Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 1677–1682, February 1996
Developmental Biology

An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid *in vitro*

(induction trapping/genetic screen)

LESLEY M. FORRESTER*†, ANDRAS NAGY*, MEHRAN SAM*‡, ALISTAIR WATT†, LOIS STEVENSON*, ALAN BERNSTEIN*‡, ALEXANDRA L. JOYNER*‡§, AND WOLFGANG WURST*¶‖

*Division of Molecular and Developmental Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, ON Canada, M5G 1X5; ‡Department of Molecular and Medical Genetics, University of Toronto, Toronto, ON Canada, M5S 1A8; §Developmental Genetics Program, Skirball Institute of Biomolecular Medicine, New York University Medical Center, 540 First Avenue, New York, NY 10016; ¶Institut für Säugetiergenetik, GSF-Forschungszentrum für Umwelt und Gesundheit, Ingolstädter Landstrasse 1, 85764 Oberschleissheim, Germany; and †Centre for Genome Research, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh, EH9 3JQ, United Kingdom

Communicated by Walter J. Gehring, University of Basel, Basel, Switzerland, October 10, 1995 (received for review June 16, 1995)

ABSTRACT **We have developed a novel induction gene trap approach that preselects *in vitro* for integrations into genes that lie downstream of receptor/ligand-mediated signaling pathways. Using this approach, we have identified 20 gene trap integrations in embryonic stem cells, 9 of which were induced and 11 of which were repressed after exposure to exogenous retinoic acid (RA). All but one of these integrations showed unique spatially restricted or tissue-specific patterns of expression between 8.5 and 11.5 days of embryogenesis. Interestingly, expression was observed in tissues that are affected by alterations in RA levels during embryogenesis. Sequence analysis of fusion transcripts from six integrations revealed five novel gene sequences and the previously identified protooncogene c-*fyn*. To date, germ-line transmission and breeding has uncovered one homozygous embryonic lethal and three homozygous viable insertions. These studies demonstrate the potential of this induction gene trap approach for identifying and mutating genes downstream of signal transduction pathways.**

Embryogenesis requires a regulated and sequential activation of genes whose protein products contribute to numerous signaling pathways. It has been possible to dissect many key signaling pathways in model invertebrates since they are readily amenable to genetic screens. In mammals, where large genetic screens are not feasible, one approach to dissect developmental pathways has been to use cross species homology. However, in order to use an unbiased screening approach, novel strategies must be applied.

We set out to determine whether it would be possible to adapt gene trap methodology in embryonic stem (ES) cells (1, 2) to identify, mutate, and analyze novel genes involved in signaling pathways in mammals. When gene trap constructs integrate into an intron of an active gene, a fusion transcript is generated containing upstream exons of the trapped gene and the *lacZ* sequence (3). There are three important consequences of such insertion events: (*i*) *lacZ* expression is regulated by the trapped gene; (*ii*) the trapped gene can be cloned using rapid amplification of cDNA ends (RACE)–PCR-based strategies; (*iii*) the insertional event has the potential to be mutagenic and so the biological consequences can be analyzed after germ-line transmission.

Gene trap approaches have been successfully used to trap genes expressed in undifferentiated ES cells (3–6). In such random screens, only 15–30% of the integrations display restricted patterns of expression during embryogenesis (5, 6). We used a retinoic acid (RA) prescreen to identify genes that respond to RA and to test whether this enriches for genes involved in embryonic development. The application of exogenous RA is known to induce profound effects on patterning of various tissues during embryogenesis (7, 8). Furthermore, RA has been shown to alter the expression of developmentally regulated pattern formation genes both in embryonal carcinoma cells in culture (9–11) and in embryos (12, 13).

We have identified integrations in 20 genes that were responsive to RA treatment. All but one of these trapped genes displayed unique expression patterns during embryogenesis, indicating that this screen greatly enriches for integrations into genes that are expressed in a spatially restricted or tissue-specific manner.

MATERIALS AND METHODS

Gene Trap Vectors. The gene trap vector PT1-ATG (14) contains the *En-2* splice acceptor site fused to the *lacZ* reporter gene with an ATG translational start site. Integration of PT1-ATG into the intron of an active gene can generate a fusion transcript between *lacZ* and an endogenous trapped gene. The bacterial neomycin-resistance gene is driven by the phosphoglycerate-1 (*PGK-1*) promoter.

ES Cell Culture and Selection of RA-Responsive Cell Lines. R1 (15) ES cells were maintained on primary embryonic fibroblasts (16). After electroporation and selection in G418 for 8 days, drug-resistant colonies were replica-plated (14) and the filters were placed in ES cell medium (16) (−LIF) containing 5% fetal calf serum (FCS) and 10^{-6} M all-*trans*-RA (Sigma) (RA-induction medium). After 42 hr, fresh RA-induction medium was added and the filters were left for an additional 6 hr before assaying for β-galactosidase (β-gal) activity (4). Blue colonies were picked from the master plate and retested in 24-well plates to differentiate between RA-responsive and constitutively expressed gene trap integrations.

Quantitative β-Gal Assay. All β-gal induction experiments were performed on ES cells grown in 35-mm plates in ES cell medium (−LIF) containing 5% FCS. RA was added at various times over a 48-hr period such that all induction time points (6, 12, 24, or 48 hr) were harvested at the same time. Control samples (0 hr) were grown in medium containing 5% FCS (−LIF, −RA) throughout the experiment. All cultures at each time point were done in duplicate (Table 1).

Cells were harvested by trypsin treatment and resuspended in 100 μl of 0.25 M Tris·HCl (pH 7.5). After three freeze/thaw cycles, the samples were centrifuged (13,000 rpm for 5 min; 15,110 × *g*), β-gal assays were performed with equivalent

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ES cell, embryonic stem cell; RACE, rapid amplification of cDNA ends; RA, retinoic acid; β-gal, β-galactosidase; LIF, leukemia inhibitory factor; dpc, days postcoitus; A/P, anterior/posterior.

‖To whom reprint requests should be addressed.

Table 1. Quantitative β-gal assays on ES cell lines after RA exposure

Cell line	RA treatment, hr					-fold induction
	0	6	12	24	48	
I.9	7 ± 1*	5 ± 0	5 ± 0	4 ± 1	87 ± 4†	12.4
I.23	20 ± 2	24 ± 3	11 ± 3	20 ± 12	29 ± 8†	1.5
I.75	8 ± 1	9 ± 1	9 ± 0	10 ± 1	14 ± 2†	1.8
I.114	20 ± 3	22 ± 0	17 ± 3	15 ± 1	51 ± 10†	2.6
I.134	2 ± 1	4 ± 1	5 ± 0	5 ± 1	46 ± 8†	20.0
I.163	10 ± 1	9 ± 0	9 ± 0	12 ± 1	45 ± 7†	4.0
I.193	13 ± 1	11 ± 1	11 ± 1	11 ± 1	112 ± 46	8.6
I.210	4 ± 1	2 ± 0	2 ± 0	5 ± 1	44 ± 18†	12.0
I.214	5 ± 1	5 ± 0	4 ± 0	17 ± 2	87 ± 4	16.4
R.24	99 ± 19	90 ± 4	92 ± 3	75 ± 8	58 ± 5	0.58
R.68	139 ± 1	132 ± 1	132 ± 5	97 ± 1	56 ± 2	0.40
R.77	41 ± 1	44 ± 1	44 ± 2	36 ± 2	15 ± 0	0.37
R.108	28 ± 0	25 ± 0	27 ± 0	26 ± 1	21 ± 2†	0.75
R.121	74 ± 3	69 ± 1	66 ± 4	57 ± 3	37 ± 4†	0.50
R.124	69 ± 10	87 ± 4	81 ± 4	44 ± 6	25 ± 5	0.36
R.140	58 ± 4	50 ± 1	49 ± 1	46 ± 1	46 ± 1	0.78
R.170	28 ± 1	27 ± 5	29 ± 2	19 ± 1	12 ± 0	0.44
R.179	36 ± 5	35 ± 1	38 ± 3	28 ± 3	16 ± 2	0.46
R.194	19 ± 1	27 ± 2	18 ± 4	11 ± 3	7 ± 1	0.37
R.213	31 ± 0	28 ± 4	25 ± 2	12 ± 0	8 ± 0	0.25

*OD_{420} ($\times 10^2$)—most values are given as mean of duplicate samples ± range.
†Mean of four independent values ± SE.

amounts of protein (0.2 mg) as described (17) and the optical density at 420 nm (OD_{420}) determined (Table 1).

Embryonic Expression Pattern Analysis and Germ-Line Transmission. ES cells were aggregated with tetraploid embryos (18), harvested at 8.5–11.5 days postcoitus (dpc), and stained for β-gal expression (4). In all cases, more than one completely ES-derived embryo was analyzed between 8.5 and 11.5 dpc to confirm the recorded expression pattern. In addition, diploid chimeras were produced for line I.75, I.210, and R.121. Some of the integrations were transmitted through the germ line by injection of the ES cells into C57BL6 blastocysts or by aggregation with CD1 eight-cell-stage host embryos. Chimeric males were bred to CD1 and/or 129/Sv females, and DNA of F_1 offspring was analyzed by Southern blotting. Heterozygous and homozygous offspring of F_1 intercrosses were distinguished by the intensity of the *lacZ* cross-hybridizing fragments compared to *En-2* as an internal standard.

Cloning and DNA and RNA Analysis. DNA and RNA preparations and Southern and Northern blots were performed according to standard procedures (19). Cloning of fusion transcripts was performed with a 5′ RACE kit (GIBCO/BRL) according to the manufacturer's instructions with 3′ oligonucleotide primers as described (3). RACE–PCR products were cloned into pAMP1 (GIBCO/BRL) or Bluescript vectors and sequenced (19). Sequences obtained were compared to the GenBank/EMBL data base.

RESULTS

Identification of RA-Responsive Integrations. R1 ES cells (6.5×10^7) were electroporated with the PT1-ATG gene trap vector (14) and 3600 G418-resistant ES cell colonies were replica plated. The replicated colonies were induced with RA for 48 hr and then stained for β-gal activity; 202 β-gal-positive colonies were picked from the master plate and retested for their response to RA. Microscopic analyses of stained colonies revealed nine cells lines in which β-gal activity was induced by RA (Table 2). Surprisingly, we also identified 13 ES cell clones in which the reporter gene was apparently repressed by RA (Table 2). Repression of β-gal was not complete and a low level of activity was detectable in some cells of these colonies, which explains why they were identified as positive in the original screen. Clearly, this strategy would not detect integrations into genes that were completely repressed after 48 hr of RA treatment.

Quantitation of the β-Gal Activity. The inductive response was confirmed for all RA-induced lines using a quantitative β-gal assay (Table 1). However, two of the microscopically identified RA-repressed lines did not show repression by the quantitative assay and were excluded from the analysis. All lines that were subsequently analyzed showed clear RA responsiveness on three independent tests when observed microscopically and some degree of response in the quantitative β-gal assay. β-Gal activity was induced 2-fold after 6 hr in line I.134 and 3-fold after 24 hr in line I.214. After 48 hr, the activity in these lines was induced 20- and 16-fold, respectively. The other seven RA-induced cell lines were found to be induced 1.5- to 12-fold between 24 and 48 hr.

A temporal and quantitative analysis of RA repression is complicated by possible differences in stability of *lacZ* fusion transcripts and proteins. Nevertheless, in nine ES cell clones, β-gal activity was reduced to <0.5 of the control value after 48 hr and showed some degree of repression after 24 hr of RA treatment. In two lines (R.108 and R.140), the level of repression observed was only 0.75 and 0.78 of the control values, respectively. Although this repression is marginal, these lines were included in the analysis as there was a clear qualitative difference in the pattern of staining when observed microscopically on three independent assays.

Northern Blot and Southern Blot Analyses. Northern blot analysis (Fig. 1) of ES cell RNA confirmed that induction of the *lacZ* fusion transcript could be observed in line I.214 (Fig. 1*A*) after 12 hr and in lines I.114 (Fig. 1*C*) and I.193 (Fig. 1*B*) after 48 hr. To date, repression of *lacZ* was confirmed at the transcriptional level for ES cell lines R.68 and R.124 (L.M.F. and Pall, unpublished data). The fusion transcripts in most of these cell lines were ≈4 kb, indicating that ≈600 bp of endogenous sequence was fused to *lacZ* and therefore that these integrations were most likely in the 5′ transcribed regions of the genes. Similar observations were made for the majority of the cell lines isolated in this study (data not shown).

The number of gene trap integrations in eight lines was assessed by Southern blot analysis using a *lacZ* probe on DNA

Table 2. Expression analysis in chimeric embryos derived from RA-induced or -repressed ES cell lines

Cell line	Days of gestation* 8.5	9.5	10.5	11.5	Staining pattern
RA induced					
I.9	—	0/3	—	3/3	Medial regions of somites, two lateral stripes in spinal cord
I.23	—	0/2	3/3	—	Dorsal midbrain, yolk sac endoderm
I.75	1/1	4/6	6/7	8/9	Yolk sac endoderm
I.114	0/4	6/6	5/5	—	Fetal liver, yolk sac endoderm
I.134	2/2	3/4	1/1	4/7	Yolk sac mesoderm
I.163	0/3	2/2	—	3/3	Posterior spinal cord, adjacent mesoderm, distal limb bud, yolk sac mesoderm
I.193	—	3/3	2/2	—	Limb bud mesenchyme, branchial arches, facial mesenchyme, posterior somites, yolk sac endoderm
I.210	—	0/4	4/6	3/3	Dorsal tip of otic vesicle, eye, yolk sac endoderm
I.214	0/1	0/1	—	2/2	Rostral regions of somites, two lateral stripes in adjacent spinal cord
RA repressed					
R.24	1/1	3/10	6/11	—	8.5 dpc, neural ectoderm, primitive gut; 10.5 dpc, telencephalon, hindbrain, eyes, branchial arches, limb bud, heart, liver, somites
R.68	2/2	2/2	2/2	—	8.5 dpc, widespread; 9.5/10.5 dpc, heart, branchial arches, craniofacial region, ventral hind- and midbrain, dorsal roof of hindbrain, limb buds, amnion
R.77	1/1	5/6	3/3	—	Yolk sac endoderm
R.108	—	2/2	5/5	4/4	8.5/9.5 dpc, branchial arches, craniofacial region, hindbrain, yolk sac; 10.5/11.5 dpc, anterior distal tip of limb buds, heart, posterior spinal cord and adjacent mesoderm, yolk sac
R.121	1/6	0/8	0/3	—	Two stripes in hindbrain, weak in heart and allantois
R.124	—	2/2	2/3	—	Heart, branchial arch and dorsal hindbrain
R.140	—	2/2	2/2	—	9.5 dpc, weak expression in heart, dorsal roof of the hindbrain, branchial arches, yolk sac; 10.5 dpc, as at 9.5 dpc plus anterior distal tip of the limb buds
R.170	—	0/5	0/4	0/4	No expression
R.179	—	4/4	3/3	3/3	Yolk sac mesoderm, amnion
R.194	—	0/2	0/2	3/4	Scattered cells in limb bud ectoderm, tail tip, heart, yolk sac
R.213	—	4/4	2/2	—	Heart, presumptive pancreatic primordium, yolk sac endoderm

—, Not determined.
*No. of *lacZ* expressing embryos/total no. of embryos analyzed.

digested with *Eco*RI, which cuts once within the gene trap vector. One (R.140), two (I.23, I.214, and I.193), three (I.163), or four (R.24 and I.114) restriction fragments were detected indicating that one to four copies of the gene trap vector had integrated into the cell lines (Table 3). For lines I.23 and I.163, only one restriction fragment segregated with the observed expression pattern in the F_2 and later generations. In lines I.114 and I.193, the multiple bands cosegregated with the expected patterns of expression for up to four generations, suggesting tandem integration into a single site (data not shown).

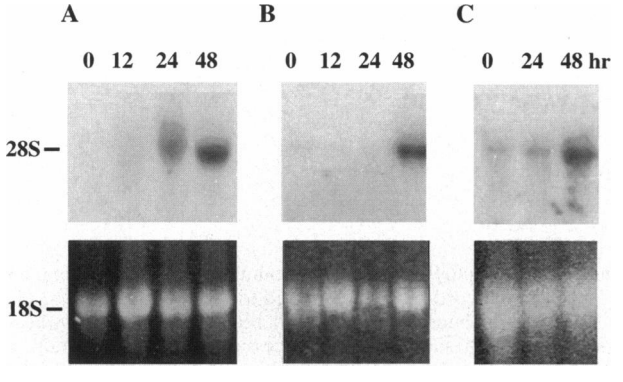

FIG. 1. Northern blot analysis of cell lines I.214 (*A*), I.193 (*B*), and I.114 (*C*) after different times (indicated above each lane) of RA exposure hybridized with a *lacZ*-specific probe. Ethidium bromide-stained gels are shown as loading controls.

Embryonic Expression Patterns of RA-Responsive Genes. In embryos derived from three RA-induced lines, *lacZ* expression was restricted to spatially defined regions along the anterior/posterior (A/P) axis in the spinal cord and adjacent somites (Table 2; Fig. 2 *A–C*). Another RA-induced line generated embryos in which *lacZ* expression was observed in the posterior somites, the branchial arches and facial mesenchyme, as well as in the limb bud (Fig. 2*D*). The remaining five RA-induced lines were expressed in specific tissues including the dorsal midbrain (Fig. 2*I*), the fetal liver (Fig. 2*L*), and extraembryonic tissues (data not shown).

Ten of the 11 lines that showed repression of β-gal activity *in vitro* showed spatially restricted reporter gene expression in chimeric embryos *in vivo* (Table 2; Fig. 2) and the patterns, in general, differed from those generated from the RA-induced lines. For example, none of the RA-repressed lines showed expression in restricted regions along the A/P axis compared to 3 of 9 RA-induced liens. In addition, 8 RA-repressed lines expressed the reporter gene in the developing heart, whereas none of the RA-induced liens was expressed in this tissue. In embryos derived from 4 of the RA-repressed lines, expression was also seen in scattered cells within the hindbrain, cranial–facial region, and branchial arches. The scattered reporter gene expression pattern and the apparent flow of these cells is

suggestive of neural crest cells. In embryos derived from line R.121, β-gal activity was restricted to two bands across the hindbrain at 8.5 dpc, but this very restricted pattern was seen only in 1 of 6 embryos at this stage. No expression was seen at later stages in 11 embryos. Three RA-repressed lines showed reporter gene expression in the developing limb buds (Fig. 2 *E*, *F*, and *H*). One of the lines also showed expression throughout the neurectoderm of 8.5-dpc embryos (Fig. 2*G*) and, after neural tube closure, this broad expression pattern was lost and expression became restricted to regions of the telencephalon, hindbrain, eyes, branchial arches, somites, limb buds, and fetal liver (Fig. 2*H*).

It is interesting to note that none of the trapped genes that responded to RA *in vitro* was ubiquitously expressed *in vivo* and only one line failed to express the reporter gene in embryos at the developmental stages analyzed.

Germ-Line Transmission of the Integrations in RA-Responsive Genes. Six of the RA responsive integrations were transmitted through the germ line and in all cases *lacZ* expression analysis of heterozygous embryos confirmed the

Table 3. Germ-line transmission and RACE–PCR cloning results

Cell line	No. of vector integrations	Fusion transcript cloned	Homozygous phenotype
I.23	2*	—	Viable
I.114	4	150-bp novel sequence	Viable
I.163	3*	120-bp novel sequence	Viable
I.193	2	256-bp novel sequence	Viable
I.214	2	200-bp homology to 5′ UTR of rat muscarinic acetylcholine receptor subtype M4	—
R.24	4	130-bp c-*fyn*	—
R.140	1	205-bp novel sequence	Embryonic lethal

—, Not determined; UTR, untranslated region.
*Segregation of one restriction fragment with the expression patterns after the F_2 generation.

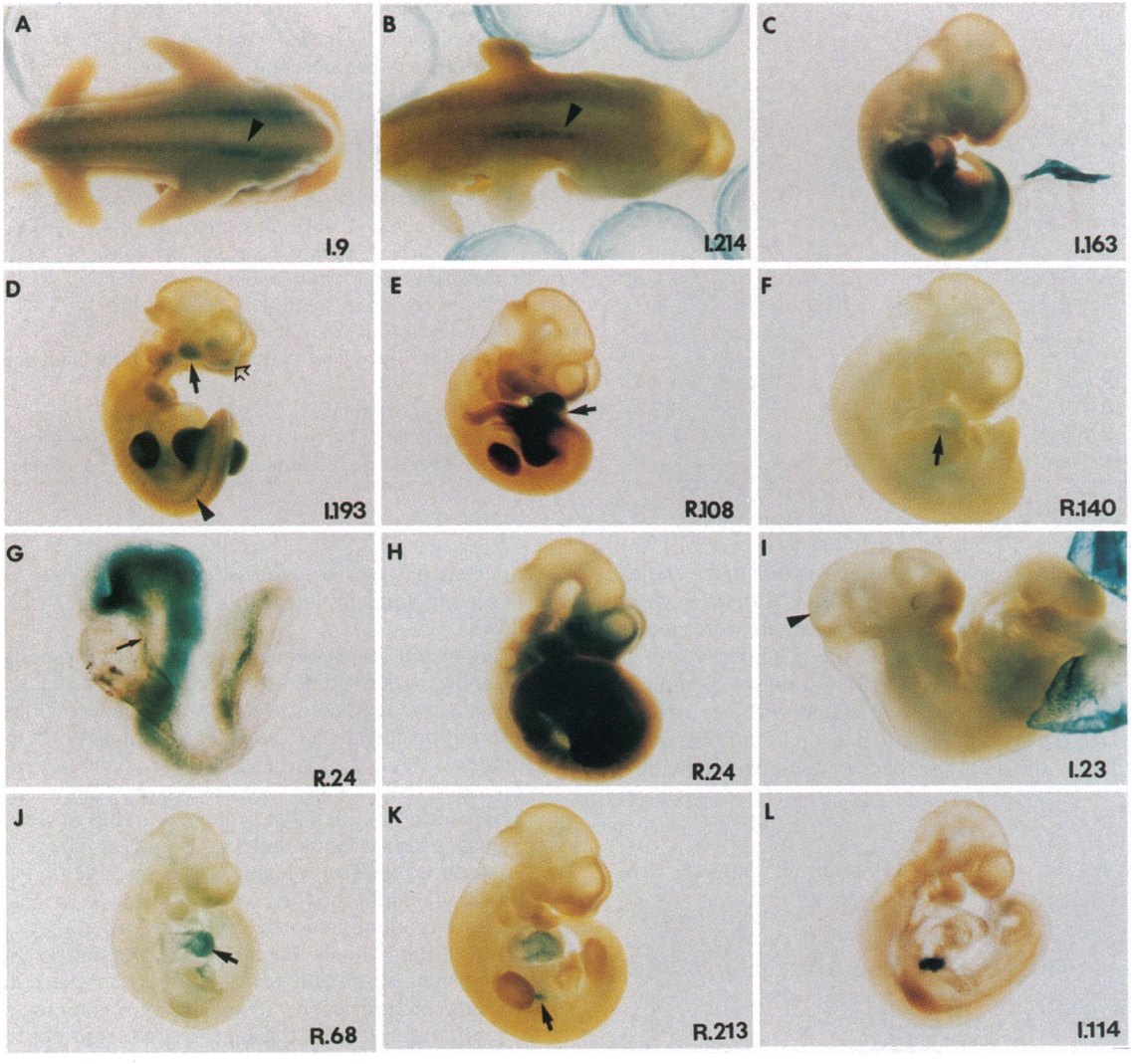

FIG. 2. Expression patterns of gene trap integrations of 8.5- to 11.5-dpc chimeric embryos. (*A*) I.9 (11.5 dpc) somites and adjacent spinal cord (arrow) from the hindbrain to the tail bud. (*B*) I.214 (11.5 dpc) somites and adjacent spinal cord (arrow) restricted to the thorax. (*C*) I.163 (11.5 dpc) posterior spinal cord, adjacent somites, and yolk sac mesoderm. (*D*) I.193 (10.5 dpc) limb buds, somites (arrow), branchial arches (arrowhead), and craniofacial region (open arrow). (*E*) R.108 (10.5 dpc) limb buds (arrow) and adjacent mesoderm (heart expression covered by the tail). (*F*) R.140 (11.5 dpc) anterior tip of the limb bud (arrow). (*G*) R.24 (8.5 dpc) neurectoderm and the primitive gut (arrow). (*H*) R.24 (10.5 dpc) telencephalon, eyes, branchial arches, limb bud, and somites (liver is covered by the tail). (*I*) I.23 (11.5 dpc) dorsal midbrain (arrow). (*J*) R.68 (9.5 dpc) heart, weak expression in scattered cells in the dorsal and ventral hindbrain, ventral midbrain, branchial arches, and the limb buds (dorsal staining in the hindbrain can be seen only in a dorsal view). (*K*) R.213 (11.5 dpc) heart, pancreatic primordium (arrow). (*L*) I.114 (9 dpc) fetal liver.

expression patterns observed in the aggregation chimeras (data not shown). To date, genotyping of offspring from F_1 intercrosses at 3 weeks of age has revealed four homozygous viable integrations (I.23, I.114, I.163, and I.193) and one homozygous embryonic lethal (R.140). The homozygous lethal phenotype of line R.140 was confirmed by comparing, using densitometry, the intensity of the *lacZ*-specific band to an internal standard probe (*En-2*). Of 146 intercross offspring, 87 (60%) were heterozygous, 59 (40%) were wild type, and no homozygotes were detected. The reliability of typing by quantitative Southern blots was assessed and confirmed by test breeding. The homozygous lethal phenotype of line R.140 was further confirmed by genotyping retarded 12.5-dpc embryos as homozygous.

Cloning of Trapped Genes. RACE–PCR cloning was used to isolate endogenous gene exon sequences upstream of the gene trap *lacZ* sequences from six cell lines (Table 3). The sequence of one clone, cell line R.24, was identical to nucleotides 160–290 of the c-*fyn* cDNA (20). This sequence included the initiating methionine codon (positions 226–228), which has been reported to lie within exon 2. Vector integration into the second intron was confirmed by Southern blot analysis of *Bam*HI-digested ES cell DNA probed with a c-*fyn* exon 2 probe (Fig. 3) (21). A 200-bp sequence generated from the *lacZ* fusion transcript in line I.214 showed 85% identity (over an 80-bp region) to the 5′ untranslated region of the rat muscarinic acetylcholine receptor (subtype M4). In addition, this clone contained sequences that were similar to an uncharacterized repeat sequence. Sequences that did not correspond to any known sequence in GenBank were generated from four lines (Table 3).

DISCUSSION

We have used RA as an inducing agent to demonstrate the general applicability of an induction gene trap screen in ES cells to identify genes that lie downstream of ligand/receptor-mediated signaling pathways. From $\approx 4 \times 10^3$ integrations, 20 ES cell integrations were identified that responded to RA *in vitro*. All but one of these lines displayed a restricted pattern of expression during late gastrulation to midgestation.

This induction gene trap screen identified a significantly higher proportion of genes that are expressed in a restricted manner compared to gene trap screens in which lines were selected solely on the basis of *lacZ* expression in undifferentiated ES cells (3–6). In such screens, only 14% of the integrations showed specific patterns of expression at 8.5 dpc and 30% displayed specific patterns of expression in midgestation embryos (5, 6). In contrast, 5 of 9 (56%) RA-responsive ES cell lines showed restricted patterns at 8.5 dpc and 19 of 20 (95%) showed such patterns between 8.5 and 11.5 dpc. Furthermore, in this induction screen, none of the integrations was constitutively expressed compared to the 32% that showed widespread expression in a screen where cell lines were selected on the basis of *lacZ* expression in undifferentiated ES cells (6). Thus, the induction gene trap approach significantly enriches for integration events into genes that respond to an inducer *in vitro* and show restricted expression *in vivo*. Future work will determine whether the response of these genes to RA *in vitro* is also observed *in vivo*.

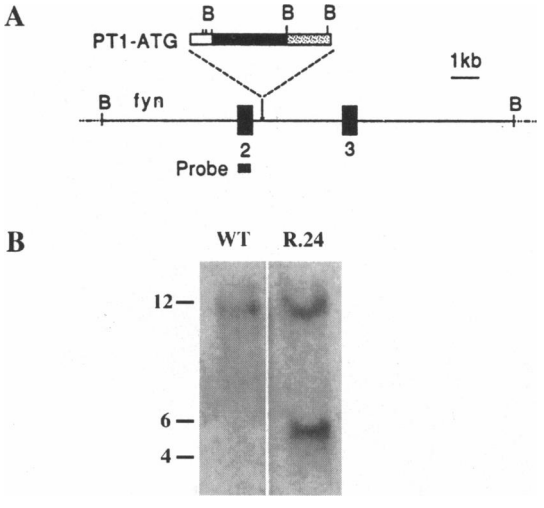

FIG. 3. (*A*) Schematic presentation of the predicted gene trap vector integration into the second intron of c-*fyn* protooncogene. Restriction enzyme sites *Bam*HT (B) and the hybridization probe used are indicated. (*B*) Southern blot analysis of wild-type (WT) and R.24 DNA digested with *Bam*HT and hybridized with c-*fyn* exon 2 probe. The 5.5-kb fragment in lane R.24 represents the mutant c-*fyn* allele.

The *in vitro* induction protocol used in this study was designed to identify genes that were either directly or indirectly responsive to RA. In one of the induced lines, the fusion transcript responded after 6 hr of RA treatment but most lines were induced between 24 and 48 hr. These observations suggest that most of the trapped genes are not direct targets of the RA receptor complexes but more likely lie further downstream in the RA pathway. However, the promoter region of the laminin B1 gene contains a RA receptor binding element and the induction of expression of this gene is not seen until 26 hr after RA treatment (22). Thus, a late inductive response is not necessarily an indication of an indirect mode of RA regulation. Clearly, to determine the mode of RA regulation of these genes, it will be necessary to characterize their transcriptional regulatory regions and cross them with mice carrying mutations in RA receptors.

We identified ES cell lines in which the reporter gene was partially repressed by RA, further demonstrating the power of this approach to delineate downstream events in genetic pathways. Interestingly, the expression patterns in embryos derived from the repressed lines were distinct from those derived from inducible lines, suggesting that there may be qualitative differences between these two classes of genes. For example, none of the nine integrations that were induced by RA were expressed in the developing heart, whereas 8 of the 11 repressed lines showed some expression in this tissue.

Three induced lines showed restricted patterns of expression in different regions along the A/P axis in the spinal cord and in the adjacent somites, reminiscent of aspects of *Hox* gene expression (23, 24). This is consistent with the fact that transcriptional regulation of *Hox* genes is modulated by RA *in vitro* (9, 10, 25, 26) and *in vivo* (12, 13, 23). Administration of RA during embryogenesis alters *Hox* expression, resulting in vertebral column transformations (23, 24) and mutation of the RA receptor γ gene leads to specific homeotic transformations in the rostral axial skeleton (27). Therefore, it is tempting to speculate that the three genes identified in this study are downstream targets of RA receptors and/or of homeobox-containing transcription factors.

Heart defects have been reported in RA-treated and RA-deprived embryos (28, 29). In addition, mice homozygous for a targeted mutation in the retinoid X receptor α gene exhibit fetal heart defects (30), providing direct evidence that RA and its receptors are involved in heart development. Interestingly, eight RA-repressed integrations identified in this study were expressed in the developing heart. It is possible that some of these integrations were in genes expressed in cardiac neural crest-derived cells because four of the heart lines also showed expression in the dorsal hindbrain and craniofacial areas.

We have successfully cloned fusion transcripts for six RA-responsive integrations and five appear to be novel genes. In ES cell line R.24, the gene trap vector integrated into the second exon of the c-*fyn* protooncogene. As expected, the LacZ expression pattern matches that reported for *fyn* (31).

The *fyn* gene is likely to be inactivated by the gene trap integration as the predicted fusion protein would include only the first 21 amino acids of Fyn. Preliminary experiments have shown that the level of normal *fyn* transcript in the R.24 ES cell line is half that of levels in control ES cells (data not shown), which would support this prediction. It is possible that a low level of normal *fyn* transcript is made due to splicing out of the gene trap sequences. Targeted disruption of *fyn* results in abnormal T-cell signaling (21) and long-term potentiation (32). It will therefore be interesting to compare the phenotype generated by this gene trap integration to the targeted mutant.

It should be possible, by scaling up an induction gene trap screen, to identify most genes in a given genetic pathway that are active in ES cells. This approach could also be applied to identify genes downstream of transcription factors, by establishing cell lines in which the expression of the gene of interest can be controlled in an inducible manner (33). Clearly, this approach has broad applications and will allow for the simultaneous identification and mutation of novel genes involved in a variety of developmental processes in mammals.

We wish to thank Drs. Janet Rossant and David Hill for stimulating discussions and critical reading of the manuscript and the anonymous reviewers for their helpful comments. We thank Lynn Anson-Cartwright and Malgosia Kownacka for their technical assistance and Dr. S. Grant for c-*fyn* probes. This work was supported by grants from the National Institutes of Health (A.L.J.), the Medical Research Council of Canada (A.L.J. and A.B.), the National Cancer Institute of Canada (A.L.J. and A.B.), and Bristol-Myers Squibb. A.L.J. was an International Scholar of the Howard Hughes Medical Institute (HHMI) and Medical Research Council Scientist and A.B. is an International Research Scholar of the HHMI. L.M.F. was partly supported by a European Molecular Biology Organization fellowship and by the British Medical Research Council. W.W. was partly supported by fellowships from the Deutsche Forschungsgemeinschaft.

1. Joyner, A. L. (1992) *BioEssays* **13,** 649–656.
2. Hill, D. P. & Wurst, W. (1993) in *Current Topics of Developmental Biology*, ed. Pederson, R. A. (Academic, San Diego), pp. 181–206.
3. Skarnes, W. C., Auerbach, B. A. & Joyner, A. L. (1992) *Genes Dev.* **6,** 903–918.
4. Gossler, A., Joyner, A. L., Rossant, J. & Skarnes, W. C. (1989) *Science* **244,** 463–465.
5. Friedrich, G. & Soriano, P. (1991) *Genes Dev.* **5,** 1513–1523.
6. Wurst, W., Rossant, J., Prideaux, V., Kownacka, M., Joyner, A. L., Hill, D. P., Guillemot, F., Gasca, S., Auerbach, A. & Ang, S.-L. (1995) *Genetics* **139,** 889–899.
7. Linney, E. (1992) *Curr. Top. Dev. Biol.* **27,** 309–350.
8. Eichele, G. (1989) *Trends Genet.* **5,** 246–251.
9. Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E. & Mavilio, F. (1990) *Nature (London)* **346,** 763–766.
10. Simone, A., Acampora, D., Nigoro, V., Faiella, A., D'Esposito, M., Stornaiuolo, A., Mavilio, F. & Boncinelli, E. (1991) *Mech. Dev.* **33,** 215–228.
11. Pruitt, S. C. (1992) *Development (Cambridge, U.K.)* **116,** 573–583.
12. Conlon, R. A. & Rossant, J. (1992) *Development (Cambridge, U.K.)* **116,** 357–368.
13. Marshall, H., Nonchev, S., Sham, M. H., Muchamore, I., Lumsden, A. & Krumlauf, R. (1992) *Nature (London)* **360,** 737–741.
14. Hill, D. P. & Wurst, W. (1993) *Methods Enzymol.* **225,** 664–681.
15. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8424–8428.
16. Wurst, W. & Joyner, A. L. (1993) in *Gene Targeting: A Practical Approach*, ed. Joyner, A. L. (IRL, Oxford), pp. 33–61.
17. Eustice, D. C., Feldman, P. A., Colberg-Poley, A. M., Buckery, R. M. & Neubauer, R. H. (1991) *BioTechniques* **11,** 739–742.
18. Nagy, A. & Rossant, J. (1993) in *Gene Targeting: A Practical Approach*, ed. Joyner, A. L. (IRL, Oxford), pp. 147–178.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
20. Cooke, M. P. & Perlmutter, R. M. (1989) *New Biol.* **1,** 66–74.
21. Stein, P. L., Lee, H. M., Rich, S. & Soriano, P. (1992) *Cell* **70,** 741–750.
22. Vasios, G., Mader, S., Gold, J. D., Leid, M., Lutz, Y., Gaub, M.-P., Chambon, P. & Gudas, L. J. (1991) *EMBO J.* **10,** 1149–1158.
23. Kessel, M. (1992) *Development (Cambridge, U.K.)* **115,** 587–601.
24. Kessel, M. & Gruss, P. (1991) *Cell* **67,** 89–104.
25. Langston, A. W. & Gudas, L. J. (1992) *Mech. Dev.* **38,** 217–228.
26. Pöpperl, H. & Featherstone, M. S. (1993) *Mol. Cell. Biol.* **13,** 257–265.
27. Lohnes, D., Kastner, P., Dierich, A., Mark, M., LeMeur, M. & Chambon, P. (1993) *Cell* **73,** 643–658.
28. Wilson, J. G. & Warkany, L. (1949) *Am. J. Anat.* **85,** 113–155.
29. Kirby, M. L. (1993) *Trends Cardiovasc. Med.* **3,** 18–23.
30. Sucov, H. M., Dyson, E., Gumeringer, C. L., Price, J., Chien, K. R. & Evans, R. M. (1994) *Genes Dev.* **8,** 1007–1018.
31. Yagi, T., Shigetani, Y., Furuta, Y., Nada, S., Okado, N., Ikawa, Y. & Aizawa, S. (1994) *Oncogene* **9,** 2433–2440.
32. Grant, S. G., O'Dell, T. J., Karl, K. A., Stein, P. L. & Soriano, P. (1992) *Cell* **70,** 741–750.
33. Gossen, M. & Bujard, H. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 5547–5551.