamental domains of gene expression during development, provide lineage markers for future embryological experiments and provide a large number of candidate mutations affecting the development of tissues marked by reporter gene expression. In this study, we report on such a screen conducted to identify and mutate genes that are expressed around the time of the establishment of the basic body plan (8.5 dpc) in the mouse embryo. We have characterized the expression patterns of nearly 300 different genes that are expressed in embryonic stem cells. We discuss the implications of this screen in terms of the types of gene expression patterns that exist in the early mouse embryo and the feasability of using gene-trap vectors in more extensive mutagenic studies.

MATERIALS AND METHODS

Vectors: The gene trap vector used in this study, PT-1, is a modification of the GT4.5 vector used by GOSSLER and coworkers (1989). The original vector contains the splice acceptor sequence from the *En-2* gene upstream of the *Escherichia coli* β-galactosidase gene (*lacZ*), lacking its own ATG. For this screen, we modified GT4.5 by replacing the β-actin promoter-driven neomycin resistance gene with a neomycin resistance gene driven by the *PGK-1* promoter and containing an SV-40 polyadenylation signal (BOER *et al.* 1990, Figure 1A). This modification resulted in a 5-fold increase in the number of *neo*^{r} colonies per electroporation without affecting the proportion of β-galactosidase expressing colonies among the re-

TABLE 1

Summary of integrations tested in 8.5-dpc embryos

Expression pattern	No. of ES cell clones
Spatially restricted expression	36 (13)
Widespread Expression	88 (32)
No expression detected	155 (55)
Total no. of clones examined	279 (100)

Values in parentheses are percentages.

sistant colonies. Before introduction into cells, the vector DNA was linearized by digestion with *Hin*dIII. The reaction mixture was heated to 90°C for 15 min and the linearized DNA was ethanol precipitated. DNA was resuspended in phosphate buffered saline at a concentration of 1 mg/ml for electroporation.

Electroporation of ES cells and *in vitro* screening: Electroporation and screening of ES cells were performed as described previously (HILL and WURST 1993b; WURST and JOYNER 1993). After 8–10 days of G418 selection, when *neo*^{r} colonies were readily apparent, the colonies were replica plated (GOSSLER *et al.* 1989; GOSSLER and ZACHGO 1993). Colonies that showed any β-galactosidase staining, either scattered or throughout the colony, were picked, expanded, retested for β-galactosidase activity and frozen away for later analysis in chimeras.

Production and analysis of chimeras: β-galactosidase positive clones were thawed, grown for one week and then injected individually into blastocysts obtained from outbred CD1 mice (Charles River Laboratories, Quebec). For 139 out of 290 clones, 30–40 blastocysts were injected with 12–15 cells each and transferred into the uteri of three recipient females on the third day of pseudopregnancy. Two recipients from each clone were sacrificed at 8.5 dpc, when embryos were at early somite stages (4–15 somites). Dissected embryos and their extraembryonic membranes were fixed and stained for β-galactosidase activity as previously described (GOSSLER and ZACHGO 1993; HILL and WURST 1993b). The numbers of embryos that expressed β-galactosidase and the pattern of the staining were recorded, and all embryos displaying developmental regulation of β-galactosidase expression were photographed. The third recipient was allowed to continue the pregnancy until 12.5 dpc when chimerism could be scored by the presence of eye pigmentation derived from the agouti D3 ES cell line in the albino CD1 host embryos (NAGY *et al.* 1990). Most of the 12.5-dpc embryos were also stained for β-galactosidase activity. The remaining 151 clones were used to generate 20–30 embryos in two recipients, which were sacrificed and analyzed for reporter gene expression on 8.5 dpc. Three chimeras showing identical expression patterns were considered an acceptable minimum because patterns were reproducible from embryo to embryo despite varying degrees of mosaicism (GOSSLER *et al.* 1989). In cases where there was doubt concerning the pattern of reporter gene expression, the injections were repeated until three chimeras showing identical patterns were obtained. Information on the majority of clones was derived from more than three chimeras, and data based on two chimeras were reported in a few cases where expression was clearly ubiquitous.

Production of ES-tetraploid chimeras: A few clones that displayed interesting patterns were subjected to analysis after aggregation with tetraploid host embryos (NAGY *et al.* 1990; NAGY and ROSSANT 1993). In such embryos, the ES cells outcompete the compromised tetraploid host cells during development, resulting in embryonic tissues that are derived solely from ES cell descendants (NAGY *et al.* 1990). The absence of mosaicism allows for more precise confirmation of the β-galactosidase staining pattern.

RESULTS

Isolation of β-galactosidase-expressing ES cell clones: The vector PT1 (Figure 1A), a modified version of GT4.5 (see MATERIALS AND METHODS) was used to generate β-galactosidase-expressing ES clones. Clones containing vector DNA were isolated by selection for expression of the neomycin resistance (*neo*^{r}) gene, driven by the PGK-1 promoter. These *neo*^{r} colonies were then replica plated and assayed for β-galactosidase activity (Figure 1B, GOSSLER *et al.* 1989). Colonies that were expressing β-galactosidase usually contained a mixture of undifferentiated and partially differentiated ES cells, and β-galactosidase staining was observed in either or both compartments. From 38,730 *neo*^{r} clones, 393 (~1%) β-galactosidase-expressing clones were identified, from these 393 clones, we were able to establish 300 cell lines that were expanded and kept as frozen stocks.

Detailed examination of subcellular localization of β-galactosidase activity was performed in 208 expressing lines and showed that β-galactosidase activity was localized in six different ways. In 122 (59%) of the clones, β-galactosidase activity was observed throughout the cytoplasm, and in 46 (22%) of the clones, staining was restricted to distinct dots in the cytoplasm. Staining in both the cytoplasm and dots was seen in nine (4%) clones. Nuclear staining was observed in 26 (13%) of the clones and nuclear plus cytoplasmic staining was seen in 4 (2%) of the clones. One clone showed β-galactosidase staining around the edges of the cells.

When these 208 clones were examined for β-galactosidase expression patterns within the partially differentiated colonies, the expression in 13 clones appeared restricted to partially differentiated cells based on gross cellular morphology. Only two clones showed β-galactosidase expression restricted to undifferentiated ES cells. The remainder of the clones showed β-galactosidase expression throughout the colony.

Reporter gene expression in 8.5-dpc chimeras: ES clones (290) expressing the reporter gene were injected into CD1 blastocysts, and chimeric embryos were assayed for β-galactosidase expression at the 4–15 somite stage of embryogenesis. Of the 290 clones injected, 279 lines provided enough chimeric embryos to satisfy our criteria for inclusion in this study (see MATERIALS AND METHODS). When chimeric embryos were assayed for β-galactosidase expression at 8.5 dpc, 36 (13%) of the lines displayed spatially regulated expression of β-galactosidase, 88 (32%) of the lines displayed wide-

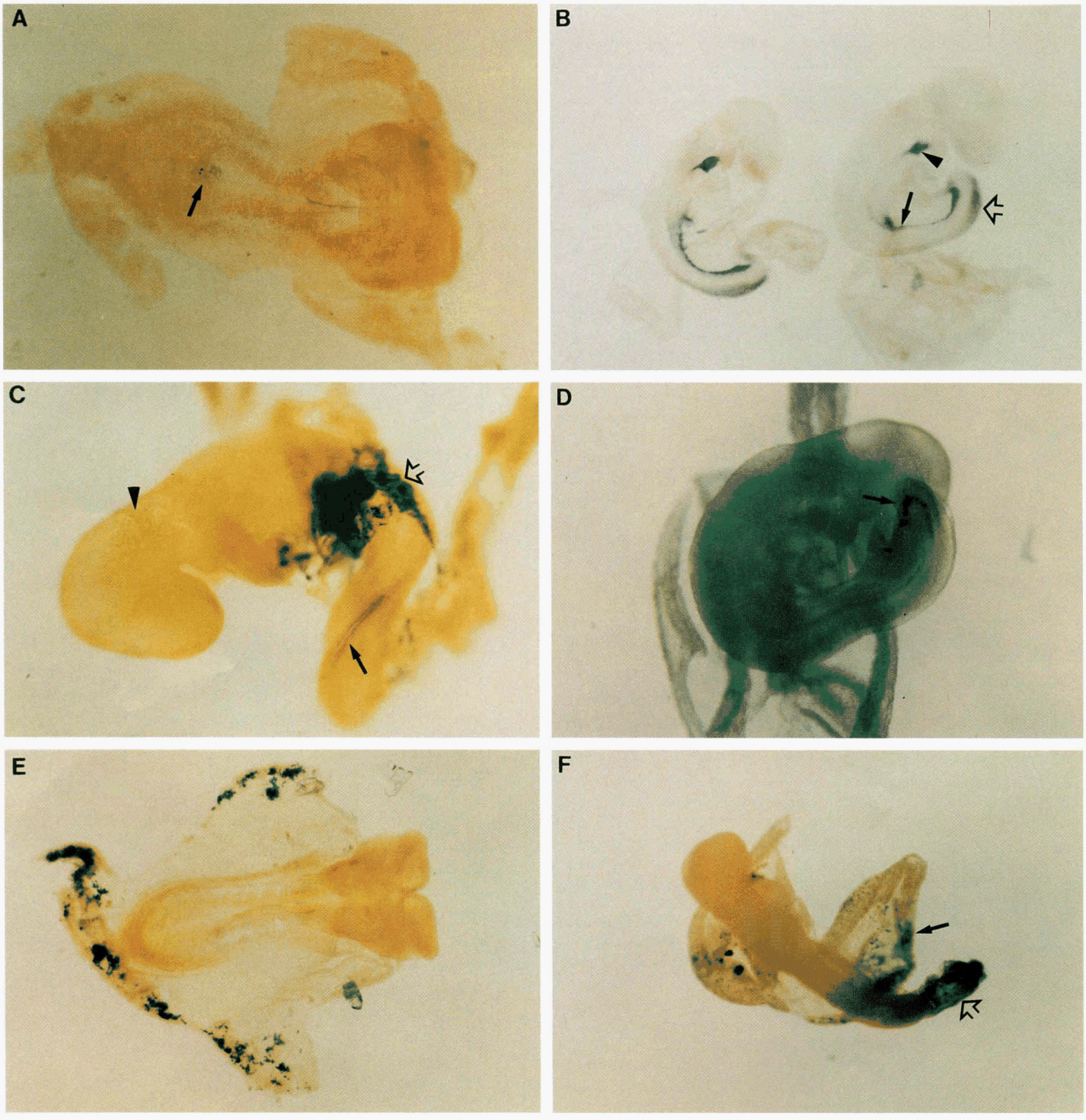

FIGURE 2.—ES-cell (8.5 dpc) chimeric embryos showing tissue specific expression. Reporter gene expression was detected in the embryo proper (A-D) or extraembryonic tissues (E and F). (A) An ~8-somite-stage chimeric embryo from ES clone PT1–19 showing node-specific (arrow) expression. (B) A late 8.5-dpc totally ES-cell-derived embryo from ES clone 6–15–2 showing gut (arrow), ventral pharynx (arrowhead) and posterior spinal cord (open arrow) expression. (C) A late 8.5-dpc chimeric embryo from clone 9–10 exhibiting longitudinal stripes of β-galactosidase-expressing cells along the posterior neural tube (arrow) and scattered cells in the head (arrowhead) and yolk sac (open arrow). (D) An ~9-dpc chimeric embryo from ES clone 14–50 showing specific expression in the dorsal hindgut (arrow). (E) An early 8.5-dpc chimeric embryo from ES clone PT1–14 showing expression in the mesoderm layer of the yolk sac probably in blood islands. (F) A 8.5-dpc embryo from ES clone 6–15–1 showing strong expression in the mesoderm layer of the yolk sac (arrow) and the allantois (open arrow).

spread or ubiquitous staining and 155 (55%) of the lines failed to express the reporter construct at detectable levels (Table 1).

After examining the chimeras from the 36 lines that showed spatially restricted patterns of expression in the 8.5 dpc embryos, we classified the expression patterns into four categories: highly restricted tissue-specific patterns (Figure 2), highly restricted region-specific patterns (Figure 3), highly restricted tissue-specific patterns plus widespread low-level expression (Figure 4), and highly restricted region-specific patterns plus widespread low-level expression (Figure 5). A description of the expression pattern for each clone is provided in Table 2.

Seventeen of the 36 clones showed tissue-specific expression (Table 2A). Nine clones showed tissue-specific expression in the embryo proper; of these lines six also showed weak expression in the extraembryonic yolk sac. Eight clones showed staining that was predominantly in extraembryonic tissues; two of these lines also showed weak expression in the embryo proper. The tissues that displayed β-galactosidase expression varied from clone to clone and included node and putative notochord (PT1–19, Figure 2A), gut (6–15–2,14–50; Figure 2, B and D), spinal chord (6–15–2, 9–10, 12–52; Figure 2, B and C), a few scattered cells in the head region (9–10, 12–27, 12–50, 14–49), yolk sac mesoderm (6–15–1, 13–11, 14–57) and distinct pockets of cells in the yolk sac mesoderm (PT1–14, 6–9-1).

Eleven of the 36 clones showed widespread expression with stronger expression in certain tissues. As in the first group, the patterns of expression ranged across a large variety of tissue types, including heart (PT1–13, 9–7,11–33, 13–17), somites (7–9-3, 9–7, 9–9; Figure 3), central nervous system (PT1–13, 9–7, 9–12, 12–2) and neural tube (7–9-3, 9–9).

Widespread low-level reporter-gene expression with stronger expression in the central nervous system was seen in 5 of the 36 clones. Three clones showed bands of differential expression across the hindbrain, indicative of expression in developing rhombomeres (9–4, 13–11, 5–8-1; Figure 4). One clone showed a specific anterior boundary of expression in the hindbrain (13–28). The fifth clone also showed spatially restricted bands of expression in the CNS, this time in longitudinal stripes along the neural tube (PT1–7).

Three of the 36 clones showed spatially restricted expression patterns along the anterior-posterior axis, without any obvious tissue-specificity. Two clones (6–16–3, 8–7-1) showed strong expression at the anterior and posterior ends of the embryo with a reduction of staining toward the middle of the embryo (Figure 5). The other clone showed a widespread low level of expression, with stronger staining in the anterior neural folds and the somites.

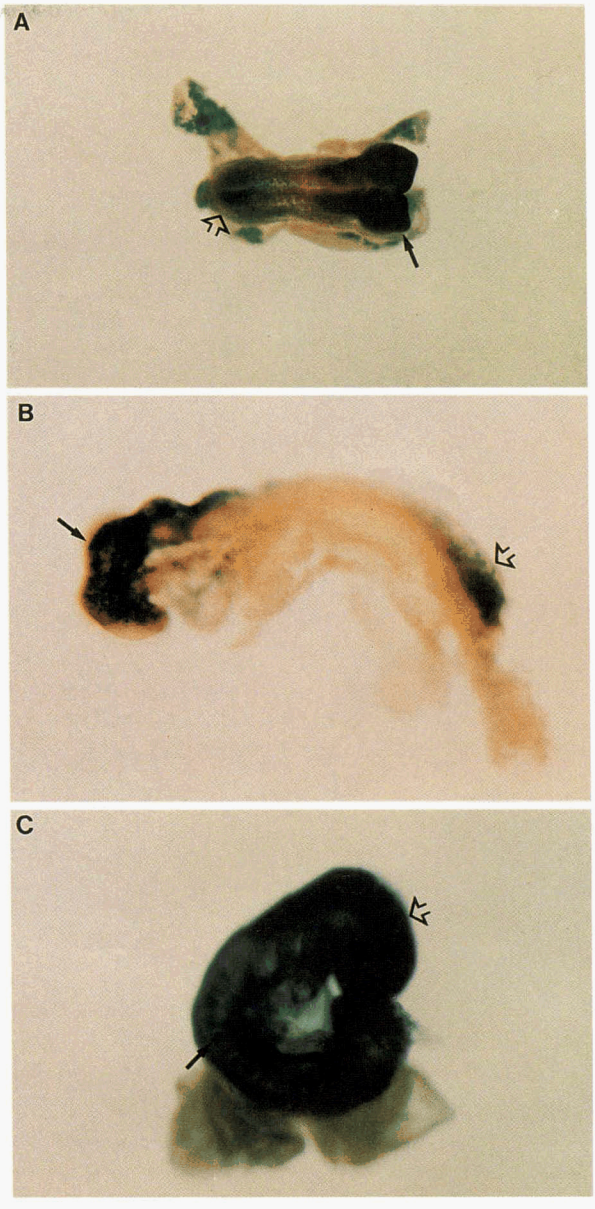

FIGURE 3.—ES-cell (8.5 dpc) chimeric embryos showing reporter gene expression in graded patterns along the anterior/posterior axis. ES cell chimeric embryos from lines 6–16–3 (A) and 8–7-1 (B) show strong reporter gene expression at the posterior (open arrow) and anterior (arrow) ends. The embryo in (B) represents a totally ES-cell-derived embryo. (C) An ES-cell chimeric embryo from line 9–3 showing stronger staining in the anterior neural (open arrow) folds and in the somites (arrow).

Reporter-gene expression in ES-tetraploid chimeras: To confirm the effectiveness of using chimeric embryos to predict patterns of reporter-gene expression, we generated a small number of aggregation chimeras using tetraploid host embryos. Such chimeras are almost entirely ES-cell-derived (NAGY *et al.* 1990; NAGY and ROSSANT 1993). Ten clones were chosen for these experiments of which four, 6–15–2, 8–7-1, 5–8-1 and

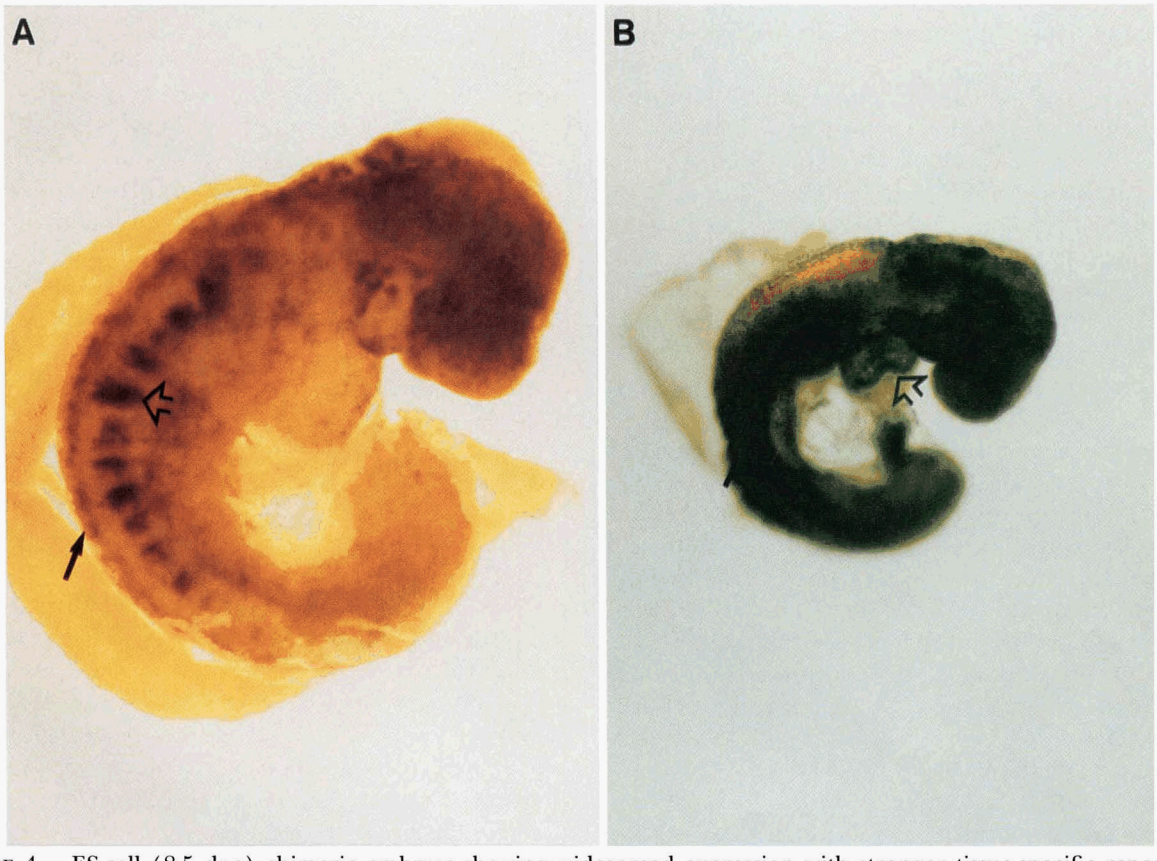

FIGURE 4.—ES-cell (8.5 dpc) chimeric embryos showing widespread expression with stronger tissue-specific reporter gene expression. (A) An ES cell chimeric embryo from line 7–9-3 showing widespread staining probably in the mesoderm with stronger staining in the somites (open arrow) and midline of the neural tube (arrow). (B) An ES-cell chimeric embryo from line 9–7 showing strongest staining in the somites (arrow), heart (open arrow) and hindbrain.

12–52, generated healthy chimeric embryos. These embryos were judged to be almost entirely ES-cell derived based on the lack of mosaicism in their β-galactosidase staining. These analyses confirmed the patterns of β-galactosidase gene expression that had been observed in the previously described diploid chimeras (Figures 2B, 3B and 4C).

In two lines, 6–15–2 and 5–8-1, chimeric embryos also were assayed at stages of development other than 8.5 dpc. At 8.5 dpc, 6–15–2 displayed reporter-gene expression in specific regions of the gut and in the posterior spinal chord. By 9.5 dpc, expression was found in these two areas as well as in a portion of the ventral pharynx. Line 5–8-1 displayed low levels of widespread expression coupled with strong staining in the midbrain and hindbrain of 8.5-dpc embryos and a lack of expression in two bands in the hindbrain, probably rhombomeres 3 and 5. When this clone was assayed in 7.5-dpc embryos, staining was seen to be restricted to the posterior region of the embryo.

Analysis of 12.5-dpc embryos: A large proportion (55%) of the β-galactosidase expressing clones failed to show reporter gene activity in 8.5-dpc embryos. To ensure that this was not because of the absence of ES cell contribution to the embryos, a subset of the lines were analyzed for their ability to generate chimeras at 12.5 dpc (Table 3). Because the D3 cell line carries the dominant agouti coat color marker, chimerism can be scored by the presence of eye pigmentation. We analyzed 139 lines for the presence of eye pigmentation at 12.5 dpc. From 19 of these lines we could not recover embryos at 12.5 dpc. Of the remaining 120 clones where embryos were recovered, 93 (78%) scored positive for eye pigmentation. This indicates that a substantial proportion of those clones showing no expression at 8.5 dpc were able to generate chimeras but were not displaying detectable reporter gene activity at 8.5 dpc.

We also assayed reporter-gene expression in 70 of the 139 lines that were tested for eye pigmentation (Table 4). When the patterns of reporter gene expression were compared between 8.5 dpc and 12.5 dpc, in 30 out of 70 lines the behavior of the reporter gene changed. Some of the lines (10 out of 26) that did not express the reporter gene at 8.5 dpc showed expression in 12.5-dpc embryos. Two of these had

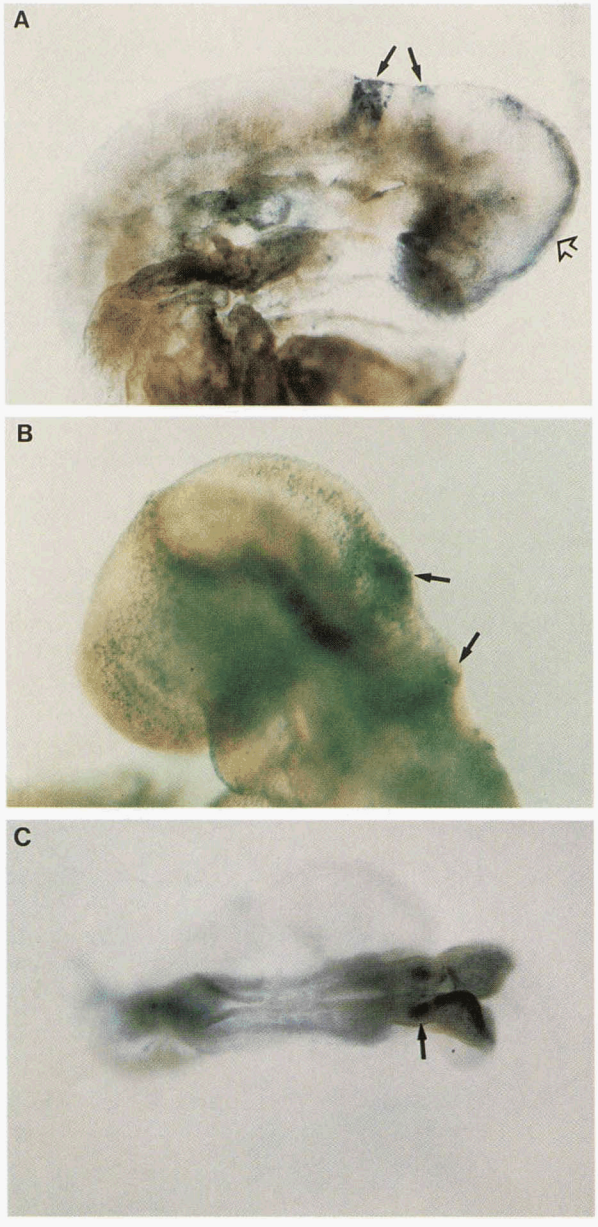

FIGURE 5.—ES-cell (8.5 dpc) chimeric embryos showing widespread low expression with stronger reporter gene expression in spatially defined domains. (A) An ES-cell chimeric embryo from line 13 to 31 showing stronger staining in the midline of the mid- and forebrain (open arrow) and strong staining in two bands in the hindbrain (arrow). (B) An ES cell chimeric embryo from line 9–4 showing stronger staining in two stripes across the hindbrain (arrow). (C) A totally ES-cell-derived chimeric embryo from ES clone 5–8-1 showing stronger staining in the hindbrain (arrow) and the midbrain.

restricted expression patterns. Five lines showed ubiquitous expression at 8.5 dpc but restricted expression at 12.5 dpc (Table 4). The restricted expression at 12.5 dpc was predominantly in the CNS (5–4-2, PT1–15 and 5–3-1) and in the developing limb buds (5–4-2, PT1–15 and 12–1).

DISCUSSION

We have generated >300 mouse ES cell clones containing potentially mutagenic integrations of a gene-trap vector that uses a β-galactosidase reporter gene to identify endogenous cellular promoters. When 279 of these clones were assayed for reporter-gene expression in 8.5-day chimeras, a range of developmental patterns was observed. Approximately one third of the clones showed widespread *lacZ* expression in 8.5-dpc embryos. Thirty-five of the 279 clones exhibited tissue-specific or spatially restricted expression patterns at 8.5 dpc. The remainder of the clones did not express *lacZ* at this stage of development, despite expression in ES cells. We have shown that the lack of expression in the latter group of embryos was not due to inability of the ES cells to contribute to chimeras, because for most of these clones, the presence of ES-derived cells could be scored by the presence of eye pigmentation when the embryos were allowed to develop to 12.5 days. In fact, we estimate that only ~10% of the ES cell clones that underwent the screening procedure were not able to contribute to chimeric embryos. When a subset of these negative clones were analyzed at midgestation (12.5 dpc), more than one third were found to express *lacZ*. Furthermore, 5/33 clones that were ubiquitously expressed at 8.5 dpc showed evidence of patterned expression at 12.5 dpc. The results of this gene-trap screen demonstrate that it is possible to identify a wide assortment of genes showing tissue-specific and spatially restricted expression during development even when the analysis is limited to genes expressed in ES cells and during one developmental stage. Furthermore, our limited analysis at 12.5 dpc indicates that if the time window of expression analysis could be broadened, then a much larger number of genes with restricted expression patterns during development would be identified.

As mentioned earlier, the expression patterns of the 36 clones that showed spatial modulation of the reporter gene at 8.5 dpc could be divided into four basic categories: highly restricted tissue-specific patterns (Figure 2), highly restricted region-specific patterns (Figure 3), highly restricted tissue-specific patterns plus widespread low-level expression (Figure 4) and highly restricted region-specific patterns plus widespread low-level expression (Figure 5).

Because this screen was limited to genes that are expressed in ES cells, the question of whether the frequency of these classes of gene expression patterns reflects the frequency of the types of genes that are developmentally regulated in the embryo remains unresolved. It is clear that some developmentally restricted genes such as *En-2* and *Hox 1.3* are expressed in undifferentiated ES cells, whereas others such as *wnt-1* and *En-1* are not (JOYNER *et al.* 1985; JOYNER and MARTIN

TABLE 2

Gene trap β-galactosidase expression patterns obtained in chimeric embryos at 8.5 dpc

Clone	Expression type	Description[a]
PT1-19	Highly restricted tissue-specific	Embryonic, node-specific, and a line of cells anterior to the node in a midline position at early somite stages
6-13-1	Highly restricted tissue-specific	Embryonic/extraembryonic, a few scattered cells in headfolds/yolk sac and allantois
6-15-2	Highly restricted tissue-specific	Embryonic/extraembryonic, specific staining in mid- and hindgut, posterior spinal cord and a portion of the ventral pharynx/very weak staining in yolk sac
9-10	Highly restricted tissue-specific	Embryonic/extraembryonic, specific stripe along the dorsal posterior neural tube, scattered cells in the head/strong staining in the yolk sac
12-27	Highly restricted tissue-specific	Embryonic/extraembryonic, strong staining in groups of cells lateral to the hindbrain (possibly neural crest), scattered cells staining at posterior and along neural tube/weak staining in yolk sac
12-50	Highly restricted tissue-specific	Embryonic, bilateral stream of cells between metencephalon and otic vessicle (possibly neural crest or paraxial mesoderm derived cells)
12-52	Highly restricted tissue-specific	Embryonic/extraembryonic, strong in the posterior spinal chord and scattered cells in heart and head/yolk sac
14-49	Highly restricted tissue-specific	Embryonic, expression in scattered cells in the heart, around the otic vesicle and in the brachial arches (possibly neural crest)
14-50	Highly restricted tissue-specific	Embryonic/extraembryonic, specific dorsal hindgut staining/allantois and yolk sac
PT1-1	Highly restricted tissue-specific	Extraembryonic, few cells staining in yolk sac and allantois
13-76	Highly restricted tissue-specific	Embryonic/extraembryonic, weak general staining in embryo/stronger yolk sac and allantois
PT1-14	Highly restricted tissue-specific	Extraembryonic, yolk sac mesoderm, probably in blood islands
6-15-1	Highly restricted tissue-specific	Extraembryonic, yolk sac mesoderm and strong staining in allantois
7-5-2	Highly restricted tissue-specific	Extraembryonic, yolk sac only
13-11	Highly restricted tissue-specific	Extraembryonic, staining in yolk sac mesoderm at the base of the allantois
14-57	Highly restricted tissue-specific	Extraembryonic, yolk sac mesoderm
6-9-1	Highly restricted tissue-specific	Embryonic/extraembryonic, few scattered cells in embryo/groups of cells in yolk sac and allantois
6-16-3	Highly restricted region-specific	Embryonic/extraembryonic, stronger expression at anterior and posterior ends of the embryo with graded reduction towards the midline/strong staining in yolk sac endoderm
8-7-1	Highly restricted region-specific	Embryonic/extraembryonic, stronger staining in the dorsal posterior and anterior regions of the embryo with graded reduction towards the middle/weak staining in yolk sac and allantois
9-3	Highly restricted region-specific	Embryonic/extraembryonic, widespread low-level expression with stronger staining in anterior neural folds and the somites/weak staining in allantois
PT1-13	Tissue-specific plus widespread low-level	Embryonic/extraembryonic, stronger staining in heart, skin, and along the midline of the spinal cord and hindbrain, scattered cells in mandible/yolk sac
7-9-3	Tissue-specific plus widespread low-level	Embryonic/extraembryonic, at early stages widespread staining probably in the mesoderm and later strongest staining in the somites and midline of the neural tube/weak staining in yolk sac
9-9	Tissue-specific plus widespread low-level	Embryonic/extraembryonic, strongest staining in the somites and dorsal neural tube/yolk sac
9-12	Tissue-specific plus widespread low-level	Embryonic, widespread but stronger staining in the ventral CNS
11-33	Tissue-specific plus widespread low-level	Embryonic/extraembryonic, strongest staining in the heart/strong staining in the yolk sac
12-2	Tissue-specific plus widespread low-level	Embryonic/extraembryonic, stronger in posterior end of the embryo and a strong band of staining in the hindbrain/weak yolk sac staining
13-15	Tissue-specific plus widespread low-level	Embryonic/extraembryonic, widespread in embryo but absent from the neural tube (may be mesoderm specific)/yolk sac
13-17	Tissue-specific plus widespread low-level	Embryonic, stronger anteriorly in head and heart with weak widespread staining
13-48	Tissue-specific plus widespread low-level	Embryonic, strongest in the head (mainly CNS)
14-59	Tissue-specific plus widespread	Embryonic/extraembryonic, strongest in the anterior end of the embryo
PT1-7	Region-specific patterns plus widespread low-level	Embryonic/extraembryonic, weak widespread staining with stronger staining in four longitudinal anterior to posterior stripes along the neural tube

TABLE 2

Continued

Clone	Expression type	Description[a]
5-8-1	Region-specific patterns plus widespread low-level	Embryonic, weak widespread expression with strongest staining in mid- and hindbrain
9-4	Region-specific patterns plus widespread low-level	Embryonic, widespread early weak staining but later stages show stronger staining in two stripes across the hindbrain
13-28	Region-specific patterns plus widespread low-level	Embryonic/extraembryonic, weak widespread staining including allantois with stronger staining in spinal cord and hindbrain with anterior boundary in the hindbrain and a stronger band of expression at this boundary
13-31	Region-specific patterns plus widespread low-level	Embryonic/extraembryonic, broad weak staining with strong specific staining in a band of cells in the hindbrain (at the level of the otic vesicle) and a weaker band more anteriorly, plus strong staining in the midline of the mid- and forebrain, heart/yolk sac

[a] Clones PT1-19, 6-15-2, 9-10, 8-7-1, 9-3, 7-9-3, 9-7, 14-50, PT1-14, 6-15-1, 6-16-3, 5-8-1, 9-4 and 13-31 are shown in Figures 2, 3, 4 or 5. Clones 6-15-2, 12-52, 5-8-1 and 8-7-1 were tested by aggregation with tetraploid embryos. Clones 13-15, PT1-13 and 13-31 express the reporter gene only in differentiated cells.

1987; McMAHON and BRADLEY 1990; JEANNOTTE *et al.* 1991). A more laborious but less biased screen would involve the analysis of all clones that had incorporated the vector into their genome. A large-scale screen of this type would be very difficult because many integrations would fall outside of transcription units and, due to the nature of the vector, even those that were within genes could only be expressed if integrated in the correct orientation and reading frame.

Many of the reporter-gene expression patterns that we observe are consistant with the behavior of ES cells in culture. Spontaneous differentiation of the D3 ES cell line has been shown to give rise to a number of identifiable cell types, including yolk sac-like structures that contain blood islands and primitive blood vessels as well as cardiac muscle cells (DOETSCHMAN *et al.* 1985). Of the 36 clones that displayed restricted expression, 24 clones showed expression in the yolk sac (Table 2); in two of the lines expression may be restricted to blood islands (PT1-14 and 6-9-1). Expression in the developing heart was observed in six clones. Because in our prescreen colonies were allowed to partially differentiate, we may have a bias toward genes that are expressed in tissues that are generated by *in vitro* differentiation. The two clones that were expressed exclusively in differentiated cells (13-31 and PT1-13) and displayed regulated expression of the reporter gene showed expression in both the yolk sac and the heart.

CNS expression was predominant in 18 out of 36 lines. This result is also not surprising because at 8.5 dpc the CNS is undergoing active organization with respect to dorsal/ventral and anterior/posterior polarity as well as with respect to segmental identity. A large number of genes responsible for these events might be expected to be active in the early embryo and ES cells and thus be identified in our screen.

This study shows that it is feasible to prescreen 300 gene-trap integrations in ES cells by expression in chimeras. Those with the most interesting developmental expression patterns then can be selected for further analysis at the molecular level and can be transmitted through the germ line for phenotypic analysis of the mutation. This prescreen allows large-scale insertional mutagenesis in mice without access to large animal colony resources. In this screen, we have limited analysis to one window of development: 8.5 dpc. This time was chosen as it is the time when the basic body plan of the embryo is being established. However, other time windows could be added, recognizing that every additional time point analyzed doubles the number of animals used for chimera analysis. In a screen of ~300 lines, it has proven possible to generate a diversity of

TABLE 3

Chimeric embryos analyzed at 12.5-dpc

Lac Z expression at 8.5 dpc	Total no. of lines assayed	No. of lines with eye pigment at 12.5 dpc	No. of lines without eye pigment at 12.5 dpc	No. embryos at 12.5 dpc
+	70	53	10	7
−	69	40	17	12
Total	139	93	27	19

TABLE 4

***Lac Z* expression in 12.5-dpc embryos**

Lac Z expression at 8.5 dpc	Total no. of embryos assayed at 12.5 dpc	Widespread expression at 12.5 dpc	Restricted expression at 12.5 dpc	No expression at 12.5 dpc
Widespread	28	21	5	2
Embryonic patterns	14	8	3	3
Extraembryonic patterns	2	2	0	0
No expression	26	8	0	0

clones whose pattern of reporter gene expression is regulated around the time point of interest. The choice of clones to follow in further detail varies with the particular interests of each individual investigator.

The effort required for this screen of 300 clones makes it difficult to envision achieving saturation mutagenesis for all genes expressed in ES cells. One factor that could influence the ability to achieve saturation mutagenesis with this type of approach is insertion of the vector itself. Although it has been reported that transfection via electroporation is a random process, SUTHERLAND and coworkers found that 2/46 integration events with a targetting vector integrated into the same locus (TONEGUZZO *et al.* 1988; RIED and SMITHIES 1992; SUTHERLAND *et al.* 1993). Although insertion of vector DNA is probably not an entirely random process, because we did not see any reiteration of patterns generated in this study, our vector is certainly not limited to insertion into only a few sites in the genome. Given the low frequency of identifying genes expressed in a particular tissue or spatial domain, for example 4/279 for hindbrain stripes, it also makes it difficult to envision using this type of global screen to target specific developmental processes.

A number of vector modifications are also possible that would allow for both improved identification and induction of mutations. Improved efficiency of isolating ES clones with gene-trap integrations in active genes has been achieved by development of the *βGeo* gene-trap vectors, in which all *neo*^{r} colonies should express β-galactosidase activity (FRIEDERICH and SORIANO 1991). The addition of a translation initiation sequence to the β-galactosidase gene makes it independent of the endogenous protein coding sequence and increases the frequency of expressing clones by at least threefold (unpublished results). Replacement of the β-galactosidase gene by a gene that can be assayed *in vivo* would eliminate the replica plating step of the screen. One candidate reporter gene is the green fluorescent protein recently used as a reporter gene in *Caenorhabditis elegans* (CHALFIE *et al.* 1994). Another useful addition to the gene-trap vector would be the insertion of sequences that respond to site-specific recombinases. Addition of these sequences would allow manipulation of the locus where insertion occurred to create genetic mosaics and revertants or place other genes under the control of the endogenous promoter. This approach already has been used to generate a tissue-specific mutation in the mouse (GU *et al.* 1994).

With these type of improvements in the gene-trap vector, the rate-limiting step of the screen becomes the production and screening of chimeric embryos. Improved ES cell lines such as R1 and improved techniques of generating chimeric embryos such as aggregation between ES cells and single embryos can reduce the number of embryos and the effort needed to generate chimeras (NAGY *et al.* 1994). A further reduction the number of chimeric embryos required for a screen could be achieved by prescreening ES clones for genes that have a higher possibility of being involved in processes of interest. For example subcellular localization of the β-galactosidase activity could be used as a criterion. Another potential prescreen would take advantage of the ability of ES cells to differentiate spontaneously (DOETCHMAN *et al.* 1985), or in response to growth and differentiation factors (HEATH and SMITH 1988; HILL and WURST 1993b). In addition, molecular characterization of the tagged genes by RACE protocols could provide further information on which to base the choice of lines to be studied (VON MELCHNER *et al.* 1990). These types of approaches, coupled with a better understanding of the behavior of ES cells and the design of better trapping vectors, should make mutagenic screens in the mouse feasible in a large number of laboratories.

The authors thank Dr. ERIC EVERETT, Dr. KENT ROBERTSON, Dr. SUSAN STROME and MONICA MCANDREWS-HILL for their critical reading of the manuscript. This work was supported by National Institutes of Health grant HD-25334 to J.R. and A.J. W.W. held a Fellowship from the Deutsche Forschungsgemeinschaft, J.R. and A.J. are International Scholars of the Howard Hughes Medical Institute, J.R. is a Terry Fox Cancer Research Scientist of the NCIC, A.J. is an MRC of Canada Scientist, D.P.H. and F.G. held MRC of Canada Fellowships.

LITERATURE CITED

BELLEN, H. J., C. J. O'KANE, C. WILSON, U. GROSSNIKLAUS, R. K. PEARSON *et al.*, 1989 P-element-mediated enhancer detection: a versitile method to study development in Drosophila. Genes Dev. **3:** 1288–1300.

BIER, E., H. VAESSIN, K. LEE, K. MCCALL, S. BARBELL *et al.*, 1989 Searching for pattern and mutation in the Drosophila genome with a P-lacZ vector. Genes Dev. **3:** 1273–1287.

BOER, P. H., H. POTTEN, C. N. ADRA, K. JARDINE, G. MULLHOFER *et al.*, 1990 Polymorphisms in the coding and noncoding regions of murine PGK-1 alleles. Biochem. Genet. **28:** 299–308.

CAPECCHI, M. R., 1989 Altering the genome by homologous recombination. Science **244:** 1288–1292.

CHALFIE, M., Y. TU, G. EUSKIRCHEN, W. WARD and D. C. PRASHER, 1994 Green fluorescent protein as a marker for gene expression. Science **263:** 802–804.

CHAMBON, P., 1993 Local alterations of Krox-20 and Hox gene expression in the hindbrain suggest lack of rhombomeres 4 and 6 in homozygote null Hoxa-1 (Hox 1.6) mutant embryos. Proc. Natl. Acad. Sci. USA **90:** 7666–7670.

DOETSCHMAN, T. C., H. EISTETTER, M. KATZ, W. SCHMIDT and R. KEMLER, 1985 The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood glands and myocardium. J. Embryol. Exp. Morphol. **87:** 27–45.

DOLLE, P., T. LUFKIN, R. KRUMLAUF, M. MARK, D. DUBOULE *et al.*, 1992 Developmental defects of the ear, cranial nerves and hindbrain resulting from a targeted mutation disruption of the mouse homeobox gene Hox 1.6. Nature **355:** 516–520.

FRIEDRICH, G., and P. SORIANO, 1991 Gene traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev. **5:** 1513–1523.

GOSSLER, A., and J. ZACHGO, 1993 Gene and enhancer trap screens in ES cell chimeras. pp. 181–227 in *Gene Targeting: A Practical Approach*, edited by A. L. JOYNER. Oxford University Press, Oxford.

GOSSLER, A., A. L. JOYNER, J. ROSSANT and W. C. SKARNES, 1989 Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science **244:** 463–465.

GU, H., J. D. MARTH, P. C. ORBAN, H. MOSSMAN, and K. RAJEWSKY, 1994 Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. Science **256:** 103–105.

HEATH, J. K., and A. G. SMITH, 1988 Regulatory factors of embryonic stem cells. J. Cell Sci. **10** (Suppl.): 257–266.

HILL, D. P., and W. WURST, 1993a Gene and enhancer trapping: mutagenic strategies for developmental studies. pp.181–206 in *Current Topics in Developmental Biology, volume 28*, edited by Roger A. PEDERSEN. Academic Press, Orlando.

HILL D. P., and W. WURST, 1993b Screening for novel pattern formation genes using gene trap approaches. Methods Enzymol. **225:** 664–681.

HOPE, I., 1991 "Promoter Trapping" in *Caenorhabditis elegans*. Development **113:** 399–408.

JEANNOTTE, L., J. C. RUIZ, and E. J. ROBERTSON, 1991 Low level of Hox1.3 gene expression does not preclude the use of promotorless vectors to generate a targeted gene disruption. Mol. Cell. Biol. **11:** 5578–5585.

JOYNER, A. L., and G. R. MARTIN, 1987 En-1 and En-2, two mouse genes with sequence homology to the Drosophila engrailed gene: expression during embryogenesis. Genes Dev. **1:** 29–38.

JOYNER, A. L., T. KORNBERG, K. G. COLEMAN, D. R. COX and G. R. MARTIN, 1985 Expression during embryogenesis of a mouse gene with sequence homology to the Drosophila engrailed gene. Cell **43:** 29–37.

KAUR, S., G. SINGH, J. L. STOCK, C. M. SCHREINER, A. KIER *et al.*, 1992 Dominant mutation of the murine Hox 2.2 gene results in developmental abnormalities. J. Exp. Zool. **264:** 323–336.

KOLLER, B. H., and O. SMITHIES, 1992 Altering genes in animals by gene targeting. Annu. Rev. Immunol. **10:** 705–730.

KORN, R., M. SCHOOR, H. NEUHAUS, U. HANSELING, R. SOININEN *et al.*, 1992 Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimeric embryos with a high frequency and detect endogenous genes. Mech. Dev. **39:** 95–109.

KRUMLAUF, R., 1993 Hox genes and pattern formation in the brachial region of the vertebrate head. Trends Genet. **9:** 106–112.

LE MOUELLIC, H., Y. LALLEMAND and P. BRULET, 1992 Homeosis in the mouse induced by a null mutation in the Hox 3.1 gene. Cell **69:** 261–264.

LUFKIN, T., A. DIERICH, M. LEMEUR, M. MARK and P. CHAMBON, 1991 Disruption of the Hox 1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. Cell **66:** 1105–1119.

MCLAIN, K., C. SCHREINER, K. L. YAGER, J. L. STOCK and S. S. POTTER, 1992 Ectopic expression of Hox 2.3 induces craniofacial and skeletal malformations in transgenic mice. Mech. Dev. **39:** 3–16.

MCMAHON, A. P., and A. BRADLEY, 1990 The Wnt-1 (Int-1) proto-oncogene is required for development of a large region of the mouse brain. Cell **62:** 1073–1085.

NAGY, A., and J. ROSSANT, 1993 Production of completely ES cell-derived fetuses, pp. 147–179 in *Gene Targeting: A Practical Approach*, edited by A. L. JOYNER. Oxford University Press, Oxford.

NAGY, A., E. GOCZA, E. M. DIAZ, V. R. PRIDEAUX, E. IVANYI *et al.*, 1990 Embryonic stem cells alone are able to support fetal development in the mouse. Development **110:** 815–822.

NAGY, A., J. ROSSANT, R. NAGY, W. ABRAMOW-NEWERLY and J. C. RODER, 1994 Derivation of completely cell culture derived mice from early-passage embryonic stem cells. Proc. Natl. Acad. Sci. USA **90:** 8424–8428.

REID, L., and O. SMITHIES, 1992 Gene targetting and electroporation, pp. 209–225 in *Guide to electroporation and electrofusion*, edited by D. C. CHANG, B. M. CHASSEY, J. A. SAUNDERS and A. E. SOWERS. Academic Press, Orlando.

RISAU, W., H. SARIOLA, H. G. ZERWES, J. SASSE, P. EKBLOM *et al.*, 1988 Vasculogenesis and angiogenesis in embryonic-stem-cell-derived embryoid bodies. Development **102:** 471–478.

SKARNES, W. C., B. A. AUERBACH and A. L. JOYNER, 1992 A gene trap approach in mouse embryonic stem cells: the lacZ reporter is activated by splicing, reflects endogenous gene expression and is mutagenic in mice. Genes Dev. **6:** 903–918.

ST. JOHNSTON, D., and C. NUSSLEIN-VOLHARD, 1992 The origin of pattern and polarity in the Drosophila embryo. Cell **68:** 201–219.

SUTHERLAND, H., R. LOVELL-BADGE and I. JACKSON, 1993 Characterisation of two identical independent non-homologous integration sites in mouse embryonic stem cells. Gene **131:** 265–268.

THISSE, C., and B. THISSE, 1992 Dorsoventral development of the Drosophila embryo is controlled by a cascade of transcriptional regulators. Development (Suppl.) 173–181.

TONEGUZZO, F., A. KEATING, S. GLYNN and K. MCDONALD, 1988 Electric field-mediated gene transfer: characterization of DNA transfer and patterns of integration in lymphoid cells. Nucleic Acids Res. **16:** 5515–5532.

VON MELCHNER, H., S. REDDY and H. E. RULEY, 1990 Isolation of cellular promoters by using a retrovirus promoter trap. Proc. Natl. Acad. Sci. USA **87:** 3733–3737.

WILSON, C., R. K. PEARSON, H. J. BELLEN, C. J. O'KANE, U. GROSSNIKLAUS *et al.*, 1989 P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. Genes Dev. **3:** 1301–1313.

WURST, W., and A. L. JOYNER, 1993 Production of targeted embryonic stem cell clones, pp. 33–61 in *Gene Targeting: A Practical Approach*, edited by A. L. JOYNER. Oxford University Press, Oxford.

Communicating editor: R. E. GANSCHOW