

A gene trap approach in *Xenopus*

Odile J. Bronchain, Katharine O. Hartley and Enrique Amaya

The frog transgenesis technique ultimately promises to make mutagenesis possible through random insertion of plasmid DNA into the genome. This study was undertaken to evaluate whether a gene trap approach combined with transgenesis would be appropriate for performing insertional mutagenesis in *Xenopus* embryos. Firstly, we confirmed that the transgenic technique results in stable integration into the genome and that transmission through the germline occurs in the expected Mendelian fashion. Secondly, we developed several gene trap vectors, using the green fluorescent protein (GFP) as a marker. Using these vectors, we trapped several genes in *Xenopus laevis* that are expressed in a spatially restricted manner, including expression in the epiphysis, the olfactory bulb and placodes, the eyes, ear, brain, muscles, tail and intestine. Finally, we cloned one of the trapped genes using 5' rapid amplification of cDNA ends polymerase chain reaction (RACE PCR). These results suggest that the transgenic technique combined with a gene trap approach might provide a powerful method for generating mutations in endogenous genes in *Xenopus*.

Address: Wellcome/CRC Institute, Department of Zoology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK.

Correspondence: Enrique Amaya
E-mail: ea3@mole.bio.cam.ac.uk

Received: 17 August 1999
Revised: 7 September 1999
Accepted: 7 September 1999

Published: 11 October 1999

Current Biology 1999, 9:1195–1198

0960-9822/99/\$ – see front matter
© 1999 Elsevier Science Ltd. All rights reserved.

Results and discussion

In this study, we wanted to determine whether the *Xenopus* transgenesis technique [1] could be adapted for insertional mutagenesis. Transgenesis in the frog is based on restriction enzyme mediated integration (REMI) and allows the generation of transgenic frog embryos expressing reporter constructs at a high frequency and in a non-mosaic fashion [1]. Because plasmid DNA is randomly inserted into the genome, this technique is probably mutagenic [2]. If the transgenic technique is to be useful for insertional mutagenesis, however, the integrated DNA must be stably transmitted through the germline.

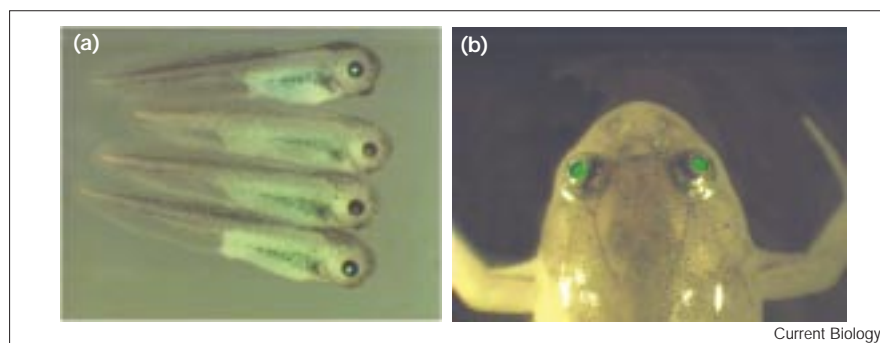
To address this issue, we have generated transgenic embryos carrying a γ 1crystallin–GFP construct, which is specifically expressed in the lens, and raised them to

maturity. The transgenic frogs were crossed to wild-type counterparts and the F1 progeny was analyzed for the expression of the γ 1crystallin–GFP construct (Figure 1; Table 1). Four independent lines were crossed and all transmitted the transgene. In lines 1 and 3, 50% of the F1 progeny expressed the transgene, suggesting that one integration event occurred during the transgenic procedure. F1 progeny derived from line 2 expressed the transgene at 75%, suggesting two independent integrations. The frequency of transmission obtained from these lines strongly suggests that the integration event in these founder animals occurred at a very early stage of development, resulting in a non-mosaic germline. In line 4, 89% of the F1 progeny expressed the transgene, suggesting three or four independent integrations. We are currently raising the GFP-expressing F1 animals from line 1 and the metamorphosed frogs still express the transgene (Figure 1b). These results confirm that the amphibian transgenesis procedure results in early, stable integration of plasmid DNA into the genome, that the germline in the F0 founder animals is not mosaic and that the transgenesis procedure is suitable for multi-generation experiments. In addition, the ability to generate more than a thousand transgenic animals following a simple mating procedure makes this technology in the frog particularly powerful (see Table 1).

Next, we investigated whether a gene trap approach would be feasible in *Xenopus*. A typical gene trap vector contains a selectable marker either fused in frame or placed on an independent translational unit to a selectable marker. Because it is devoid of promoter sequences, an insertion into a transcriptionally active unit is required for the marker to be expressed [3–5]. In the mouse, a selection step occurs following transfection of embryonic stem cells. Rather than utilizing a selection step, we have modified the existing gene trap vectors to allow for rapid screening of insertions in living embryos. We replaced the commonly used β -galactosidase reporter gene with the GFP gene [6–7], the expression of which can be detected in living embryos at all stages of development following the start of zygotic transcription. Integration of a gene trap vector containing GFP into an actively transcribed gene acts as a locus-specific marker of the gene in living embryos and provides a tag for identifying the disrupted gene.

We generated a variety of gene trap vectors and tested them in transgenic *X. laevis* embryos. The basic exon trap (ET) vector contains little more than the GFP coding sequence. This vector must integrate into exons of actively transcribed genes in the correct orientation (and reading frame, if within the translated region) for GFP fluorescence to be recovered. To allow detection of integrations

Figure 1



Germline transmission of the γ 1crystallin-GFP transgene in *X. laevis*. (a) Image of four F1 stage 30 embryos from transgenic line 1. Note that two out of the four embryos express GFP in the lens. (b) Metamorphosed F1 froglet expressing GFP in the lens.

into introns, we have added splice acceptor (SA) sequences upstream of GFP; these SA sequences were from the adenovirus late major transcript [5] (the SAGT vector) or the murine *engrailed 2* gene [8] (the SEGT vector). Because we wanted to screen more specifically for insertions into translated genes, we mutated the GFP initiation codon in the SEGT vector (SE Δ GT) so that GFP translation would occur only when fused in frame within the translated region of a trapped gene product.

Our screening strategy relies on generating fusion transcripts between endogenous genes and GFP, in some instances resulting in GFP fusion proteins. For a gene trap approach to be generally useful, therefore, it is essential that GFP fluorescence will not be adversely affected when randomly fused to other proteins. To avoid potential misfolding problems, we inserted a stretch of glycine residues upstream of the GFP coding sequence. This bridge would be expected to provide a flexible hinge between GFP and upstream sequences, allowing the formation of independent protein modules.

To determine the efficacy of a gene trap approach in *Xenopus*, we generated several hundred transgenic

embryos with each of our gene trap vectors. The embryos were screened for GFP expression from the gastrula to tadpole stages. Although most embryos failed to express GFP, a small percentage of them (between 0.5% and 3%) expressed GFP either ubiquitously or in a spatially restricted pattern. In general, a higher percentage of embryos generated with the SAGT and SEGT vectors expressed GFP than embryos generated with ET or SE Δ GT. Figures 2 and 3 show GFP expression in different gene trap embryos. Embryos in Figure 2a,b and Figure 3d were generated using the SAGT vector. SA1 expressed GFP ubiquitously (Figure 2a and data not shown). Figure 2a shows a close-up view of the expression pattern of this gene trap insertion in the head. SA2 expressed GFP in the tail somites (Figure 2b) and nervous system (data not shown). SA3 expressed GFP weakly throughout the brain (Figure 3d). Figure 3b shows the expression pattern of a tadpole (ET1) generated with the exon trap vector. ET1 expressed GFP exclusively throughout the olfactory system, including the placodes, bulb and nerves. SE1 (Figure 3c), a tadpole generated with SEGT, showed strong expression in the pineal gland and weaker expression in the midbrain-hindbrain border. Finally, Figure 2c-e shows the expression pattern of three embryos (SE Δ 1-3) generated with SE Δ GT. SE Δ 1 showed fluorescence in the inner ear, most probably the otoconia (Figure 2c). SE Δ 2 expressed GFP in nerves flanking the spinal cord (Figure 2d), and SE Δ 3 expressed GFP specifically in the coiled intestine (Figure 2e).

Table 1

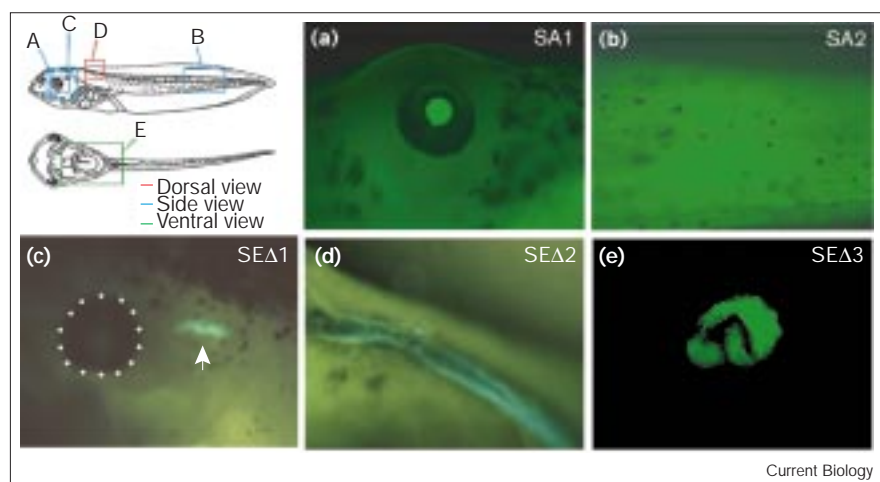
GFP expression in γ 1crystallin-GFP F1 embryos.

	GFP+	GFP-	% of GFP+ embryos	Number of integrations
Line 1	1412	1379	50.7	1
Line 2	812	266	75.5	2
Line 3	844	856	49.6	1
Line 4	2179	269	89.0	3-4

Four independent transgenic F0 founder animals were mated to wild-type counterparts and the F1 progeny were assayed for GFP expression at stage 30. Positive (GFP+) and negative (GFP-) embryos were scored. The percentage of expressing embryos is shown and the number of independent integrations was based on this percentage as predicted by simple Mendelian segregation.

A significant advantage of an insertional mutagenesis scheme is the relative ease of identifying mutant alleles. A gene trap approach has the added advantage that trapped genes can be identified from isolated mRNA, rather than genomic DNA. Even though an embryo may contain multiple insertions (Table 1), therefore, only ones that are actively transcribed will be identified. To determine whether we could clone some of the trapped genes, we isolated mRNA from individual gene trap embryos and performed 5' RACE PCR using nested primers within GFP. We have cloned and analyzed several 5' RACE-PCR

Figure 2

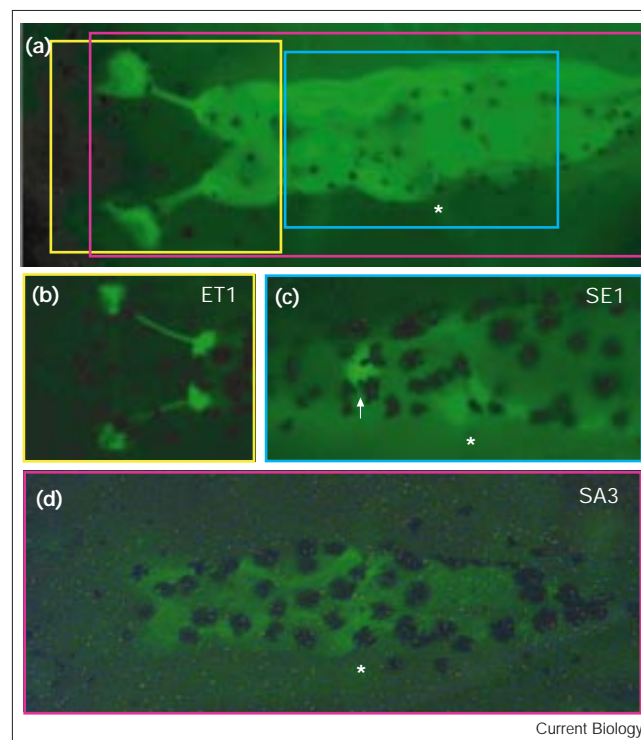


Gene trap embryos expressing GFP in different tissues. The top left panel shows a side view and ventral view of two tadpoles. The colored boxes outline the regions and orientation used to take each picture. (a,b) SA1 and SA2 were generated using the SAGT vector. GFP expression in SA1 was ubiquitous. (a) A magnified view of the head from SA1. Note strong expression in the lens. (b) View of the tail from SA2. Note strong expression in the somites. (c–e) SE Δ 1, 2 and 3 were three separate embryos containing the SE Δ GT gene trap vector. Each embryo expressed GFP in a distinct manner. (c) SE Δ 1 expressed GFP in the inner ear (otoconia; arrow), (d) SE Δ 2 expressed GFP in neurons flanking the spinal cord and (e) SE Δ 3 expressed GFP in the coiled intestine. The white dots outline the eye in (c).

products. Two of them were isolated from embryos generated with the SEGT vector. In both cases, we found that the *engrailed 2* splice acceptor is used properly to create new GFP fusion transcripts (Figure 4). The two 5' RACE products showed homology to previously identified cDNAs in the database. One showed 100% identity to the 5' end of the *Xenopus RelB* gene encompassing the 5' UTR and part of the coding region [9] (Figure 4b). The junction between *RelB* and *GFP* maintained the correct reading frame, resulting in a *RelB*–*GFP* fusion product. It is probable, therefore, that the insertion has interrupted the normal coding sequence of the *RelB* gene. Another product that we cloned was 100% homologous to the mouse *Mym* gene (M. Sam and A. Bernstein, unpublished observations; GenBank accession number AF019615). We later identified this sequence within the mouse *Engrailed 2* intron present in the SEGT and SE Δ GT vectors, so we conclude that mouse *Mym* is a cryptic exon within the second intron of the mouse *Engrailed 2* gene. The two cases in which we cloned upstream sequences from embryos generated with the SAGT vector, sequence analysis showed that the expected adenovirus splice acceptor site was not used (data not shown). Instead, the splice junction occurred around 10 bp upstream of the GFP initiation codon, where a cryptic splice acceptor is present. At least in *Xenopus*, therefore, gene trap vectors containing *engrailed* splice acceptor sequences might be preferable to those containing adenovirus sequences.

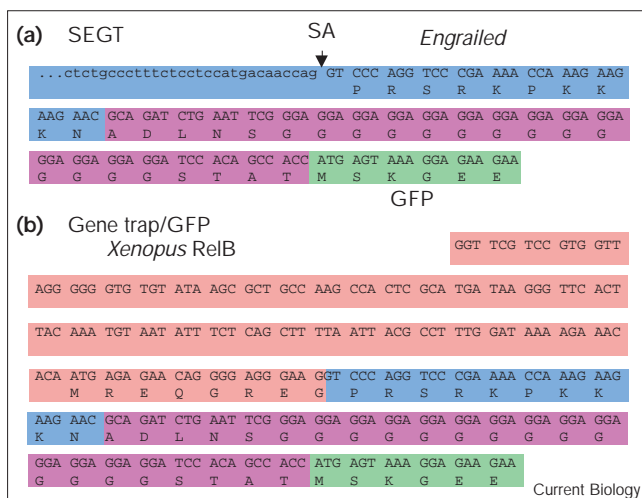
In conclusion, we have shown that the transgenesis technique can be used to generate stable transgenic lines in *X. laevis* and that the germline of the F0 founder animals is not mosaic. In addition, we show that the technique can be adapted for insertional mutagenesis using a gene trap strategy. Using this approach, we have trapped genes expressed in many tissue types and cloned some of the trapped genes using 5' RACE PCR. Indeed, we believe

Figure 3



Gene trap embryos expressing the GFP within the brain. (a) Dorsal view of the head of a transgenic embryo containing a τ –GFP fusion under the control of the neural β -tubulin promoter (NBT– τ GFP). GFP expression is observed throughout the brain, including the cranial nerves. This embryo serves as a reference for positioning the expression of the different gene trap embryos along the anterior–posterior axis. The colored boxes outline the regions of the head photographed in (b, yellow), (c, blue) and (d, red). (b) ET1 was generated using the exon trap vector ET. GFP fluorescence was restricted to the olfactory placodes, nerves and bulbs. (c) SE1 was generated using the SEGT gene trap vector and expressed GFP in the pineal gland (arrow) and the midbrain–hindbrain junction (asterisk in panels (a), (c) and (d)). (d) SA2 was generated using the SAGT gene trap vector. Expression was observed weakly throughout the brain.

Figure 4



Cloned gene trap sequences. **(a)** Nucleotide and predicted amino acid sequences of the *Engrailed 2/gly-GFP* junction in the SEGT gene trap vector. A vertical arrow indicates the splice site. The *Engrailed 2* SA sequence is shown in blue with the intron in lowercase letters and the exon in uppercase letters. The glycine bridge is shown in purple and GFP in green. **(b)** Sequence obtained from a trapped gene obtained using a 5' RACE PCR approach (shown in pink). The predicted amino acid sequence is indicated underneath the nucleotide sequence.

the *Xenopus* system is particularly suited for a gene trap approach to insertional mutagenesis. Firstly, the transgenic technique can be used to produce several hundred transgenic embryos in a few hours [1]; therefore, generating embryos containing trapped genes requires a very modest investment of time. Secondly, given that the embryos develop externally, expression of the marker gene GFP can be monitored in living embryos at any stage of development, in any tissue. In fact, a screen can be tailored such that only embryos expressing the marker gene at a particular stage or tissue type are maintained and raised to maturity. This will greatly decrease the need to maintain a large colony, as one can select from the start which embryos to raise. Thirdly, when establishing families of lines, one can easily recognize animals carrying the integration (and putative mutation). Heterozygous animals can be selected by expression of the GFP marker and only these are raised at each generation. The ability to generate several thousand embryos from a simple mating makes this system particularly powerful for genetic purposes, including careful analysis of phenotypes, gene mapping, and segregation analysis.

We should also mention two major disadvantages of using *X. laevis* for genetic experiments, however. The generation time is long (around 1–2 years), and *X. laevis* is essentially tetraploid [10–12]. Mutations in some developmental loci might not, therefore, result in obvious phenotypes because of functional redundancy in closely related paralogs. Given

that teleosts also underwent an additional genome duplication relative to most other vertebrates [13–15], this concern also holds true for zebrafish. Fortunately, there is a closely related diploid frog in the same genus as *X. laevis*. *Xenopus tropicalis* has all the advantages of its larger cousin, including our ability to make transgenic lines [12], but has the additional advantages that it is diploid and has a generation time of around five months [10–12]. In the future we plan to perform a large-scale gene trap screen in *X. tropicalis*.

Supplementary material

Supplementary material including an additional figure showing a diagram of the different gene trap vectors and Materials and methods are available at <http://current-biology.com/supmat/supmatin.htm>.

Acknowledgements

We thank John Dixon and Gordon Stott for helpful discussions, and Mark Carlton and Aaron Zorn for comments on the manuscript. We thank Rob Grainger for providing us with the *Xenopus* γ -crystallin promoter region. This work was supported by a category 20 fellowship from the European Community (OJB), two MRC studentships (O.J.B., K.O.H.), an NIH grant (RR13221) (E.A.) and a Wellcome Trust Senior Research Fellowship (E.A.)

References

- Kroll KL, Amaya E: Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. *Development* 1996, 122:3173-3183.
- Kuspa A, Loomis WF: Tagging developmental genes in *Dictyostelium* by restriction enzyme-mediated integration of plasmid DNA. *Proc Natl Acad Sci USA* 1992, 89:8803-8807.
- Gossler A, Joyner AL, Rossant J, Skarnes WC: Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* 1989, 244:463-465.
- Brenner DG, Lin-Chao S, Cohen SN: Analysis of mammalian cell genetic regulation *in situ* by using retrovirus-derived "portable exons" carrying the *Escherichia coli lacZ* gene. *Proc Natl Acad Sci USA* 1989, 86:5517-5521.
- Friedrich G, Soriano P: Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* 1991, 5:1513-1523.
- Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC: Green fluorescent protein as a marker for gene expression. *Science* 1994, 263:802-805.
- Zernicka-Goetz M, Pines J, Ryan K, Siemering KR, Haseloff J, Evans MJ, Gurdon JB: An indelible lineage marker for *Xenopus* using a mutated green fluorescent protein. *Development* 1996, 122:3719-3724.
- Skarnes WC, Moss JE, Hurlley SM, Beddington RS: Capturing genes encoding membrane and secreted proteins important for mouse development. *Proc Natl Acad Sci USA* 1995, 92:6592-6596.
- Suzuki K, Yamamoto T, Inoue J: Molecular cloning of cDNA encoding the *Xenopus* homolog of mammalian RelB. *Nucleic Acids Res* 1995, 23:4664-4669.
- Bisbee CA, Baker MA, Wilson AC, Haji-Azimi I, Fischberg M: Albumin phylogeny for clawed frogs (*Xenopus*). *Science* 1977, 195:785-787.
- Cannatella DC, de Sa RO: *Xenopus laevis* as a model organism. *Syst Biol* 1993, 42:476-507.
- Amaya E, Offield MF, Grainger RM: Frog genetics: *Xenopus tropicalis* jumps into the future. *Trends Genet* 1998, 14:253-255.
- Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, et al.: Zebrafish *hox* clusters and vertebrate genome evolution. *Science* 1998, 282:1711-1714.
- Postlethwait JH, Yan YL, Gates MA, Horne S, Amores A, Brownlie A, et al.: Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* 1998, 18:345-349.
- Meyer A, Malaga-Trillo E: Vertebrate genomics: more fishy tales about *Hox* genes. *Curr Biol* 1999, 9:210-213.

A gene trap approach in *Xenopus*

Odile J. Bronchain, Katharine O. Hartley and Enrique Amaya

Current Biology 11 October 1999, 9:1195–1198

Figure S1

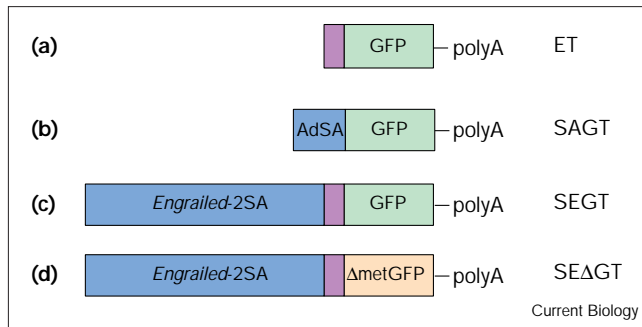


Diagram of the GFP gene trap vectors. (a) The basic exon trap vector (ET) contains a stretch of 14 glycine residues subcloned in frame to the GFP coding sequence (Gly-GFP). (b) The adenovirus gene trap construct (SAGT) contains a 120 base pair fragment, including the intron1-exon2 boundary of the adenovirus major late transcript splice acceptor, subcloned upstream of the GFP open reading frame. (c) The engrailed gene trap construct (SEGT) contains a 2.2 kb DNA fragment, including the *En-2* intron/homeobox containing exon boundary fragment. This fragment was subcloned in frame upstream of Gly-GFP. (d) The engrailed gene trap construct with the mutated initiation codon of GFP (SE Δ GT) is derived from SEGT. Purple boxes, glycine bridges; Ad, adenovirus; *En-2*, engrailed; SA, splice acceptor.

Supplementary materials and methods

Transgenic frog embryos were generated as previously described [S1,S2]. The gene trap vectors were linearized using *HindIII/NotI*. The REMI experiments were conducted using 100 ng of linearized plasmid DNA per reaction and the nuclei were treated with *HindIII*.

Plasmids

The γ 1crystallin-GFP construct was generated by subcloning a 2.2 kb fragment of the γ 1crystallin genomic DNA encompassing the promoter region (received from Robert Grainger, [S3]) upstream of an enhanced version of GFP [S4]. The adenovirus splice acceptor construct (SAGT) was generated by inserting a 120 base pair *NcoI-BamHI* DNA fragment from pSA-GFPNeo (Dixon and Evans, unpublished observations) into a GUAS-GFP3 construct linearized with *BamHI-NcoI* (K.H. and E.A., unpublished observations). The engrailed splice acceptor constructs (SEGT, SE Δ GT) were generated by inserting a 2.2 kb *HindIII-BglII* DNA fragment from pGTM1 [S5] containing the *En2* splice acceptor (SA) sequence into the *HindIII-BglII* sites of either the Gly-GFP construct or Gly- Δ GFP*. The Gly-GFP construct is derived from the pCS-GFP3 construct, which contains a modified GFP [S4] subcloned in pCS2+ [S6,S7]. An oligonucleotide encoding a stretch of 14 glycine residues (5'-GAATTCG[GGA]₁₄GGATCC-3') was subcloned in frame into the *EcoRI-BamHI* sites of pCSGFP3. The Δ met-GFP was generated by PCR. The amino terminal portion of GFP3 was amplified using a primer homologous to base pairs 3-34 (5'-CGGGATCCAAAGGAGAA-GAATTTCACTG-3') with addition of a *BamHI* restriction site in 5' for cloning purposes, and a primer homologous to the cDNA sequence localized at base pairs 426-448 (5'-CGTTGTGGGAGTTGTAGTTG-TAT-3'). The PCR product was subsequently digested with *BamHI* and *NcoI* restriction enzymes and subcloned into the *BamHI-NcoI* sites of the Gly-GFP construct. To avoid potential read-through from the polymerase, we have exchanged the single SV40 polyadenylation (polyA)

sequence with a double polyA sequence (*Xenopus* 3' β -globin UTR/SV40 polyA) from the pCSXFD construct [S1]. The double polyA was subcloned as an *XbaI-NotI* fragment from pCSXFD into the *XbaI-NotI* sites of the Gly- Δ GFP to give the final Gly- Δ GFP* construct.

5' RACE PCR

Total cellular RNA from individual embryos was isolated by the guanidinium thiocyanate-phenol-chloroform method [S8]. A 5' RACE PCR was performed as previously described [S9]. The first strand cDNA was reverse transcribed using a GFP-specific primer homologous to base pairs 426-448 (5'-CGTTGTGGGAGTTGTAGTTG-TAT-3'). As we were starting from individual embryos that might express the marker gene in only a few cells, the amount of total RNA used for the reverse transcription reaction varied from sample to sample and was estimated on the basis of the expression pattern of the marker gene and the stage at which the embryos were sacrificed. On average we used 2-10 μ g of total RNA. The total RNA was annealed to 10 ng of the GFP-specific oligonucleotide and incubated for 1 h with superscript II (Gibco) according to the manufacturer, recommendations. The synthesis of the second strand cDNA was performed as previously described [S9]. A series of nested PCR reactions were then performed using the 5' RACE primer (5'-GGTTGTGAGCTCTTCTAGATGG-3') [S9] and a succession of GFP-specific primers, each homologous to a sequence closer to the 5' end of the GFP cDNA: PCR1: 5'-CGTGTCTGTAGTTCCCGTCGTC-3', PCR2: 5'-GAAAAGCATTGAACAC-CATAAGT-3', and PCR3: 5'-GCATCACCTTCACCCTCTCCACT-3'. Usually, two rounds of nested PCR were sufficient to recover the trap gene fusion transcripts. The PCR reactions were performed using the expand high fidelity PCR system (Boehringer Mannheim) in a total volume of 50 μ l, using 600 nM of each primer, 500 μ M dNTPs, in 1 \times Buffer 3 supplemented with MgCl₂ at a final concentration of 5 mM. The amplification was carried out as followed: a 2 min hot start at 94°C, followed by 30 cycles (30 s at 94°C, 1 min 30 s at 60°C, and 3 min at 72°C). The reaction was completed by a final extension step at 72°C for 10 min, subcloned using the pGEM-T Vector system I (Promega) and final products were sequenced.

Supplementary references

- S1. Kroll KL, Amaya E: Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. *Development* 1996, 122:3173-3183.
- S2. Amaya E, Kroll KL: A method for generating transgenic frog embryos. In *Methods in Molecular Biology*, Sharpe P, Mason I, Humana Press Inc.: Totowa, NJ; 1999, 97:393-414.
- S3. Smolich BD, Tarkington SK, Saha MS, Stathakis DG, Grainger RM: Characterization of *Xenopus laevis* gamma-crystallin-encoding genes. *Gene* 1993, 128:189-195.
- S4. Zernicka-Goetz M, Pines J, Ryan K, Siemering KR, Haseloff J, Evans MJ, et al.: An indelible lineage marker for *Xenopus* using a mutated green fluorescent protein. *Development* 1996, 122:3719-3724.
- S5. Skarnes WC, Moss JE, Hurlley SM, Beddington RS: Capturing genes encoding membrane and secreted proteins important for mouse development. *Proc Natl Acad Sci USA* 1995, 92:6592-6596.
- S6. Rupp RA, Snider L, Weintraub H: *Xenopus* embryos regulate the nuclear localization of XMyoD. *Genes Dev* 1994, 8:1311-1323.
- S7. Turner DL, Weintraub H: Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev* 1994, 8:1434-1447.
- S8. Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987, 162:156-159.
- S9. Townley DJ, Avery BJ, Rosen B, Skarnes WC: Rapid sequence analysis of gene trap integrations to generate a resource of insertional mutations in mice. *Genome Res* 1997, 7:293-298.