ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/11843884

Sorting nexin-14, a gene expressed in motoneurons trapped by an in vitro preselection method

ARTICLE *in* DEVELOPMENTAL DYNAMICS · SEPTEMBER 2001

Impact Factor: 2.38 · DOI: 10.1002/dvdy.1163 · Source: PubMed

CITATIONS
17

READS
15

5 AUTHORS, INCLUDING:

Patrick Carroll
French Institute of Health and Medical Rese…
62 PUBLICATIONS **5,532** CITATIONS
SEE PROFILE

Odile Gayet
French Institute of Health and Medical Rese…
21 PUBLICATIONS **664** CITATIONS
SEE PROFILE

Available from: Patrick Carroll
Retrieved on: 26 January 2016

DEVELOPMENTAL DYNAMICS 221:431–442 (2001)

Sorting Nexin-14, a Gene Expressed in Motoneurons Trapped by an In Vitro Preselection Method

PATRICK CARROLL, YANNICK RENONCOURT, ODILE GAYET, BÉATRICE DE BOVIS, AND SERGE ALONSO*

INSERM U.382, Developmental Biology Institute of Marseille (IBDM), CNRS/INSERM/Université de la Méditerranée/AP de Marseille, Campus de Luminy, Marseille, France

ABSTRACT **A gene-trap strategy was set up in embryonic stem (ES) cells with the aim of trapping genes expressed in restricted neuronal lineages. The vector used trap genes irrespective of their activity in undifferentiated totipotent ES cells. Clones were subjected individually to differentiation in a system in which ES cells differentiated into neurons. Two ES clones in which the trapped gene was expressed in ES-derived neurons were studied in detail. The corresponding cDNAs were cloned, sequenced, and analysed by in situ hybridisation on wild-type embryo sections. Both genes are expressed in the nervous system. One gene, *YR-23*, encodes a large intracellular protein of unknown function. The second clone, *YR-14*, represents a *sorting nexin* (*SNX14*) gene whose expression in vivo coincides with that of LIM-homeodomain *Islet-1* in several tissues. Sorting nexins are proteins associated with the endoplasmic reticulum (ER) and may play a role in receptor trafficking. Gene trapping followed by screening based on in vitro preselection of differentiated ES recombinant clones, therefore, has the potential to identify integration events in subsets of genes before generation of mouse mutants.** **© 2001 Wiley-Liss, Inc.**

Key words: ES cells; gene-trapping; in vitro differentiation; in vitro preselection; neurons; Islet; sorting nexin

INTRODUCTION

Functional genomics will be a major challenge for biologists in the future as the complete mouse genome sequence becomes available. Presently, the mouse is the most amenable mammal for genetic dissection of in vivo biological functions and providing models for human diseases. We anticipate that new techniques adapted to large-scale mutagenesis of the mouse genome will complement the more straightforward but less rapid "knockout" techniques currently in use (Wurst et al., 1995). In this context, gene trapping in embryonic stem (ES) cells provides a tagged insertional mutagenesis suited for large-scale studies (reviewed by Zambrowicz and Friedrich, 1998). Gene trapping by random insertional mutagenesis tags genes in ways that allows facile determination of their expression pattern and DNA sequence, as well as the generation of mouse mutants (Gossler et al., 1989; reviewed by Evans et al., 1997; Evans, 1998).

A further development of gene trap strategies has involved the in vitro differentiation of ES cells toward defined cell lineages to detect genes expressed in certain cell types (reviewed by Keller, 1995). This type of system theoretically allows the prediction of potential in vivo sites of expression of the genes of interest before embarking on the expensive process of deriving mutant mouse lines.

Here, we describe a modified gene trap strategy by using a new vector and based on the in vitro preselection of ES recombinant clones in which the trapped gene is expected to be expressed in the nervous system. Gene-trapped ES cell clones were screened for the activity of β-galactosidase (β-gal) in embryoid bodies at day 7 of differentiation (EB7), a time window in which postmitotic neurons are generated in large numbers in our differentiation system (Renoncourt et al., 1998). Subsequent characterization of reporter gene expression after EB dissociation and cell plating allowed us to identify two genes expressed in neurons. In vivo expression analysis of these genes showed that both were expressed in the embryonic mouse nervous system. Sequence analysis identified one gene as a member of the sorting nexin family; proteins thought to be involved in intracellular protein trafficking (Haft et al., 1998). We have analysed the expression of this gene, *sorting nexin 14* (*SNX14*) during embryonic development in the mouse.

RESULTS

In Vitro Preselection Strategy for the Identification of Integration Events in Genes Expressed in the Nervous System

To trap genes in totipotent ES cells irrespective of transcriptional activity, we designed a "gene and polyA

Drs. Carroll and Renoncourt made equal contributions to this work.

Dr. Renoncourt's present address is Department of Microbiology, Umea University, S-90187 Umea, Sweden.

*Correspondence to: Serge Alonso, INSERM U.382, Developmental Biology Institute of Marseille (IBDM), CNRS/INSERM/Université de la Méditerranée/AP de Marseille, Campus de Luminy - Case 907, 13288 Marseille Cedex 09, France. E-mail: alonso@ibdm.univ-mrs.fr

Received 7 September 2001; Accepted 20 April 2001

© 2001 WILEY-LISS, INC.

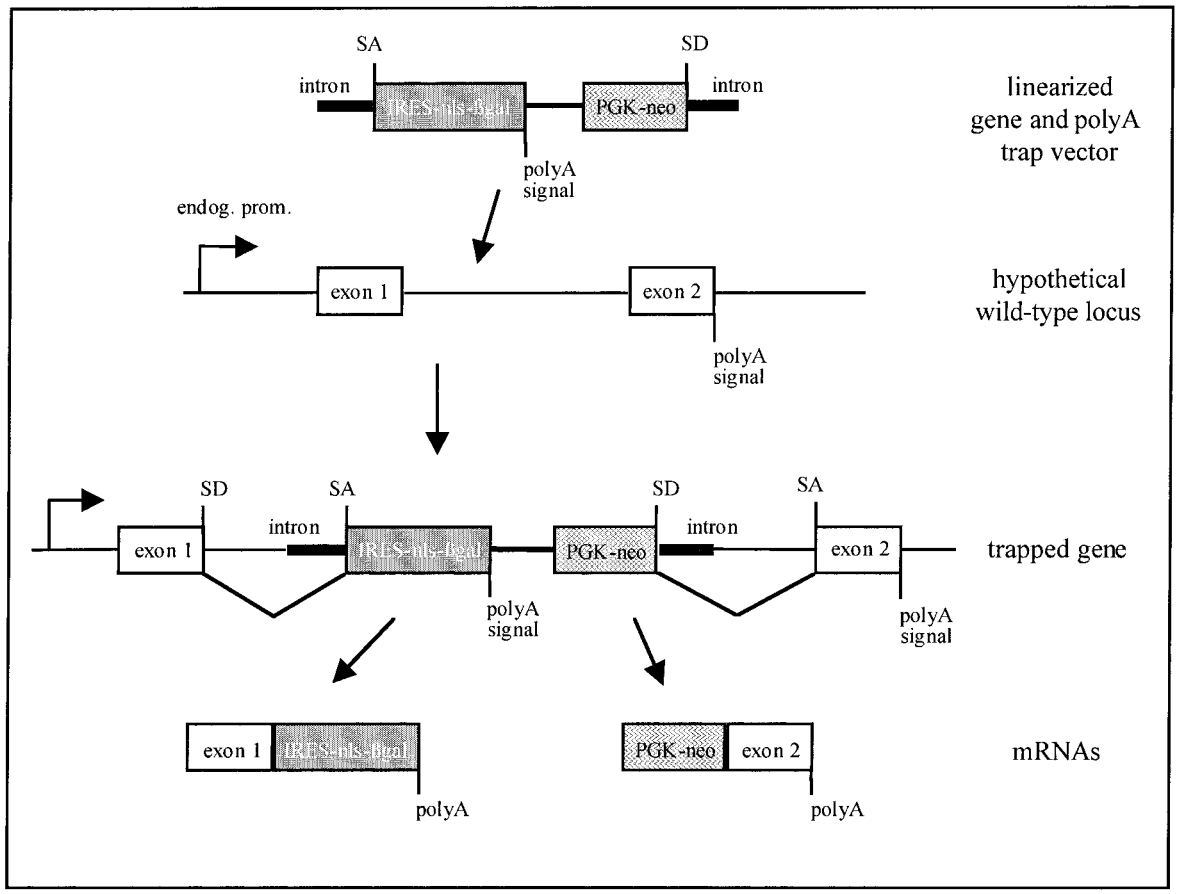

Fig. 1. Hypothetical integration of the gene and polyA trap vector. The schematic diagram shows the IRESnlsβgalNeo(-pA) vector after integration into the first exon of a hypothetical two-exon gene. The vector contains the promoterless β-galactosidase reporter gene and the PGK-Neo (polyA-) gene flanked at the 5′ and 3′ ends by engrailed-2 intron/splice acceptor (SA) and splice donor(SD)/fyn intron sequences respectively, such that an integration into an intron of a transcriptional unit in the correct 5′→3′ orientation can give rise to functional expression of both genes. The arrow on the genomic locus indicates the endogenous promoter of the trapped gene. The resulting integration would produce two fusion transcripts, one containing sequences from the trapped gene at the 5′ end and the β-gal sequence at the 3′ end, and another transcript containing the neomycin sequence at the 5′ end and sequences of the trapped gene at the 3′ end. PGK, phosphoglycerate kinase-1 promoter; neo, neomycin gene; IRES, internal ribosome entry site; nls, nuclear localisation signal; SA, splice acceptor; SD, splice donor.

trap" vector, IRESnlsβgalNeo(-pA), in which *lacZ* expression is under the control of the trapped promoter and the expression of the neomycin resistance gene (neo^R) is controlled by a constitutive promoter (Fig. 1). Because neo^R lacks its own polyA signal, the protein is absent unless its RNA transcript, containing a 3′-splice donor (3′-SD), is spliced downstream to exons that include a polyadenylation (polyA) signal. Second, the vector contains an RNA 5′-splice acceptor (5′-SA) upstream of the promoterless *lacZ* reporter gene whose expression depends on the trapped *cis*-regulatory sequences. A nuclear localization signal (nls) is fused in frame to the coding sequence of *lacZ* to concentrate the enzyme in the nucleus and, thereby, enhances the sensitivity of detection.

Undifferentiated ES cells were electroporated and G418-resistant ES cell clones were isolated and assayed for β-gal activity by Xgal histochemical staining. At this step, 25% (26 of 104) neo^R clones tested were blue (β-gal+ ES cells in Table 1). Most of these β-gal+ undifferentiated ES colonies (18 of 26) resulted from the integration of multiple copies of the vector in tandem. In the concatemer-containing clones, *lacZ* may be constitutively expressed under the control of the PGK promoter from the upstream neo^Rgene.

To enrich for genes expressed in the developing nervous system, we used the previously described in vitro system in which ES cells treated with retinoic acid differentiate into neurons (Renoncourt et al., 1998). The 8 nonconcatemer β-gal+ ES colonies plus 78 β-gal− ES colonies were submitted individually to in vitro neuronal differentiation through ES cell aggregation in embryoid bodies (EBs) treated by retinoic acid (RA) and EB7 were analyzed by X-gal staining. Six of the eight β-gal+ ES colonies were X-gal− in EB7, whereas the two remaining (clones YR-21 and YR-23) gave rise to blue EB7 (Table 1). Five of the 78 β-gal− ES colonies clones contained β-gal+ cells after differ-

TABLE 1. Cell data[a]

Total neoR ES clones	Concatemers	β-gal⁻ EB7	β-gal⁺ EB7	β-gal⁺ neurons
β-gal⁺ ES cells				
26	18	6	2	
			YR-21, YR-23	*YR-23*
β-gal⁻ ES cells				
78	3	73	2	
			YR-6, YR-14	*YR-14*
104	21	79	4	

Clones were analysed at different stages: undifferentiated ES cells, embryoid bodies at day 7 (EB7), and on cultured neurons. Left hand column and the last row correspond to total numbers of examined clones of each type. Only clones YR-14 and YR-23 give rise to *lacZ*-expressing neurons on culture.
[a]ES, embryonic stem; EB, embryoid body.

entiation, including 3 that were concatemers and 2 that were nonconcatemers (clones YR-6 and YR-14). The four clones (YR-21, YR-23, and YR-6, YR-14) that were blue at EB7 were then dissociated and plated on polyornithine/laminin-coated dishes (see Renoncourt et al., 1998). After neurite outgrowth 24 hr after plating, cells were stained with X-gal staining. Only two (YR-14 and YR-23) of the four clones were found to express *lacZ* in individual neurons: *YR-23* was expressed in approximately 60% of neurons as well as in non-neuronal cells, whereas the expression of clone YR-14 was restricted to only 1% of neurons. The other two clones (YR-6 and YR-21) had only blue non-neuronal cells in culture. Based on this analysis, we chose clones YR-14 and YR-23 for further study.

Characterization of *lacZ*-Expressing Neurons Derived In Vitro From Preselected YR-14 and YR-23 Gene-Trapped Clones

EB7 from clones YR-14 and YR-23 were dissociated, cells were plated on polyornithine/laminin-coated dishes and 24-hr cultures were double immunostained by an anti–β-gal antibody and neuronal antibodies. Only data for clone YR-14 are presented in Figure 2. Peripherin is a type III intermediate filament that has been shown to be expressed in motor, sensory, and sympathetic neurons (Djabali et al., 1991). Postmitotic neurons were identified by antibodies against the neurofilament NF145 subunit and peripherin. Neurons coexpressing either NF145 and β-gal (Fig. 2A,B) or peripherin and β-gal (Fig. 2C,D) were identified in both clones YR-14 and YR-23.

The two homeoproteins Islet-1 and Islet-2 are present in several neuronal populations, including motoneurons in the CNS, sensory, and sympathetic neurons in the PNS as well as in certain non-neuronal cells in the developing pituitary, pancreas, and gut. We used an antibody that recognizes both Islet-1 and Islet-2 proteins; the staining obtained with this antibody will be referred to as Islet staining. For both clones YR-14 (Fig. 2E–G) and YR-23 (data not shown), a subset of β-gal+ neurons also expressed the marker Islet. To test whether some of the {Islet+, β-gal+} neurons were motoneurons, we used antibodies to the homeobox protein HB9 that is expressed in somatic motoneurons but not in other Islet+ neurons (Harrison et al., 1994; Tanabe et al., 1998). As shown in Figure 2H–J, double-labeled neurons were identified, indicating that the trapped gene *YR-14* was active in neurons expressing a marker of somatic motoneurons. A subset of β-gal+ neurons also expressed the homeoprotein Phox2b, indicating that *YR14*-expressing neurons might include cranial motoneurons and/or sensory neurons (Fig. 2K–M; Pattyn et al., 1997).

Cloning and Identification of cDNA Sequences of *YR-14* and *YR-23* Trapped Genes

We used 3′-RACE (rapid amplification of cDNA ends) and 5′-RACE (Frohman et al., 1988) protocols and library screening to obtain sequence information from clones YR-14 and YR-23 (Fig. 3). For YR-14, two 3′-RACE variants (clones 14a and 14b) were obtained with identical sequences except for the presence of a 224-bp insert in clone 14b. The 5′-RACE product for YR-14 was used to screen a rat cDNA library from E14 ventral spinal cord. A rat cDNA sequence of 3,090 bp with a single open reading frame of 946 amino acids was isolated and corresponds to clone 14a without insert. The rat nucleotide sequence is 94% identical to the mouse sequence. Sequence comparison confirmed that the 5′- and 3′-trapped mouse cDNA sequences represent contiguous sequences in the corresponding rat cDNA.

The full-length cDNA rat sequence of YR-14 was compared with gene and EST database sequences. Two rat EST sequences (accession nos. AA817895 and AI146138; Bonaldo et al., 1996) that matched almost perfectly (99% identical) to the 3′ end of our rat sequence (YR-14a) over more than 500 bp were found. The predicted protein sequence was checked for the presence of particular motifs by using the Searching Protein Sequence Motifs program. This analysis suggests that the trapped gene in clone YR-14 may encode an endoplasmic reticulum associated membrane protein with a single transmembrane helix. By using the SMART (Simple Modular Architecture Research Tool; Schultz et al., 1998) protein analysis software, it seems that this protein is similar in its domain organisation to several proteins that contain, from the NH_2-terminal, a transmembrane helix, a PXA domain (associated with PX proteins), an RGS domain (regulator of G-protein signalling), and a PX domain (domain of unknown function present in Phox proteins, PLD isoforms, and a p13K isoform). Subsequently, we identified this clone as “sorting nexin 14”, a member of a family of at least 14 molecules shown to bind to the intracellular part of several types of cell surface receptors and may be involved in sorting receptors within the cell (Haft et al., 1998). Hereafter, this gene will be referred to as *sorting nexin 14* (*SNX14*).

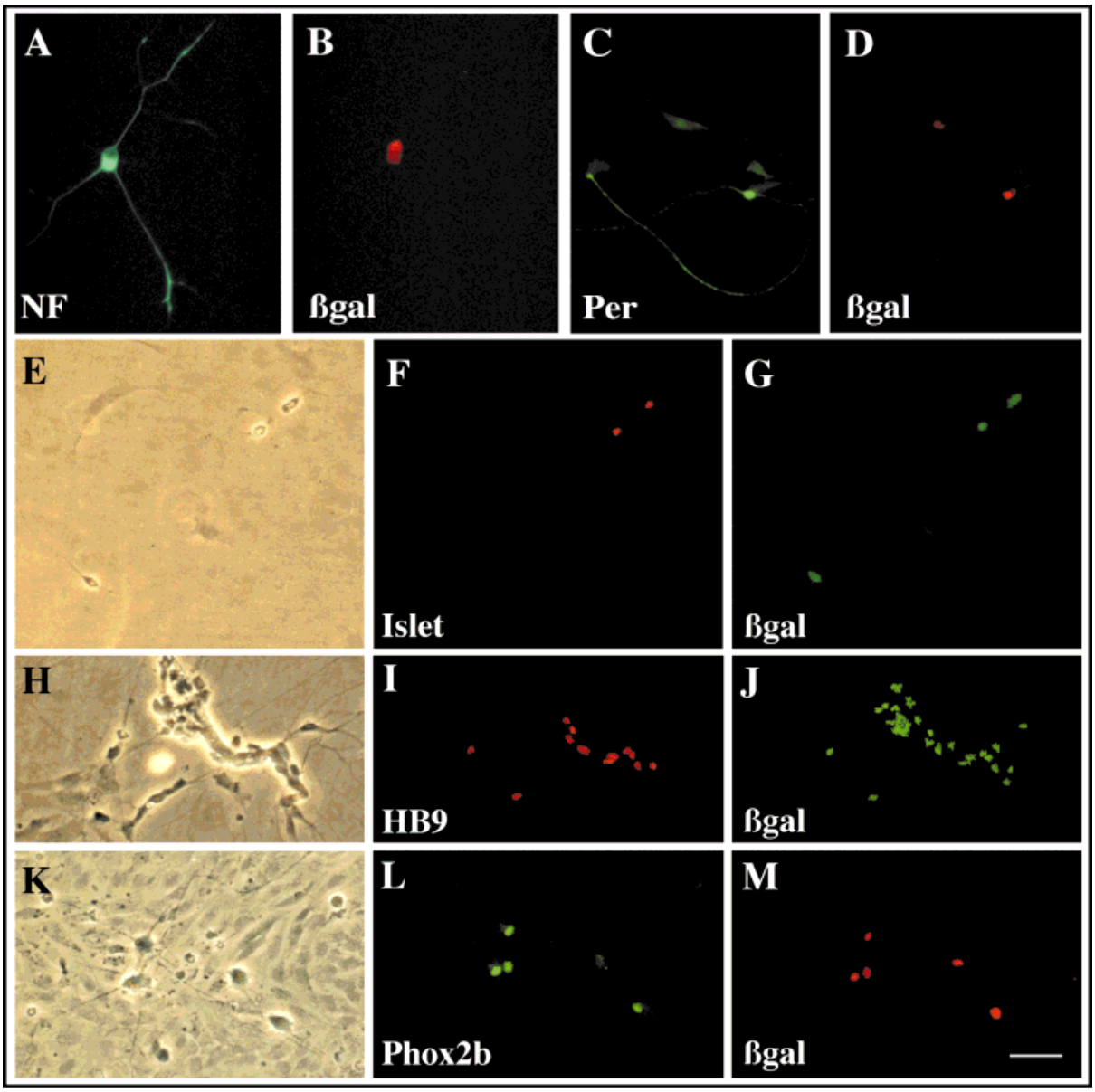

Fig. 2. Expression of neuronal markers in embryonic stem (ES) cell–derived plated neurons from clone YR-14. Double immunolabeling for the indicated antigens were performed after in vitro differentiation of clone YR-14. Pictures were chosen for their high colocalization of β-gal with neuronal markers and, thus, represent an overestimation of the real number (1%) of neurons expressing *YR-14*. **A,B:** Colocalization of β-gal with the neurofilament (NF) subunit, NF145. **C,D:** Colocalization of β-gal with peripherin (per). **F,G** Colocalization of β-gal with Islet. **I,J:** Colocalization of β-gal with the homeobox protein HB9. **L,M:** Colocalization of β-gal with Phox2b. **E,H,K:** Corresponding phase-contrast fields to right panels. Scale bar = 100 μm in M (applies to A–M).

By cloning and sequencing 5′- and 3′-RACE products from the trapped sequence in mouse clone YR-23 (Fig. 3), we found that these two independent sequences were almost (99%) identical to a mouse brain cDNA sequence of 6271 bp (accession no. AF060565) encoding a putative protein of 1,829 residues. Sequence alignment of 5′- and 3′-RACE products with the sequence from databases showed that they represent a contiguous sequence expressed in the mouse brain cDNA. A corresponding human cDNA sequence of 6178 bp (91% identical to clone YR-23) isolated from the KG-1 myeloblast cell line (accession no. D87446; referred as KIAA0257 in Nagase et al., 1996) was subsequently identified by database searching. The predicted mouse and the human protein sequences are almost 100% identical. By Northern blot analysis, Nagase and colleagues (1996) showed that the human mRNA was widely distributed in adult tissues. A search for protein motifs by using the Searching Protein Sequence Motifs program suggests that the protein may be a membrane

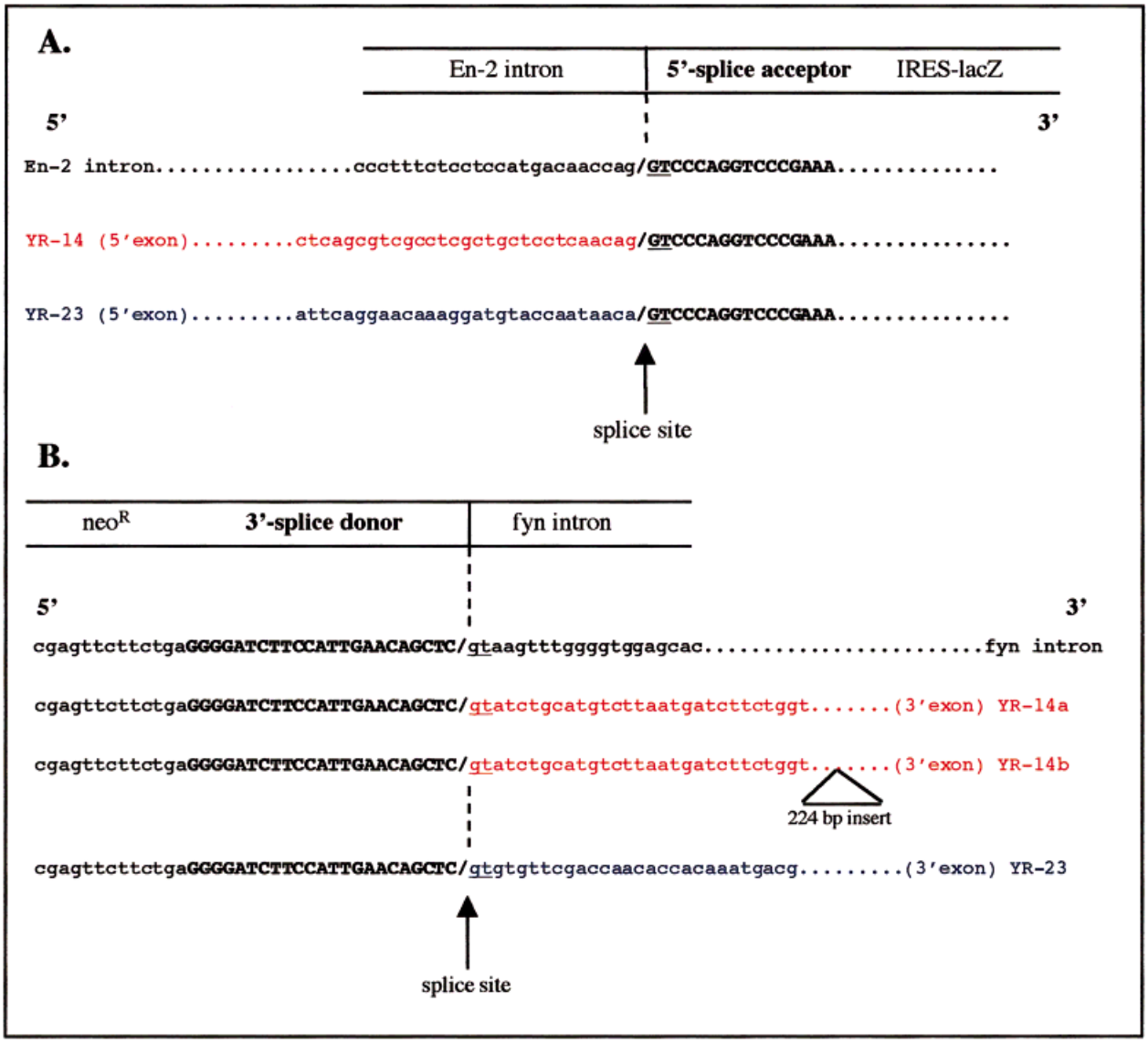

Fig. 3. Partial sequences of gene trap cDNAs YR-14 and YR-23. Sequences of the intron-exon borders after 5′-RACE (**A**) and 3′-RACE (**B**) analysis of clones YR-14 (in red) and YR-23 (in blue). For YR-14, two 3′-RACE variants (14a and 14b) were obtained showing identical sequences except for the presence of a 224-bp insert in the clone YR14b. The 5′-splice acceptor and 3′-splice donor sequences are printed in boldface type.

protein with three transmembrane helices. This sequence analysis demonstrates the correct usage of both the 5′-splice acceptor and 3′-splice donor sequences in our vector.

Expression of *SNX14* and *YR-23* Genes in the Embryo

We focused on the expression of both trapped genes in the developing spinal cord. Total RNA from spinal cord of Swiss mouse embryos at E11.5, E12.5, and E13.5 was hybridized with cDNA probes from clones SNX14 and YR-23 on Northern blot (data not shown). A single 6-kb band was detected with probe YR-23 representing the endogenous mRNA and a single 4-kb band was revealed with SNX14. For each clone, transcripts were detected at comparable levels in E12.5 and E13.5 embryos but at a lower level at E11.5. On Northern blot of different adult tissues, we found *SNX14* expression in cerebellum and hippocampus and at much lower levels in cortex, muscle, liver, lung, and heart (data not shown).

To determine the expression pattern of the trapped genes, we carried out in situ hybridization (ISH) analysis on mouse embryos by using YR-23 and SNX14 cDNA sequences. The expression profiles of endogenous genes *YR-23* and *SNX14* were analyzed on sec-

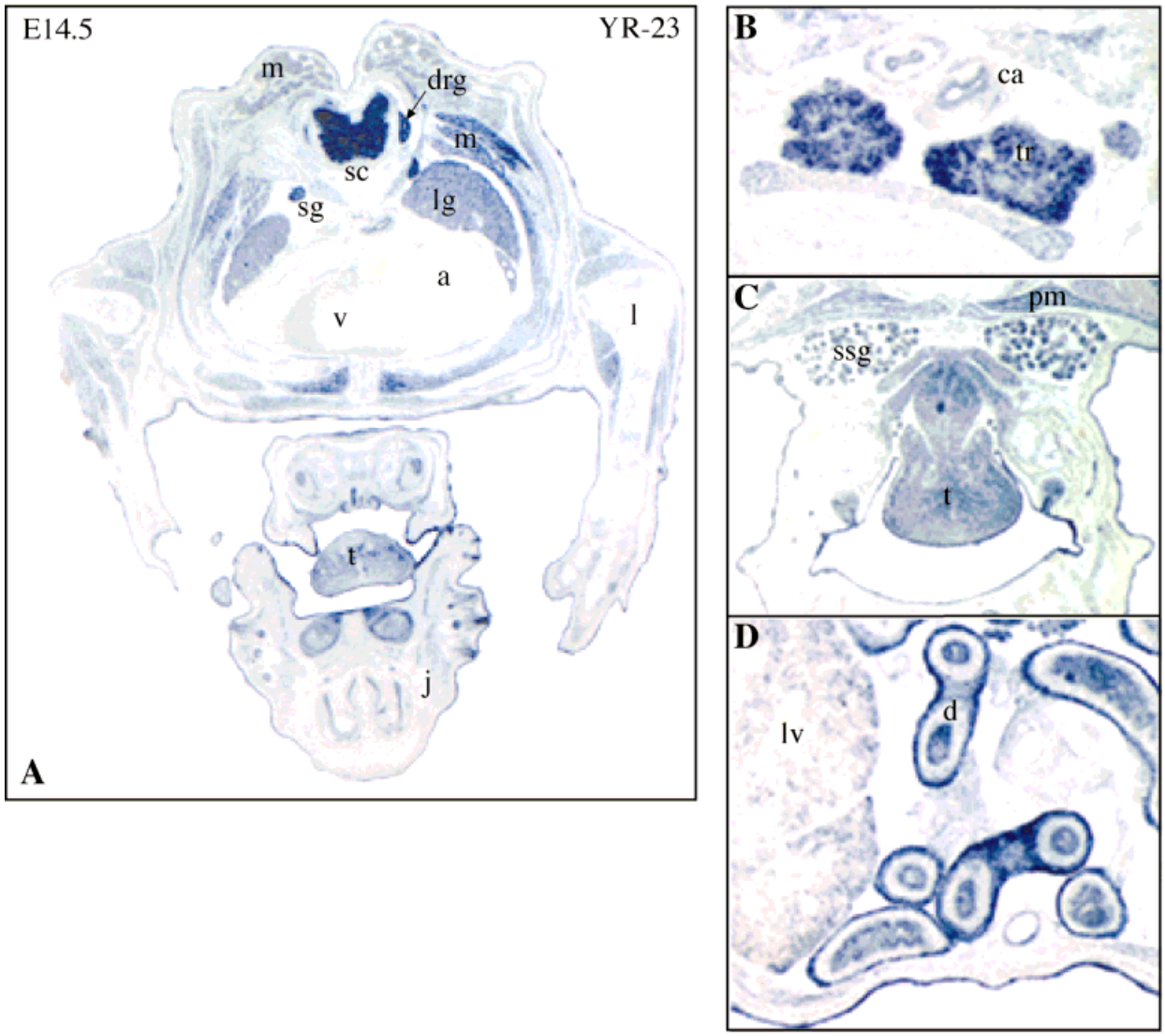

Fig. 4. Expression of *YR-23* in the mouse embryo. In situ hybridisation by using digoxigenin-labeled YR-23 probe on transverse sections of embryonic day (E) 14.5 mouse embryo. **A:** At forelimb (l) level, expression is detected in the spinal cord (sc), dorsal root ganglia (drg), sympathetic ganglia (sg), skeletal muscle (m), lung (lg), tongue (t), and primordium of teeth in the jaw (j). There is no expression of *YR-23* neither in ventricles (v) or atria (a) of the heart. **B:** Expression in the lobes of thymic rudiment (tr) and weak expression in the common carotid arteries (ca). **C:** Expression in the pectoral muscle (pm), the sublingual/submandibular glands (ssg), and the tongue (t). **D:** Expression in the duodenum (d) and weak expression in the liver (lv).

tions of wild-type mouse embryos as well as in whole embryos at different developmental stages. Both genes were expressed in the nervous system (Figs. 4, 5). *YR-23* mRNA exhibited widespread expression on transverse sections at E14.5 in various tissues such as spinal cord, DRGs, sympathetic ganglia, skeletal muscle, lung, primordium of teeth and nasal cavity but was not expressed in heart and cartilage (Fig. 4). At higher magnification, *YR-23* mRNA was detected in the thymic rudiments (Fig. 4A), the sublingual/submandibular gland and the tongue (Fig. 4B), and also in the hindgut (Fig. 4C).

For SNX14, in situ hybridization was performed on sections of mouse embryos between E9.5 and E14.5 and a peak of expression was observed at E12.5. As suggested by the in vitro expression profile of *SNX14*, on sagittal sections of E12.5 embryo, its in vivo expression was restricted mainly to neuronal lineages such as the spinal cord (Fig. 5). On transverse sections of the hindbrain at E12.5, *SNX14* was expressed in the ventral ventricular region and the floor plate as well as in cranial ganglia (Vth and VIIIth), pituitary, and eye (retina and lens; data not shown) comparable to the expression of *Islet-1* (Fig. 6A,B). In the spinal cord, *SNX14* was expressed at a low level mainly in motoneurons at the forelimb level and also weakly in DRGs (Fig. 6C) compared with *Islet-1* (Fig. 6D). Other sites of expression were observed such as the semicir-

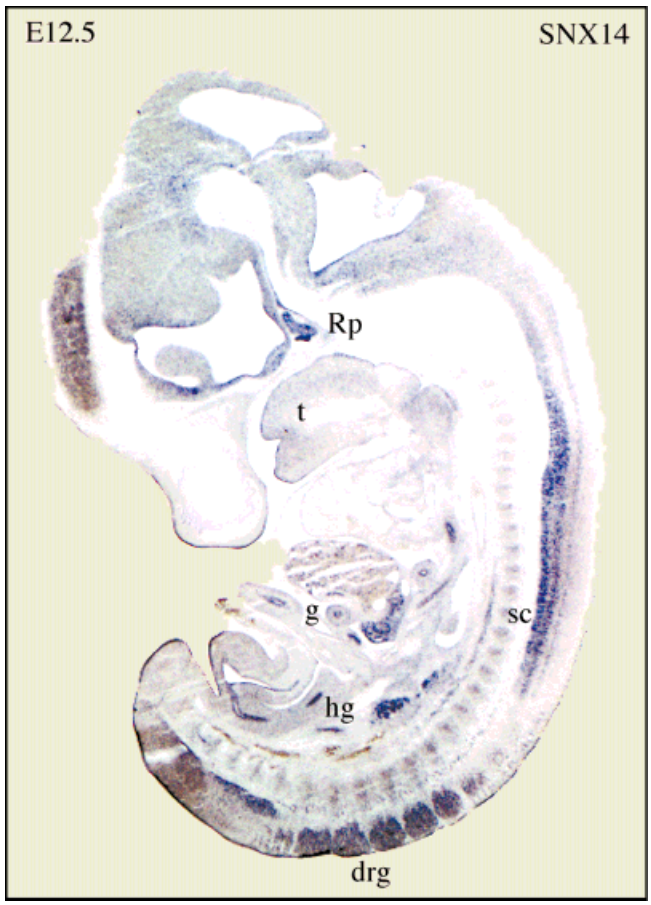

Fig. 5. Expression of *SNX14* in the mouse embryo. In situ hybridisation by using digoxigenin-labeled SNX14 probe on a sagittal section of embryonic day (E) 12.5 mouse embryo. Expression of *SNX14* is detected in the spinal cord (sc), Rathke's pouch (Rp), gut (g), dorsal root ganglia (drg), and lumen of hindgut (hg). There is no expression in muscles (see the tongue, t).

cular canal and the saccule of the inner ear (Fig. 6E). We were unable to detect by in situ hybridisation any signal in sympathetic ganglia, lung, skeletal and cardiac muscles, and liver.

In E12.5 embryos, *SNX14* was expressed in the gut (Fig. 6G) and more strongly in the developing anterior pituitary gland at two main sites: Rathke's pouch and particularly strongly in the developing pars distalis, which will generate the anterior pituitary lobe secreting several trophic hormones (Fig. 6H). *SNX14* was coexpressed with the homeodomain transcription factor Islet-1 in the pars distalis (data not shown for Islet-1).

DISCUSSION

The principle of gene trapping is based on the random integration into the genome of a promoterless reporter gene (for example *lacZ*) such that its expression is under the control of the endogenous trapped promoter (Skarnes et al., 1992). Selection of transfected ES cells requires the use of a selectable marker gene. In earlier screens, the selection gene was an independent transcriptional unit, under the control of a constitutive promoter and carrying its own polyA signal (Gossler et al., 1989); thus, the majority of integrations led to a resistant clone. To enrich for true gene trap events, the promoterless *lacZ* was fused to the neomycin resistance gene (neo^R) such that the selection of the resistant clones was dependent on the trapped gene being active in undifferentiated ES cells (Friedrich and Soriano, 1991). Recently, polyA addition traps have been introduced in vectors for gene trapping. In this case, selection is independent of the activity of the trapped gene, but the selection gene, lacking its own polyA signal, is linked by gene splicing to the 3′ exons and, thus, the polyA signal of the trapped gene (Niwa et al., 1993; Yoshida et al., 1995; Salminen et al., 1998; Zambrowicz et al., 1998). After selection, neo^RES cell clones are picked, expanded, and frozen as an in vitro library of ES cell clones, each one corresponding to a particular mutation. The generation of fusion transcripts by gene trapping allows the cloning of the endogenous flanking exons directly from ES cell cultures by using the RACE protocol (Frohman et al., 1988) without having to isolate any genomic sequence.

The major drawback of gene trapping, in terms of time and cost, is the generation of large number of mouse chimaeras by blastocyst injection or even transgenic lines to identify a particular restricted expression pattern. To circumvent this problem, in vitro preselection strategies were introduced to select integration events in subsets of genes: genes encoding secreted proteins (Skarnes et al., 1995), genes expressed in restricted cell lineages after differentiation of ES cells (Baker et al., 1997), or genes expressed in response to particular inducer molecules (Forrester et al., 1996; Bonaldo et al., 1998). In vitro differentiation of ES cells usually involves culturing them in aggregates called embryoid bodies (EBs), which differentiate and generate cells of the hematopoietic, endothelial, muscle, neuronal, and glial lineages (reviewed by Keller, 1995). Efficient generation of neurons from EBs requires the addition of retinoic acid (Bain et al., 1995; Strübing et al., 1995; Fraichard et al., 1995). We developed recently an in vitro differentiation system in which the ES-derived neurons produced are mainly of the CNS subtype and a fraction of them express homeoproteins characteristic of motoneurons and interneurons (Renoncourt et al., 1998).

Our gene and polyA trap vector, IRESnlsβgalNeo(-pA), is designed to produce expression of the selectable neomycin marker in ES cells only when the vector was integrated in the correct orientation within a gene locus. Roughly 8% (8 of 104) of the undifferentiated resistant colonies (excluding concatemers) obtained with our polyA trap vector expressed β-galactosidase. This number is in the same range as the proportions reported in previous studies in which different polyA trap vectors were used (7.7% in Niwa et al., 1993; 2.7% in

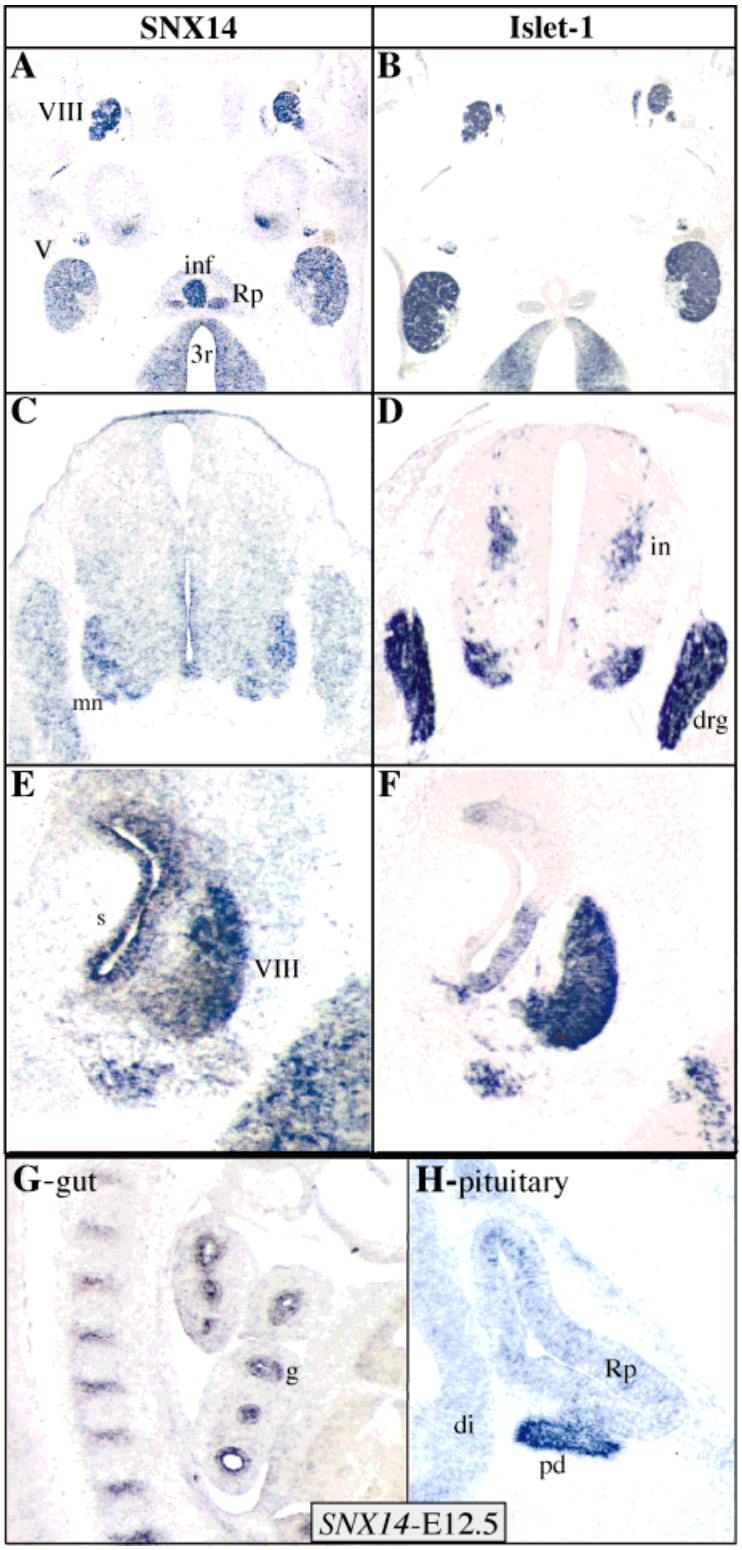

Fig. 6. Comparison of *SNX14* and *Islet-1* expression in the E12.5 mouse embryo. In situ hybridization by using SNX14 and Islet-1 digoxigenin-labeled probes on sections. Expression of *SNX14* (**A**) and *Islet-1* (**B**) on coronal sections in the head region. Coexpression of both markers is observed in the trigeminal (V) ganglia, the vestibulocochlear (VIII) ganglia, and the ventral part of the third ventricle (3r). In the pituitary gland, *SNX14* is expressed in the infundibulum (inf) and the Rathke's pouch (Rp) where *Islet-1* expression is weaker. In the spinal cord, *SNX14* is expressed at low levels mainly in lateral motoneurons (mn) at the forelimb level (**C**) but not in dorsal interneurons (in) as *Islet-1* (**D**). *SNX14* (**E**) and *Islet-1* (**F**) mRNAs colocalize in the saccule (s) of the inner ear and the vestibulocochlear (VIII) ganglia. On sagittal sections, *SNX14* is expressed in the epithelium of the gut (g) (**G**) and in the developing pituitary gland (**H**) mainly in the pars distalis (pd) when compared with the diencephalon (di) or the Rathke's pouch (Rp).

Yoshida et al., 1995; 11% in Salminen et al., 1998) and should correspond to genes transcribed in undifferentiated ES cells. The new generation of polyA trap vectors increases the proportion of true gene trap events in the bank of G418 neomycin resistant clones (see Niwa et al., 1993; Yoshida et al., 1995; Salminen et al., 1998).

One disadvantage, in our hands, of such a vector is the occurrence of concatemer-containing clones, never-

theless both concatemers and housekeeping genes could be eliminated by discarding blue undifferentiated ES colonies. All concatemer clones were β-gal+ at this stage, except three clones that turned β-gal+ only at the EB7 stage. This finding means that less than 3% (3 of 104) of neo^Rcolonies containing concatemers were not detected at the ES cell stage by our method. For these clones, we suppose that the PGK promoter may have been under the influence of a *cis*-acting negative regulatory element at the ES cell stage. One possibility for the avoidance of concatemer formation is to use a retroviral polyA trap vector, in which case the single integrations of the vector are favoured (Zambrowski et al., 1998). Importantly, for the two gene-trap clones that we examined in detail, we confirmed that both the splice acceptor and splice donor sites of our vector had been correctly utilised by the trapped gene.

Our interest was to test the feasibility of using an in vitro preselection step to identify trapped genes expressed in the nervous system. From our previous observation that neurons expressing the LIM-homeodomain protein Islet-1 are mainly generated during a short window of time between the 6th and 7th days of differentiation of retinoic acid treated embryoid bodies, we reasoned that screening for expression of β-galactosidase in embryoid bodies at day 7 of in vitro differentiation would allow us to enrich for gene-trap events in neuronally expressed genes. Thus, the library of trapped ES cell clones was screened first for clones expressing *lacZ* in EB7. In a second step, β-galactosidase-positive clones were screened for the presence of β-gal+ neurons after dissociation and plating cells from EB7. From four *lacZ*-expressing clones at EB7, two gave rise to β-gal+ neurons after plating, clone YR-23 which is expressed in a range of cell types including neurons and clone YR-14 (subsequently found to represent the mouse homologue of human sorting nexin 14) whose expression is mainly found in cells of the neuronal lineage. Thus, from 104 ES neo^Rclones, our in vitro preselection at EB7 allowed the identification of one clone whose expression was clearly restricted to neuronal subpopulations. This strong selection (50% of positive clones contained β-galactosidase–expressing neurons) for genes expressed in the neuronal lineage demonstrated the validity of choosing EB7 as the assay point. Nevertheless, this did not preclude the possibility that a fraction of 73 β-gal− EB7 represented neuronal trapped genes expressed later during EB development.

From the point of view of specificity of expression pattern, *sorting nexin 14* (*SNX14*) was the most interesting clone detected in the limited screen of 104 ES cell clones. *lacZ* expression after in vitro differentiation of this clone was restricted to neurons and included Islet-positive neurons. The highest expression in vivo of *SNX14* was in the pituitary gland at E12.5, but expression was also detected at lower levels in brachial motoneurons and other neurectoderm derivatives. In the developing pituitary gland, the site of strongest expression of *SNX14* was the pars distalis in the ventral-most region of the anterior pituitary which originates from an outpocketing of the roof of the embryonic pharynx, known as Rathke's pouch (RP). The developing pars distalis forms the anterior lobe secreting several trophic hormones. From the point of view of our gene trap strategy, it is striking that the pars distalis is a site of expression of *Islet-1* at E12.5 in the pituitary (Ericson et al., 1998; Treier et al., 1998). This result opens up the possibility that a fraction of Islet-positive cells generated in our in vitro differentiation system may, in fact, be pituitary endocrine cells. A similar correspondence in the site of expression of *SNX14* and *Islet* is observed in the ventral spinal cord, where *Islet* is expressed in motoneurons and is essential for their generation and columnar determination. Colocalization of *SNX14* and *Islet-1* mRNAs was observed in several other tissues including, cranial and spinal sensory ganglia, the semicircular canal, and the saccule of the inner ear, retina, and gut. The expression of *SNX14* and *Islet* both in the pituitary gland and motoneurons is interesting since several other markers have also been shown to be coexpressed in both tissues. In fact, strong parallels exist in the sites of expression of patterning molecules such as Sonic Hedgehog (Shh) and BMP-4 and transcription factors such as Pax-6, Islet-1, Lhx3, and Lhx4, during the development of the pituitary gland and spinal motoneurons (Ericson et al., 1998; Treier et al., 1998).

Sorting nexins represent a family of at least 14 members and have been shown to associate with the intracellular part of members of several families of cell surface receptors, including receptor tyrosine kinases and cytokine receptors (Kurten et al., 1996; Haft et al., 1998). These molecules may be involved in recycling and targeting of receptors for degradation. For example, overexpression of *SNX1* led to accelerated degradation of EGF receptors (Kurten et al., 1996). There is also some evidence that different sorting nexins may have differential affinity for particular receptors (Haft et al., 1998); thus, differential expression of sorting nexins might reflect a role of specific isoforms in regulating the availability of certain receptor types in particular cells.

Our long-term aim is to develop the system described here to preselect gene-trap clones expressed in specific subpopulations of neurons. Gene trap screens generally concentrate on expression pattern as a clue to potential gene function. Trapping strategies were combined with in vitro preselection to select for subsets of genes encoding, for example, genes responsive to specific growth or differentiation factors (Forrester et al., 1996; Baker et al., 1997; Bonaldo et al., 1998), genes expressed in the embryonic nervous system (Shirai et al., 1996, this study), and also genes induced in apoptotic cells by using the site-specific recombination Cre/loxP (Russ et al., 1996). Another innovative use of gene-trapping was the screen devised by Skarnes et al. (1995), in which genes encoding secreted proteins were specifically targeted such as *netrin-1* (Serafini et al.,

1996). At least in the case of screens involving identification of genes expressed in subclasses of cells after in vitro differentiation of gene-trapped ES clones, our data and other reports have confirmed that the profiles of *lacZ* expression obtained in vitro allows a reasonable prediction of the tissues in which the gene is expressed in vivo (Baker et al., 1997; Xiong et al., 1998). The future development of these types of screens will involve technical modifications that improve the efficiency of the in vitro differentiation systems to allow high throughput screening. Recently, an in vitro differentiation protocol was described that does not require the rather laborious step of EB formation and retinoic acid induction (Kawasaki et al., 2000). Such systems will dramatically improve the potential for high-throughout screening. Also of potential interest in this respect is the isolation, from cultures of in vitro differentiated ES cells, of cells showing characteristics of neuronal precursor cells (Okabe et al., 1996).

EXPERIMENTAL PROCEDURES

Construction of the "Gene and polyA Trap" Vector

The vector was constructed by using the gene-trap vector described by Yoshida et al. (1995) as a backbone. The pPAT vector, kindly provided by Dr. Yoshida (Tokyo), was modified by exchanging the cytoplasmic form of the *lacZ* gene with the nuclear localized form (nls), introduction of an IRES sequence upstream of the β-galactosidase coding sequence by replacing the fyn intron and splice acceptor site in pPAT with the EN-2 intron, splice acceptor site, and IRES sequence as used in the gene trap vector GT1.8IRESβgeo described in Skarnes et al. (1995), vector kindly provided by Dr. A. Smith (Edinburgh).

The cloning steps were as follows: a 2-kb *Eco*RI fragment from the vector GT1.8IRESβgeo containing the EN-2 intron and splice acceptor site followed by an IRES sequence was cloned into the vector ETL/LTNL (P. Mombaerts, New York) that contains the tau-β-galactosidase fusion protein coding sequence. The sequence encoding the N-terminal part of the tau-β-galactosidase fusion was replaced by the N-terminal part of the nlsβ-gal coding sequence by exchanging a 2-kb *NcoI*-ClaI fragment with a 1-kb *NcoI*-ClaI fragment from the vector pSKT-nls-lacZ (S. Tajbakhsh, Paris). Finally, a 3-kb *XhoI*-ClaI fragment from the above plasmid containing the En-2 intron and splice acceptor site, IRES sequence, and N-terminal part of the nlsβ-gal coding sequence was exchanged for a *SalI*-ClaI fragment in the vector pPAT. The plasmid was called IRESnlsβgalNeo(-pA) (Fig. 1) and linearized by digestion with the restriction enzyme *Pac*I before electroporation into ES cells.

ES Cell Culture and Electroporation of the Vector

The ES cell line CCE (Robertson et al., 1986) was grown in standard medium containing leukemia inhibitory factor (Gibco-Esgro LIF) at 1,000 U/ml, and cells were harvested in the exponential phase (see Renoncourt et al., 1998). A cell suspension of 10^6 cells was electroporated with 2.5 μg of linearized IRESnlsβgalNeo(-pA) plasmid in a volume of 0.8 ml of cold PBS. Cells were plated at a density of 2.5×10^5 per 10-cm tissue culture dish in standard medium, and 1 day later, electroporated cells were selected on medium containing the antibiotic G418 (300 μg/ml) for 7 to 10 days. Resistant colonies were picked under a microscope put into a hood and transferred to a 96-multiwell dish containing 25 μl of 0,025% (0.5×) trypsin-EDTA solution in which cells disaggregate rapidly at room temperature. Medium (75 μl) was added and cells were dissociated by gentle pipetting. The mixture was seeded onto triplicate 96-well dishes in medium containing G418. One dish was used to freeze ES clones in 10% DMSO and was kept at −80°C, the second was used for X-gal staining of undifferentiated ES cells, and the third to carry out in vitro differentiation experiments. After 3–4 days, the clones were ready to be frozen, stained, or differentiated.

In Vitro Differentiation of Embryoid Bodies and X-Gal Staining of EB7

Undifferentiated ES clones were grown on gelatin-coated tissue culture dishes in standard ES cell culture medium containing LIF (day 0) and were then cultured in hanging drops during 2 days on nonadhesive bacteriological Petri dishes in which ES cells aggregate to form embryoid bodies (EBs). At day 2, EBs were transferred to a nonadhesive Petri dish in culture medium containing 2 μM retinoic acid (RA) to promote neuronal differentiation according to our protocol (Renoncourt et al., 1998).

EBs were collected at day 7 (EB7) of culture, washed in PBS, fixed with 1% formaldehyde at 4°C for 15 min. Fixed EB7 were washed with PBS and stained in X-gal solution.

Antibodies and Immunocytochemistry

EB7 were collected and dissociated following a protocol found to minimize cell death. Dissociated cells plated on 3.5-cm polyornithine/laminin–coated dishes were washed in basal F12 medium and fixed with 4% paraformaldehyde at 4°C for 15 min and processed for immunocytochemistry with a set of primary antibodies (protocol described in Renoncourt et al., 1998).

The following primary antibodies were used: rabbit anti–NF-145 (1:500, Chemicon), rabbit anti-peripherin (1:200, a gift from M.M. Portier), rabbit anti-Phox2b (1:1,000, a gift from C. Goridis), monoclonal anti-HB9 (1:20, a gift from T. Jessell), monoclonal anti-Islet1/2 was a mixture of two antibodies from Developmental Studies Hybridoma Bank, 4D5 (1:100) and 2D6 (1:2), rabbit anti–β-galactosidase (1:1,000, a gift from J. Sanes), and monoclonal anti-β-galactosidase (1:1,000, Promega, #Z3781). We used the following secondary

antibodies: donkey anti-rabbit FITC-conjugated (1:50), biotin-SP–conjugated donkey anti-rabbit IgG (1:1,000, Jackson ImmunoResearch Lab.), Cy2-conjugated streptavidin at 1:1,000 (Amersham), and Cy3-conjugated donkey anti-mouse at 1:400 (Jackson ImmunoResearch Laboratory). For double immunostaining, control cultures were performed by omitting primary or secondary antibodies showing no specific labeling on cells.

Cloning the cDNAs of Trapped Genes by 5′- and 3′-RACE

The gene and polyA trap vector IRESnlsβgalNeo(-pA) containing 5′-splice acceptor site (5′-SA) and 3′-splice donor site (3′-SD) allows cloning of the fusion transcripts taking advantage of the RACE protocol (Frohman et al., 1988). Total RNA was extracted either from undifferentiated ES cells to perform 3′-RACE or from EB7 to perform 5′-RACE (Chomczynski and Sacchi, 1987). A series of nested antisense oligonucleotide primers from the region in the gene-trap vector downstream of the 5′-SA were used to perform 5′-RACE by using a 5′-RACE kit according to the manufacturer's instructions (GibcoBRL, nb 18374-058). The sequences of the gene trap vector-specific nested primers were as follows: SSP1 : 5′-ACCTGTTGGTCTGAAACTCAGCC; SSP2 : 5′-CCGCTTGTCCTCTTTGTTAGG. Reverse transcription was performed on total RNA from ES cells by using an oligo-dT primer and the Superscript II reverse transcriptase (Gibco/BRL). 3′-RACE was carried out by using a series of sense primers based on the sequence of the 3′-end of the *neo*^{R}gene: neo-sense-1, TGAAGAGCTTGGCGGCGAAT; neo-sense-2, CGTTGGCTACCCGTGATATTGCT; oligo-dT primer, TCTAGAATTCAGCGGCCGC$(T)_{30}$VN. After nested amplifications, Taq-amplified polymerase chain reaction products were cloned into the pCR-2.1 plasmid (Invitrogen) and transformed into DH5α competent cells (GibcoBRL). Several colonies were picked for each ES cell line, and inserts were sequenced and compared with sequences from the GenBank databases.

Sequencing and Search for Homologues in Sequence Databases

Sequencing was carried out on an Automated Sequencer either at the Genome Express or INSERM U.491 sequencing facilities (France). The sequence databases at NCBI were searched by using the BLAST programs (Altschul et al., 1997).

In Situ Hybridization on Mouse Embryo Sections

In situ hybridization was performed on 16-μm-thick frozen transverse and sagittal sections prepared from E12.5 mouse embryos fixed with 4% paraformaldehyde in PBS buffer and cryopreserved in 15% sucrose, 0.12 M phosphate buffer, pH 7.2, before embedding in OCT (Schaeren-Wiemers and Gerfin-Moser, 1993). Antisense digoxigenin (DIG) -labeled riboprobes for Islet-1, SNX14, and YR-23 were produced from respective cDNA pCR2.1 plasmids (Invitrogen) by using a DIG-RNA labeling kit (Boehringer, Nb 1277 073). Probes were hybridized on sections at 70°C overnight at the following dilutions: Islet-1 (1:1,000), SNX14 and YR-23 (1:400), revealed with anti-DIG-alkaline-phosphatase(AP) –conjugated antibody (1:2,000, from Boehringer, Nb 1093 274), and sections were stained with a NBT/BCIP solution (Sigma B-1911). Slides were washed in TE and water before mounting with Aquamount (BDH 36226), and photomicrographs of slides were taken on an Axiophot microscope.

ACKNOWLEDGMENTS

We thank members of INSERM Unité 382 for many helpful discussions and especially C. Henderson (INSERM U.382) and J. Sanes (St. Louis, MO) for comments on the manuscript. We thank N. Kebaili for her help in ISH experiments. We thank also S. Tajbakhsh (Institut Pasteur, Paris) for the gift of the pSKT-nls-lacZ vector, M. Yoshida (Kumamoto, Japan) for the gift of the pPAT vector, A. Smith (Edinburgh, UK) for the gift of the GT1.8-IRES-βgeo vector, and P. Mombaerts (NY) for the gift of the ETL/LTNL vector. We acknowledge generous gifts of antibodies from T.M. Jessell (anti-HB9), M.M. Portier (anti-peripherin), C. Goridis (anti-Phox2b), and J. Sanes (anti–β-gal). The two monoclonal anti-Islet antibodies (39.4D5 and 40.2D6) developed by T.M. Jessell were obtained from the Developmental Studies Hybridoma Bank maintained by the Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, and the Department of Biological Sciences, University of Iowa, Iowa City, IA 52242, under contract NO1-HD-2-3144 from the NICHD. Our laboratory is supported by INSERM, CNRS, Association Française contre les Myopathies (AFM), and by the European Community (EC Biotechnology grant BIO4-CT96-0433). P.C. was supported by an EC fellowship from the TMR Work Programme and by a bourse from the Féderation pour la Recherche Médicale (FRM).

REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.

Bain G, Kitchens D, Yao M, Huettner JE, Gottlieb DI. 1995. Embryonic stem cells express neuronal properties in vitro. Dev Biol 168: 342–357.

Baker RK, Haendel MA, Swanson BJ, Shambaugh JC, Micales BK, Lyons GE. 1997. In vitro preselection of gene-trapped embryonic stem cell clones for characterizing novel developmentally regulated genes in the mouse. Dev Biol 185:201–214.

Bonaldo MF, Lennon G, Soares MB 1996 Normalization and subtraction: two approaches to facilitate gene discovery. Genome Res 6:791–806.

Bonaldo P, Chowdhury K, Stoykova A, Torres M, Gruss P. 1998. Efficient gene trap screening for novel developmental genes using IRES beta geo vector and in vitro preselection. Exp Cell Res 244: 125–136.

Chomczynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. Anal Biochem 162:156–159.

Djabali K, Portier MM, Gros F, Blobel G, Georgatos SD. 1991. Network antibodies identify nuclear lamin B as a physiological attachment site for peripherin intermediate filaments. Cell 64:109–121.

Ericson J, Norlin S, Jessell TM, Edlund T. 1998. Integrated FGF and BMP signaling controls the progression of progenitor cell differentiation and the emergence of pattern in the embryonic anterior pituitary. Development 125:1005–1015.

Evans MJ. 1998. Gene trapping—a preface. Dev Dyn 212:167–169.

Evans MJ, Carlton MB, Russ AP. 1997. Gene trapping and functional genomics. Trends Genet 13:370–374.

Forrester LM, Nagy A, Sam M, Watt A, Stevenson L, Bernstein A, Joyner AL, Wurst W. 1996. An induction gene trap screen in embryonic stem cells: identification of genes that respond to retinoic acid in vitro. Proc Natl Acad Sci USA 93:1677–1682.

Fraichard A, Chassande O, Bilbaut G, Dehay C, Savatier P, Samarut J. 1995. In vitro differentiation of embryonic stem cells into glial cells and functional neurons J Cell Sci 108:3181–3188.

Friedrich G, Soriano P. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev 5:1513–1523.

Frohman MA, Dush MK, Martin GR. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer Proc Natl Acad Sci USA 85: 8998–9002.

Gossler A, Joyner AL, Rossant J, Skarnes WC. 1989. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science 244:463–465.

Haft CR, de la Luz Sierra M, Barr VA, Haft DH, Taylor SI. 1998. Identification of a family of sorting nexin molecules and characterization of their association with receptors. Mol Cell Biol 18:7278–7287.

Harrison KA, Druey KM, Deguchi Y, Tuscano JM, Kehrl JH. 1994. A novel human homeobox gene distantly related to proboscipedia is expressed in lymphoid and pancreatic tissues. J Biol Chem 269: 19968–19975.

Kawasaki H, Mizuseki K, Nishikawa S, Kaneko S, Kuwana Y, Nakanishi S, Nishikawa SI, Sasai Y. 2000. Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. Neuron 28:31–40.

Keller GM. 1995. In vitro differentiation of embryonic stem cells. Curr Opin Cell Biol 7:862–869.

Kurten RC, Cadena DL, Gill GN. 1996. Enhanced degradation of EGF receptors by a sorting nexin, SNX1. Science 272:1008–1010.

Nagase T, Seki N, Ishikawa K, Ohira M, Kawarabayasi Y, Ohara O, Tanaka A, Kotani H, Miyajima N, Nomura N. 1996. Prediction of the coding sequences of unidentified human genes: VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain. DNA Res 3:321–329; 341–354.

Niwa H, Araki K, Kimura S, Taniguchi S, Wakasugi S, Yamamura K. 1993. An efficient gene-trap method using poly A trap vectors and characterization of gene-trap events. J Biochem (Tokyo) 113:343–349.

Okabe S, Forsberg-Nilsson K, Spiro AC, Segal M, McKay RD. 1996. Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. Mech Dev 59:89–102.

Pattyn A, Morin X, Cremer H, Goridis C, Brunet JF. 1997. Expression and interactions of the two closely related homeobox genes Phox2a and Phox2b during neurogenesis. Development 124:4065–4075.

Renoncourt Y, Carroll P, Filippi P, Arce V, Alonso S. 1998. Neurons derived in vitro from ES cells express homeoproteins characteristic of motoneurons and interneurons. Mech Dev 79:185–197.

Robertson E, Bradley A, Kuehn M, Evans M. 1986. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature 323:445–448.

Russ AP, Friedel C, Ballas K, Kalina U, Zahn D, Strebhardt K, von Melchner H. 1996. Identification of genes induced by factor deprivation in hematopoietic cells undergoing apoptosis using gene-trap mutagenesis and site-specific recombination. Proc Natl Acad Sci USA 93:15279–15284.

Salminen M, Meyer BI, Gruss P. 1998. Efficient poly A trap approach allows the capture of genes specifically active in differentiated embryonic stem cells and in mouse embryos. Dev Dyn 212:326–333.

Schaeren-Wiemers N, Gerfin-Moser A. 1993. A single protocol to detect transcripts of various types and expression levels in neural tissue and cultured cells: in situ hybridization using digoxigenin-labelled cRNA probes. Histochemistry 100:431–440.

Schultz J, Milpetz F, Bork P, Ponting CP. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci USA 95:5857–5864.

Serafini T, Colamarino SA, Leonardo ED, Wang H, Beddington R, Skarnes WC, Tessier-Lavigne M. 1996. Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system. Cell 87:1001–1014.

Shirai M, Miyashita A, Ishii N, Itoh Y, Satokata I, Watanabe YG, Kuwano R. 1996. A gene trap strategy for identifying the gene expressed in the embryonic nervous system. Zool Sci 13:277–283.

Skarnes WC, Auerbach BA, Joyner AL. 1992. A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. Genes Dev 6:903–918.

Skarnes WC, Moss JE, Hurtley SM, Beddington RS. 1995. Capturing genes encoding membrane and secreted proteins important for mouse development. Proc Natl Acad Sci USA 92:6592–6596.

Strübing C, Ahnert-Hilger G, Shan J, Wiedenmann B, Hescheler J, Wobus AM. 1995. Differentiation of pluripotent embryonic stem cells into the neuronal lineage in vitro gives rise to mature inhibitory and excitatory neurons. Mech Dev 53:275–287.

Tanabe Y, William C, Jessell TM. 1998. Specification of motor neuron identity by the MNR2 homeodomain protein. Cell 95:67–80.

Treier M, Gleiberman AS, O'Connell SM, Szeto DP, McMahon JA, McMahon AP, Rosenfeld MG. 1998. Multistep signaling requirements for pituitary organogenesis in vivo. Genes Dev 12:1691–1704.

Wurst W, Rossant J, Prideaux V, Kownacka M, Joyner A, Hill DP, Guillemot F, Gasca S, Cado D, Auerbach A, et al. 1995. A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice. Genetics 139:889–899.

Xiong JW, Battaglino R, Leahy A, Stuhlmann H. 1998. Large-scale screening for developmental genes in embryonic stem cells and embryoid bodies using retroviral entrapment vectors. Dev Dyn 212: 181–197.

Yoshida M, Yagi T, Furuta Y, Takayanagi K, Kominami R, Takeda N, Tokunaga T, Chiba J, Ikawa Y, Aizawa S. 1995. A new strategy of gene trapping in ES cells using 3′RACE. Transgenic Res 4:277–287.

Zambrowicz BP, Friedrich GA. 1998. Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse. Int J Dev Biol 42:1025–1036.

Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C, Sands AT. 1998. Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature 392:608–611.