Minireview

Gene trap: a way to identify novel genes and unravel their biological function

Francesco Cecconi*, Barbara I. Meyer

Department of Molecular Cell Biology, Max-Planck-Institute of Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany

Received 25 April 2000

Edited by Gianni Cesareni

Abstract The gene trap methodology is a powerful tool to characterize novel genes and analyze their importance in biological phenomena. It is based on the use of mouse embryonic stem cells and reporter vectors designed to randomly integrate into the genome, tagging an insertion site and generating a mutation. Theoretically, all the 100 000 genes present in the mouse genome could be tagged and functionally inactivated at the same time. Here we describe the basic concepts and perspectives of this methodology and show some results obtained by the gene trap approach used to study molecular cascades in basic cell biology and in developmental processes.

Key words: *Apaf1*; Genomics; Library; *netrin 1*; Transgenic; Trap vector

1. Introduction

The worldwide effort aimed at understanding the structural and the informational content of human and murine genomes is near to completion. Based on the sequence information, general conclusions about the genetic complexity of these organisms can be drawn, but the knowledge of gene function and its relevance for human diseases will still be unknown.

The general opinion maintains that every gene should be studied in its natural biological context. In this respect, advances are tremendous, due to the advent of gene targeting in mice. Using that approach, mutants have been generated for a huge variety of genes and their function largely revealed. This methodology depends a priori on the isolation and molecular analysis of a given gene. As an example, the specific inactivation through gene targeting of particular genes, such as transcriptional factors or signaling molecules, has shed new light on the functional complexity of mammalian development ([1]; for a compendium about targeted mutations, see [2–4]). Various strategies have been used to isolate such genes, starting from the classic genetic studies in lower organisms, as *Drosophila melanogaster* and *Caenorhabditis elegans*. The *Hox* genes, for example, were initially isolated from the mouse genome on the basis of their sequence similarities to conserved sequence motifs within homologous genes of *Drosophila*.

However, the vast majority of the estimated 100 000 genes present in the mouse genome have not yet been cloned, which precludes their disruption by homologous recombination. To identify and mutate new murine genes, an excellent experimental strategy is the gene trap [5–8]. In our laboratory and in many others, several genes important for cell metabolism and embryonic development have been captured and functionally analyzed in this way. The purpose of this review is to summarize the principles, problems and new results of the gene trap approach. Moreover, we will discuss the advantages and disadvantages connected with the establishment of large-scale gene trap strategies and the striking possibility to adapt the gene trap approach to another organism, *Xenopus laevis* [9]. The final chapter of this review will be dedicated to reporting the capture and simultaneous inactivation of two genes important for basic cell biology and developmental processes, *Apaf1* and *netrin 1*, as two examples of successful gene trapping in our lab.

In our opinion, the generation of gene trap libraries will be a fundamental step in the future of functional genomics. Making these libraries available to the scientific community and using them to study gene function are primary objectives of several public and private laboratories. The establishment of new screening devices for the trapping events will optimize this technique, decreasing the workload and thereby providing researchers with an efficient tool for discovering drug targets.

2. Gene trap methodology

The gene trap approach, represented in Fig. 1, is based on the use of murine embryonic stem cells (ES) and a class of reporter vectors, that has been designed containing a splice-acceptor site upstream of the β-galactosidase (*lacZ*) gene and the neomycin resistance gene (*neo*). Integration of these vectors into a genomic locus downstream of a functional promoter results in the generation of a fusion transcript between the endogenous gene and the *lacZ* gene. Fusion transcripts from insertion of these vectors mimic endogenous gene expression at the insertion locus. This expression can be monitored by visualizing *lacZ* activity [6,10]. The tagged genes can be identified by the use of anchored PCR procedure [11], performed on the selected ES cell clones. Furthermore, usually the gene trap vectors will also act as insertional mutagens, disrupting the endogenous gene function. ES cell lines can be used as a bridge between genetic manipulation in vitro and biological analysis in vivo. Indeed, after genetic manipulation (electroporation or retroviral infection, for a review, see [12]), the cells remain pluripotent and, when reintroduced into morulae by embryo aggregation (or into blastocysts by microin-

*Corresponding author. Fax: (49)-551-2011504.
E-mail: fceccon@gwdg.de

0014-5793/00/$20.00
PII: S0014-5793(00)01779-8

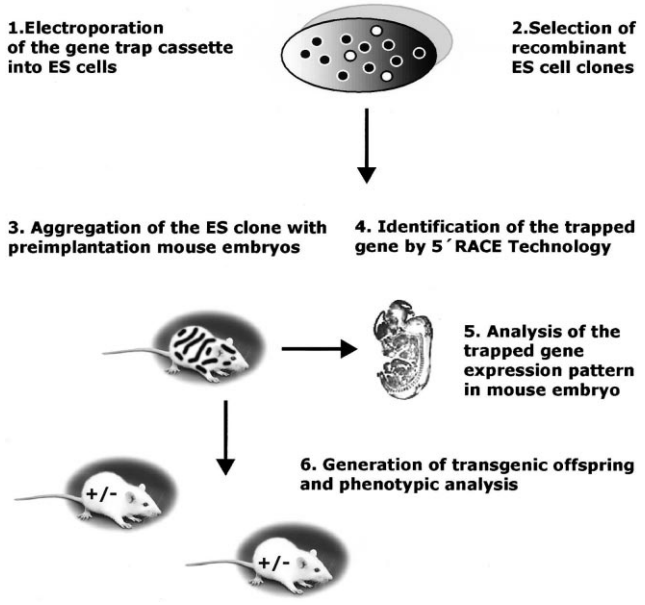

Fig. 1. Strategy of the gene trap approach. The scheme represents the strategy used in our laboratory. As an alternative, the insertion of exogenous DNA into the ES cells genome can be carried out by retroviral infection. Prescreening approaches could be applied before selection of recombinant ES clones (2.). The transgenic offspring at (6.) is derived by breeding the chimeric animal (labeled by the dark stripes) with an inbred strain C57bl/6 or an outbred strain (NMRI). The progeny will be tested for germ line transmission and hybrid crosses between F1 littermates will give rise to the homozygous specimen.

jection), they can contribute to all tissues of the mouse including the germ line [13]. Using this strategy, the spatial and temporal expression patterns of the endogenous genes during embryogenesis and adulthood can be analyzed. The phenotypic consequences of the gene trap mutation can also be investigated.

We have been interested in analyzing the genetic control of the development of mammalian central nervous system (CNS). The gene trap is an experimental strategy particularly promising in mammals, and therefore suited to this purpose.

2.1. Gene trap vectors

As described above, the rationale behind the use of reporter constructs is to tag and detect *cis*-regulatory sequences by locating the reporter gene within an endogenous gene. In certain integrations, the reporter construct can also disrupt endogenous gene function and can therefore act as a mutagen. All trapping vectors devised so far consist of modifications on one basic structure. A transition zone serves to place endogenous genomic sequences next to a reporter gene, which is then brought under their influence. The vector can also allow the selection of insertion events by carrying a reporter gene that itself serves as a selectable marker. Fig. 2 shows the schemes of seven selected gene trap vectors used successfully worldwide. In the vectors that we have been using, the transition zone contains a splice-acceptor site derived from the *engrailed 2* gene (ei and ee, *eng2* intron element and exon element, respectively). The reporter gene is *lacZ* and the selectable marker is neomycin phosphotransferase (*neo*) fused together into the bi-functional *lacZ/neo* gene (β-geo). The presence of an internal ribosomal entry site (IRES) derived from the encephalomyocarditis virus ensures the reporter and resistance activity without requiring an in-frame fusion with the coding region of the endogenous gene (Fig. 2, IRESβgeo, [14,15]).

To capture genes that are not active in undifferentiated ES cells and could be induced by in vitro differentiation (see below), Salminen et al. [16] have constructed a new poly(A) trap vector. In IRESβgalNeo(−pA), the transcription of *neo* is under the control of the β-actin constitutive promoter (βa), while the β-gal expression is dependent on the activity of the trapped gene. The *Pax2* splice-donor element follows the bicistronic structure. In IRESβgalNeo(+pA), the independently transcribed *neo* gene has its own poly(A) signal. The authors elegantly showed that the deletion of the poly(A) structure reduced the number of integrations occurring outside of the transcription units. Many important genes are known to be inactive in undifferentiated ES cells [17]. When the ES cells are cultured under conditions which allow differentiation, a 3-fold increase in positive stained clones is observed with IRESβgalNeo(−pA). As a consequence, this vector will be very useful in the search for genes responding to specific regulatory factors in vitro or in the identification of genes specific to early steps in differentiation pathways.

The plasmid vector PT1βgeo is one of the first gene trap vectors generated by the group of W.C. Skarnes in 1989 and was successfully used in several laboratories. However, the lack of the IRES element renders this vector susceptible to silencing by non-in-frame fusion events within the endogenous transcript. Wiles and collaborators [18] have recently reported the establishment of a large-scale gene trap screen in ES cells using plasmid and retroviral vectors (Fig. 2, PT1βgeo and U3βgeo). The introduction of a trap vector by a retrovirus to ES cells also has some advantages. The insertional capa-

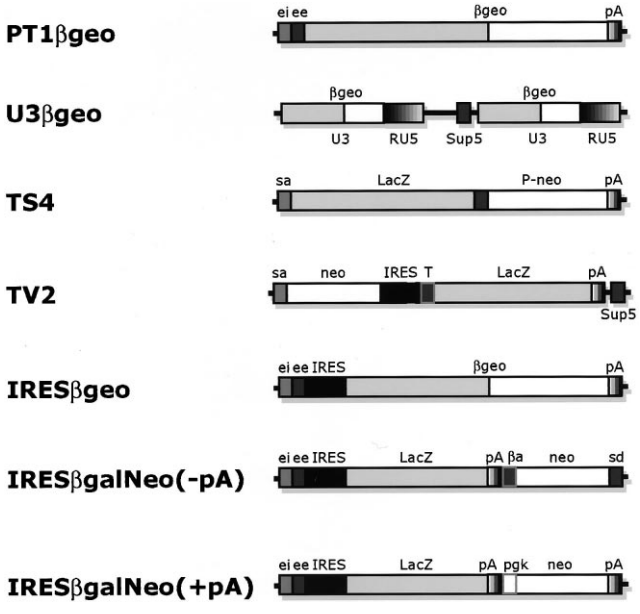

Fig. 2. Examples of gene trap vectors. For references, see text. ei: *eng2* intron; ee: *eng2* exon; βgeo: *lacZ–neo* fusion; pA: polyadenylation signal; U3: U3 region of enhancerless Moloney murine leukemia; Sup5: *E. coli supF* tRNA gene; sa: *Zfp-40* splicing acceptor; IRES: internal ribosomal entry site; T, SV40 nuclear transporting signal; P-neo, *neo* gene driven by *HSV-tk* promoter; βa: human β-actin promoter; sd: mouse *Pax2* splice-donor site; pgk: phosphoglycerate kinase promoter.

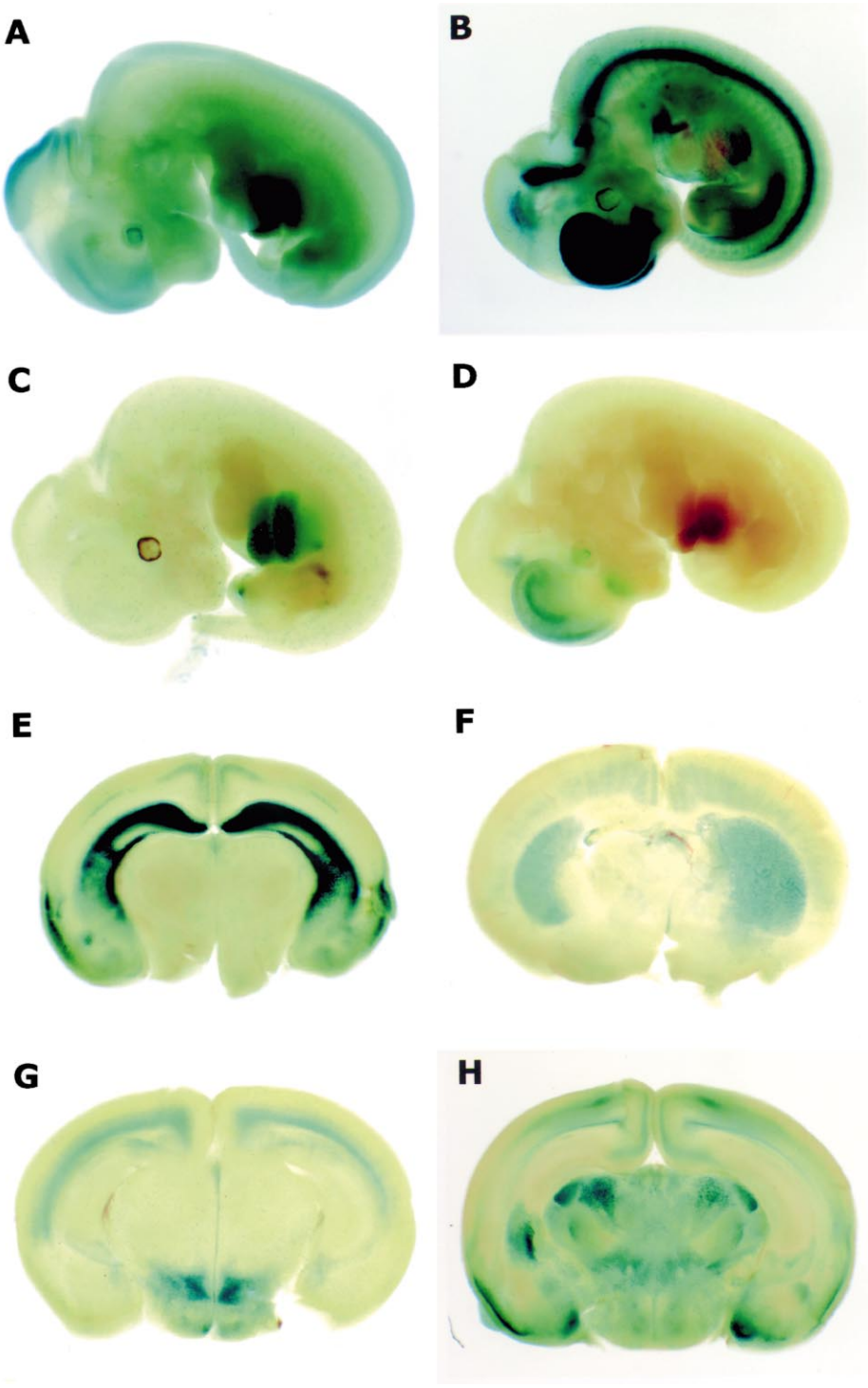

Fig. 3. Expression of different gene trap lines visualized by β-gal activity. (A–D) show a lateral view on day e12.5 embryos. (E–H) show a view of cross-sectioned pn9 mouse brains. (A) Gene trap line 6B-4 has a widespread expression throughout the embryo, while the dorsal CNS shows the highest activity level. (B) Gene trap line 9A-33 demonstrates very strong restricted expression in the tegmentum, in the ventral metencephalon and the spinal cord. The most dominant activity is located in the cortical hemispheres. Additional expression domains are heart, liver, spinal nerves and limbs. (C) Gene trap line 3C-40 reveals restricted activity in the developing liver. (D) Gene trap line 2C-98 has restricted expression in the dorsal telencephalon, the anterior tectum and the eye. (E) Gene trap line 3C-178 shows strong expression in the hippocampus and weaker activity in the cingular and piriform cortex of the early postnatal brain. (F) Gene trap line 6A-126 exhibits restricted expression into the striatum and a fainter level of activity in the subventricular zone of the cingular cortex. (G) Gene trap line 9A-196 expression is located in the subventricular zone of the somatosensory cortex and the hypothalamus. (H) Genetrap line 7B-79 shows, beside regional restricted expression in the cingular and motor cortex, a broader activity in the thalamus.

bility is increased to almost 100% and rearrangements at the site of insertion have been slight [5,19,20]. However, hot spots for retroviral insertions retain the potential to bias a screen based on this class of vectors. As an example, the U3βgeo vector, from the retroviral class, carries the βgeo cassette in the U3 region of an enhancerless Moloney murine leukemia virus LTR [18,21,22]. When U3βgeo integrates into an intron, an exon or a 5′ uncoding element of a genomic locus, it results in a fused βgeo transcript, which induces dysfunction of the tagged gene. By the parallel use of PT1βgeo and U3βgeo, Wiles et al. [18] generated 12 000 targeted ES cell lines, most of them suitable for generation of transgenic animals.

Finally, vectors TS4 and TV2, constructed by Takeuchi et al. [20,23], represent interesting modifications of the described master-vectors. TS4 contains the *neo* gene driven by HSV-tk promoter (P-*neo*) and TV2 displays the sequence for the nuclear transporting peptide derived from the SV40 large T gene at the 5′ end of the *lacZ* sequence. The 'T' element is expected to carry the β-gal to the nucleus, facilitating staining detection.

It should be mentioned that many other vectors have been successfully used in the gene trap methodology, vectors that contain only a promoterless *neo* without *lacZ* [24,25] and vectors containing other drug resistant genes (i.e. hygromycin or phleomycin [26,27]). Skarnes et al. [14] have constructed a vector (pGT1.8TM) expressing the transmembrane portion of the CD4 type I protein (TM) fused to βgeo. Integration of these vectors into genes that contain a signal sequence, to cause accumulation into the endoplasmic reticulum (ER), produces βgeo insertions in the membrane in an active configuration. Integration of the vectors in genes that lack a signal sequence produces a fusion protein with an internal TM domain. Insertions of this kind expose βgeo to the lumen of the ER, where β-gal activity is lost. By this method, it is possible to detect integrations in genes that encode cell surface proteins, a class of genes previously missed by conventional gene trap vectors [14]. Good examples of genes captured by this technique are *netrin 1* and *neuropilin 2* [28,29].

Finally, Thorey et al. [30] have engineered a gene trap vector encoding the Cre recombinase, which should be electroporated in a particular ES cell line containing in its genome the *neo* gene flanked by the *loxP* elements and followed by the *lacZ* gene. Integration of that construct into the genome of this specific line will allow LacZ expression only when the *neo* gene has been removed by Cre-mediated recombination. This method is devised to trap genes only transiently expressed in the ES cells, allowing tracking of the cell lineages within cell populations and tissues in the developing embryo.

2.2. ES cells manipulation and construction of transgenic mouse lines

Integration of the gene trap vectors in ES cells allows the establishment of mutated lines and the transmission of these mutations through the germ line by chimera production (see Fig. 1). By electroporation, hundreds of potentially mutated ES cell lines expressing β-gal can easily be established and identified in vitro. ES cells are derived from the inner cell mass of a normal or implantation-delayed blastocyst. They can be maintained in culture in a pluripotent state with a normal karyotype on a fibroblastic feeder layer. Following manipulation or screening in vitro (see below for selective prescreening methodologies in vitro), they can be reintroduced into morulae (by embryo aggregation) from mice of a different coat color, where they intermingle with host cells. The ES cells readily aggregate with morulae; therefore, all that is required for chimera production is contact of the two cell populations [31]. Upon transfer to a pseudo-pregnant recipient, the ES cells participate in normal development of the chimeric embryo and contribute to all cell types, including the germ line. Once germ line chimeras are obtained, they will be used as a source for heterozygous mice carrying the inserted trapping vector. The activity of β-gal can be detected in whole embryos and in sectioned postnatal tissues by a chemical reaction that leads to a blue coloring of positive cells (blue cleaved X-gal substrate). Staining patterns are then classified for both spatial and temporal differences between strains with particular attention to the domain of interest of the laboratory. In Fig. 3, we show a collection of expression patterns obtained by tracking X-gal staining in embryonic and postnatal brain tissues. Since this method utilizes an extremely simple enzymatic reaction, the full characterization of the expression pattern of the captured gene can be completed in an efficient manner. Noteworthy is that our study demonstrates a clear correlation between the specific vector used and the corresponding expression patterns in the established mouse lines, as shown in Table 1. As expected from previous studies [16], the IRESβgal-Neo(−pA), which has been shown to produce more restricted LacZ staining in vitro than the IRESβgeo vector, also exhibits a high percentage of restricted expression patterns during mouse development and in postnatal brains.

It should be mentioned that the expression of the reporter gene does not faithfully mimic that of the endogenous trapped gene in some cases. This could be caused, for example, by the insertional disruption of intronic regulative elements of the endogenous gene. It is therefore essential to confirm the expression of the trapped gene, once it has been successfully cloned by in situ hybridization [32].

Table 1
Statistical analysis of LacZ expression patterns shown from mouse gene trap lines generated with different vectors

Vector	Negative (%)	Extended (%)	Restricted (%)	CNS (%)
IRESβgeo[a] (45)	20	44	36	31
IRESβgalNeo(−pA)[b] (56)	29	7	64	43

ES cell clones selected on the basis of a restricted LacZ activity pattern (blue staining in only 10% of the colonies in a 2 cm dish and in 50% of the cells within a colony) under standard/undifferentiated culture conditions have been used to generate the mouse lines. The expression patterns have been analyzed during embryonic development at stage e9.5, e12.5 and e15.5 and on brain samples from postnatal day 9 (pn9). The individual expression patterns of the mouse lines have been classified into negative, extended (activity in different cell types, organs and structures) and restricted (activity restricted to one cell type, organ or structure). Within the restricted patterns, the subgroup of lines showing expression in the CNS has been determined. The number of analyzed mouse lines is indicated in brackets.
[a][14,15].
[b][16].

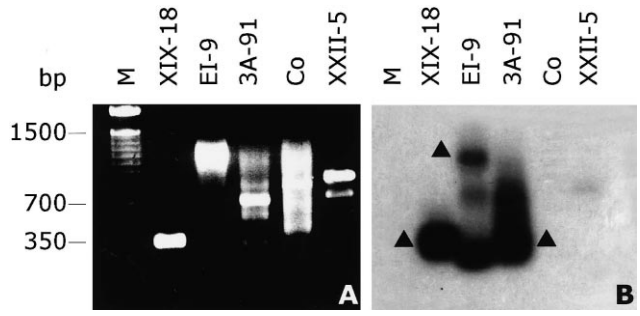

Fig. 4. Selection of trapped cDNAs by nested PCR and Southern blotting. Electrophoresis on 2% agarose gel of the 5′ RACE products from several gene trap lines (A). Co: control on standard DNA. To size select cDNAs, in order to amplify the most sequence-informative fragments, the first strand RT reactions on ES cells mRNA were micro-dialyzed. This allowed the discharge of DNA fragments less than 200 bp. Size-selected cDNAs then underwent two rounds of nested PCR amplification and were micro-dialyzed after each round. Lines XIX-18 (*Apaf1*), EI-9 and 3A-91 were obtained by electroporation of the IRESβgeo vector into ES cells. Autoradiography of the Southern blot of the same gel, labeled with a radioactive probe corresponding to the ee region of the vector (B) (see Fig. 2). The arrowheads indicate the fragments positive to the hybridization, which were selected for cloning and consequent identification of the trapped gene. The line XXII-5 is shown as an example of an unsuccessful 5′ RACE approach.

2.3. Molecular analysis of the trapped gene

At this point, tagged genes can be identified by the use of anchored (rapid amplification of cDNA ends (RACE)) PCR procedures [11,15,33]. Products could then be cloned, sequenced and compared with databases to verify the novelty of the trapped gene. Fig. 4 shows an example of the RACE protocol outcome performed in our laboratory. After the first strand reaction of the reverse transcriptase (RT), the cDNA template obtained from the ES cells mRNA is amplified by two subsequent rounds of nested PCR. Primers were derived from the IRES sequence (Fig. 4A). The amplified products from several ES cell lines were then blotted and hybridized with a probe specific for the *eng2* splice-acceptor site, 5′ upstream of the IRES (see Fig. 2, ee of IRESβgeo). In Fig. 4B, we point out the PCR products, which were successfully cloned and correspond to portions of the trapped gene. There are two major problems connected with this approach. First, the relative amount of fusion mRNA transcribed in the ES cells could not be high enough to allow a successful first strand polymerization. Second, tandem insertion of the vector into the genome will increase the background of the nested PCR reactions.

Another problem, connected with the capability of the gene trap methodology to generate insertional mutagenesis, is the potential of a genomic locus to undergo splice-around events. The splice-around phenomenon occurs when the splice-acceptor site contained in the vector is ignored by the spliceosome. This could happen in 100% of the cases, rendering the integration completely silent and therefore not interfering with the screening procedure. In the worst scenario, a splice-around event could occur only in a time- or space-restricted manner during developmental or adult tissues, generating a complete absence of phenotype. For this reason, it is essential to perform a molecular analysis of the transcriptional fusion at the integration site, by RT-PCR or Northern blot, whenever an expression pattern is considered to be interesting for phenotypic analysis. However, the possibility to generate a hypomorphic allele could be intended as an advantage, providing investigators with a variety of different insertional mutations of the same gene.

3. Prescreening of gene trap events

On the basis of the expression patterns revealed by the tagged genes, a large screen of the adult CNS has recently been performed by Steel and coworkers [34], taking advantage of mouse lines previously generated [35]. Four trapped genes showed interesting expression patterns in the mouse hippocampus. Those were phenotypically analyzed for involvement in synaptic plasticity (see below). In our laboratory, several gene trap lines have been selected for expression in the embryonic and postnatal CNS and are undergoing phenotypic analysis (see Table 1; [36,37]).

However, the possibility to screen by expression pattern analysis in embryonic or adult tissues is severely hampered by space limitation of the mouse colony, which could render the strategy too difficult for specific purposes.

An area of growing interest is the possibility of directing ES cells along specific lineages, such as the neuronal lineage. In the absence of feeder layers or DIA/LIF, ES cells have the capacity to differentiate into many cell types in vitro, either spontaneously or induced by chemical agents such as neurotrophic factors [10,38]. In aggregates of differentiated ES cells, embryoid bodies contain other various differentiated cells such as hematopoietic cells, muscle cells and chondrocytes [39–42]. The gene trap methodology has been adapted to use after a prescreening for genes activated during these differentiating conditions [24,43–47].

ES cell clones could also be a priori classified, according to their subcellular staining distribution: diffuse, nuclear, cytoplasmic and surface-localized. For instance Skarnes and collaborators [14] have planned a novel gene trap strategy, named secretory trap, which is based on a new class of vectors (see above) and this subcellular localization prescreening.

These approaches will comprehensively allow for identification and insertional mutation of genes whose expression is modulated during development and the differentiation of several tissues and organs.

4. Gene trap in *Xenopus*

The amphibian transgenesis procedure, set up by Kroll and Amaya [48], is a rapid and efficient protocol suitable for multi-generation experiments, which does not give rise to mosaic germ line in the founder [9]. The *Xenopus* transgenic technique is not based on ES cell establishment, but on a restriction enzyme-mediated integration by introduction into condensed sperm nuclei in vitro (REMI [48]). It must be mentioned that attempts are ongoing to establish in other organisms cell lines similar to murine ES cells (i.e. in chicken [49]).

To apply the *Xenopus* transgenic generation to a more wide gene trap approach, the authors generated a new vector replacing the *lacZ* gene with the green fluorescent protein gene (GFP), whose expression could be detected in living individuals [9]. The efficacy of the approach was demonstrated by the generation of hundreds of transgenic embryos, in a few hours, expressing GFP in many tissue types. The main advantage of using *Xenopus* as a target organism is that the embryos develop externally and expression of the marker (GFP) can be

monitored almost limitlessly. Genotyping of heterozygotes would also be an easier task, since the positive embryos could be selected by monitoring GFP expression. The possibility of obtaining thousands of mutant embryos from a single experimental mating makes this methodology a plausible and powerful tool for a large-scale functional genomics program. The tetraploidity of *Xenopus*, however, can be a negative factor, as the organism would have a more complex genome exhibiting functional genetic redundancy [9].

5. Large-scale strategies

Two major strategies have been used in the last 5 years for creating libraries of gene trap lines and sequences of the tagged genes. The first strategy has been organized by a German consortium, which has generated a reference library of gene trap sequence tags (GTST [18]). The second is lead by Lexicon Genetics (TX, USA). Both approaches are based on a gene trap sequence screening automated in the 96-well format and both have the potential to represent insertional mutations for most of the mammalian genes in mouse ES cells [18,50,51]. While those libraries are on their way to completion, new databases are being created which will provide a minimal amount of required data needed, before proceeding to study the organism in vivo. In the age of functional genomics, those databases will be an invaluable tool for investigators.

6. Phenotypic analysis: some examples

In favorable cases, considering the redundancy of the mammalian genome and the potential of a splice-around event, the gene trap insertions would lead to a phenotype.

The number of gene trap mouse lines responsible for phenotypes is rapidly increasing, from early embryonic lethality to adult fertility impairment.

Table 2 summarizes the most relevant trapped genes with a clear phenotype isolated so far.

Friedrich and Soriano [5] were the first to take advantage of the gene trap methodology, in their attempt to identify developmentally regulated genes. To determine whether recessive phenotypes were associated with mutations in the 'trapped genes', heterozygous mice from 24 strains were intercrossed and the offspring genotyped. Homozygous offspring could not be isolated from nine strains, indicating that the insertions led to a recessive lethal phenotype. Cloning of one of the trapped genes, which led to recessive lethal phenotype, indicated that insertion disrupted the transcriptional enhancer factor 1 (*TEF-1*) gene. The characterized strain exhibited a phenotype associated with heart dysfunction; thus their analysis has shown that *TEF-1* plays an essential role in the maturation stage of cardiogenesis [52]. Another example of the efficiency of the gene trap strategy has been provided by DeGregori et al. [25]: an embryonic lethal mutation has been accounted for by the insertional disruption of the gene *fug1*, similar to the *RNA1* gene of *Saccharomyces cerevisiae*. Using a gene trap strategy, Torres et al. [38] have isolated a mouse mutation for the gene encoding the α-E-catenin. Catenins are proteins associated with the cytoplasmic domain of cadherins, a family of transmembrane cell adhesion molecules. The cadherin–catenin adhesion system is involved in morphogenesis during development and in maintaining the integrity of different tissue types. Another mouse mutation, termed *jumonji* (*jmj*), was also generated by a gene trap strategy [23]. Expression of *jmj*, as monitored by X-gal staining, was detected predominantly at the midbrain–hindbrain boundary and in the cerebellum. The phenotype showed dramatic brain malformations. More recently, Ross and coworker [53] identified the anti-apoptotic factor bcl-w by the gene trap method. The expression of this gene in the testis appears to be restricted to elongating spermatids and Sertoli cells. The authors suggest a role of bcl-w for germ cell differentiation.

Finally, neuropilin 2, trapped by the secretory trap [14], is a necessary receptor for a class 3 semaphorin (Sema3F) and plays important roles in establishing several axonal trajectories in the developing CNS [29].

In the final portion of this review, we will focus on two genes trapped in our laboratory (*Apaf1* and *netrin 1*) which turned out to be key factors in basic cell biology functions, cell death and cell adhesion.

6.1. Apaf1

Among several trapped genes isolated and analyzed so far in our laboratory, *Apaf1* (apoptotic protease activating factor 1, reviewed in [54]) is likely to be the most important example. The gene trap line XIX-18 was selected during an expression screening in mouse embryos performed at several stages of development. Strong LacZ expression was detected in the differentiating layers of the neural tube, in the ganglion cells layer of the neuro-retina and in the limbs. After identification of the 5′ RACE product from the gene trap line XIX-18 as the murine *Apaf1* cDNA homolog (see Fig. 4A,B, second lanes), we cloned the full length mouse *Apaf1* cDNA and analyzed the embryonic phenotype in mice homozygous for the insertional mutation [36]. The gene trap insertional mutation generated a null allele, as demonstrated by Northern and Western blot analysis.

Table 2
Examples of genes isolated and mutated by gene trap

Gene	Product	*LacZ*[a]	Vector	Reference
TEF-1	transcriptional factor	extended	Rosaβ-geo	[52]
Fug1	*RNA1*-like	extended	U3neo	[25]
Netrin 1	secreted factor	restricted	PGT1.8TM	[28]
			IRESβgeo	[37]
α-E-Catenin	adhesion molecule	restricted	PGT1.8geo	[38]
Jmj	novel	restricted	TV2	[23]
Bcl-w	anti-apoptotic protein	restricted	Rosaβ-gal	[53]
Apaf1	pro-apoptotic protein	extended	IRESβgeo	[36]
HSP90β	Heat-shock protein	extended	PGT1.8geo	[68]
Neuropilin 2	TM receptor	extended	PGT1.8TM	[29]

[a]LacZ expression pattern analyzed at several stages of development.

In *C. elegans*, *CED-9* inhibits cell death [55], *CED-3* and *CED-4* promote cell death [56,57]. Apaf1 is a CED-4 homolog and it participates in the cytochrome *c*/dATP-dependent activation of a mammalian CED-3 homolog, caspase 3 (Casp3; [58]), by prior activation of caspase 9 (Casp9; [59]).

Mice in which Casp3 has been deleted by targeted disruption die perinatally with a massive cell overgrowth in CNS, as a result of an apoptosis deficiency in the neuroepithelial cells [60]. The Casp9 knockout mice showed the same abnormality [61]. Obviously, Apaf1 was a good candidate to play a crucial role in mammalian development, being upstream of Casp9 and Casp3.

Abnormal embryos were found only at e12.5 and later. All abnormal embryos were homozygous for the Apaf1 mutation, which is lethal around e16.5. *Apaf1*$^{-/-}$ embryos at e14.5 show a characteristic cranio-facial phenotype whose major traits are facial cleft, absence of the skull vault and of all vomer and ethmoidal elements, and rostral or dorsal exencephaly (Fig. 5A,B). In addition, *Apaf1* mutants show alterations of the retina, lens and eye vascular system. At e14.5, a hyperplasic folded retina occupies most of the optic cup (Fig. 5C,D; retinal pigmented epithelium is indicated by the arrowheads). Apoptosis has been described as a regulator of the cell number during normal development of the retina (histogenetic cell death).

From e12.5 onwards, the mutant brain shows important morphological distortions. The telencephalic vesicles are folded and reduced in size. Anatomically, the brain hyperplasia of Apaf1 mutants is particularly intense in the diencephalon and the midbrain. Histologically, the hyperplasia is due to the abnormal proliferation of neural precursor cells and their irregular distribution in the tissue, as shown by the proneural marker Delta-1 (Fig. 5E,F). The brain phenotypes in *Apaf1*, *Casp3* and *Casp9* null mutations [36,60,61] are extremely similar. In the wild-type CNS but not in the mutant, we observed numerous active Casp3-immunoreactive cells. These findings strongly indicate that these proteins could be components of the same apoptotic pathway during brain development. This is consistent with their functional interactions shown in vitro [58,59].

In summary, Apaf1 is involved in histogenetic cell death (control of cell number in the developing retina and brain), morphogenetic cell death (in the neural tube, lens, skull, face and limbs) and phylogenetic cell death (elimination of the hyaloid artery system in the developing eye).

The apoptotic machinery or apoptosome, which involves Apaf1 and many other anti- and pro-apoptotic factors such as Bcl-X_L, Bax and other caspases, could contribute to the pathogenesis of some neurodegenerative diseases and could be involved in cancer suppression [62–64]. The identification and functional analysis of this gene performed by gene trap could have provided us with a new drug target for those diseases.

6.2. Netrin 1

By using the secretion trapping protocol, Skarnes et al. [14] generated a hypomorphic allele of the *netrin 1* gene which had previously been shown to be involved in axon guidance in the developing neural tube [65,66]. Consistent with these results, the *netrin 1* mutant mice have defects in several major commissures of the brain and show misguidance of the spinal commissural as well as retinal ganglion cell axons [28,67].

Salminen and coworkers [37] independently generated another *netrin 1* gene trap mouse line and focused on the importance of netrin 1 in inner ear development. Fig. 5G,H shows sections of wild-type and mutant inner ears of an e12.5 mouse embryo. In the mutant ear, the posterior semicircular duct (psd) does not form properly, due to the lack of local disruption in the basement membrane and to the epithe-

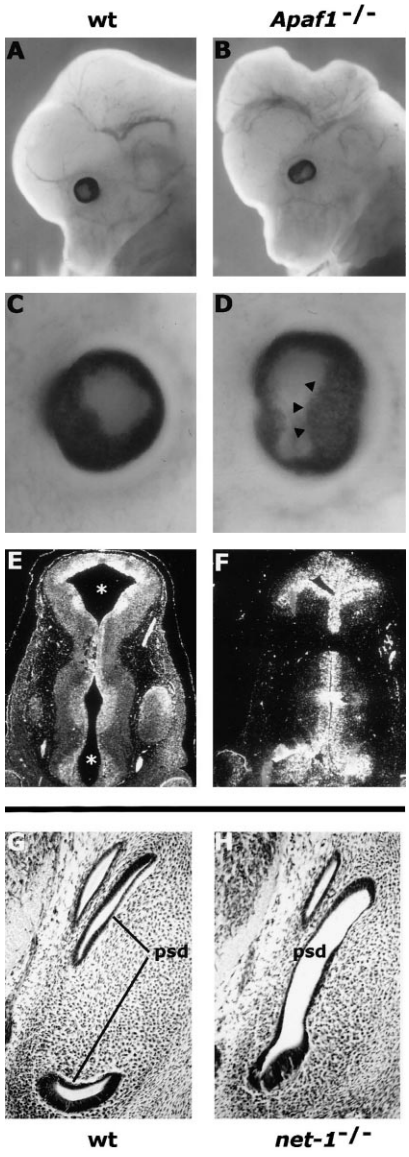

Fig. 5. Two phenotypes deriving from gene trap insertional mutagenesis. The trapped gene Apaf1 regulates apoptosis in mammalian development (A–F). The Apaf1 deficient phenotype exhibits dorsal exencephaly in a homozygous e14.5 embryo (B) compared with the wild-type littermate (A). The retinal hyperplasia in the same embryos is also shown. The arrowheads point to the external margin of retinal pigmented epithelium (C, D). The overgrowth of brain tissues in cross-sections of e12.5 embryos is due to proliferation of neural precursor cells, as shown by in situ hybridization with the proneural marker Delta-1. The asterisks show the telencephalic vesicles, occluded by cells in the mutant (E, F). The secreted molecule netrin 1 (net-1), also trapped by gene trap in our laboratory, is important for proper development of the inner ear (G, H). Comparison of transverse sections through the wild-type (G) and the mutant (H) inner ear of an e13.5 embryo shows impaired formation of the posterior semicircular duct (psd). Psd is part of a precisely organized epithelial labyrinth essential for balance.

lial–mesenchymal mis-signaling caused by *netrin 1* deficiency. These findings imply a new morphological role for netrin 1.

7. Concluding remarks

The production of Apaf1 and netrin 1 gene trap mutants represents a good example of how genetically modified ES cell libraries can be used to rapidly and efficiently generate and analyze mouse mutants. New vectors, designed for more specific approaches, together with novel prescreening techniques, will contribute to optimize this strategy by reducing workload and time consumption. The large-scale efforts carried out worldwide to create gene trap libraries, which could saturate the mammalian genome, will increase the ability to study gene activity in depth, and will provide us with the compelling perspective needed to associate such genes to human diseases.

Acknowledgements: We would like to thank Peter Gruss for continuous support and encouragement. We are also grateful to I. Stoltmann for critically reading the manuscript and M. Salminen for the images of the *netrin 1* mutant presented in Fig. 5. F.C. was supported by a Human Frontiers Science Program (HFSP) fellowship and is a Marie Curie Research fellow (TMR Program of the European Union). This work was supported by Amgen, Inc. (Thousand Oaks, CA, USA) and by the Max Planck Society.

References

[1] Capecchi, M.R. (1989) Science 244, 1288–1292.
[2] Brandon, E.P., Idzerda, R.L. and McKnight, G.S. (1995) Curr. Biol. 5, 873–881.
[3] Brandon, E.P., Idzerda, R.L. and McKnight, G.S. (1995) Curr. Biol. 5, 758–765.
[4] Brandon, E.P., Idzerda, R.L. and McKnight, G.S. (1995) Curr. Biol. 5, 625–634.
[5] Friedrich, G. and Soriano, P. (1991) Genes Dev. 5, 1513–1523.
[6] Gossler, A., Joyner, A.L., Rossant, J. and Skarnes, W.C. (1989) Science 244, 463–465.
[7] Skarnes, W.C., Auerbach, B.A. and Joyner, A.L. (1992) Genes Dev. 6, 903–918.
[8] Wurst, W., Rossant, J., Prideaux, V., Kownacka, M., Joyner, A., Hill, D.P., Guillemot, F., Gasca, S., Cado, D., Auerbach, A. and Ang, S.-L. (1995) Genetics 139, 889–899.
[9] Bronchain, O.J., Hartley, K.O. and Amaya, E. (1999) Curr. Biol. 9, 1195–1198.
[10] Pascoe, W.S., Kemler, R. and Wood, S.A. (1992) Biochim. Biophys. Acta 1114, 209–221.
[11] Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[12] Friedrich, G. and Soriano, P. (1993) Methods Enzymol. 225, 681–701.
[13] Robertson, E., Bradley, A., Kuehn, M. and Evans, M. (1986) Nature 323, 445–448.
[14] Skarnes, W.C., Moss, J.E., Hurtley, S.M. and Beddington, R.S. (1995) Proc. Natl. Acad. Sci. USA 92, 6592–6596.
[15] Chowdhury, K., Bonaldo, P., Torres, M., Stoykova, A. and Gruss, P. (1997) Nucleic Acids Res. 25, 1531–1536.
[16] Salminen, M., Meyer, B.I. and Gruss, P. (1998) Dev. Dyn. 212, 326–333.
[17] Rohwedel, J., Maltsev, V., Bober, E., Arnold, H.H., Hescheler, J. and Wobus, A.M. (1994) Dev. Biol. 164, 87–101.
[18] Wiles, M.V. et al. (2000) Nat. Genet. 24, 13–14.
[19] von Melchner, H. and Ruley, H.E. (1989) J. Virol. 63, 3227–3233.
[20] Takeuchi, T. (1997) Dev. Growth Differ. 39, 127–134.
[21] Yoshida, M., Yoshida, M., Yagi, T., Furuta, Y., Takayanagi, K., Kominami, R., Takeda, N., Tokunaga, T., Chiba, J., Ikawa, Y. and Aizawa, S. (1995) Transgenic Res. 4, 277–287.
[22] Niwa, H., Araki, K., Kimura, S., Taniguchi, S., Wakasugi, S. and Yamamura, K. (1993) J. Biochem. (Tokyo) 113, 343–349.
[23] Takeuchi, T., Yamazaki, Y., Katoh-Fukui, Y., Tsuchiya, R., Kondo, S., Motoyama, J. and Higashinakagawa, T. (1995) Genes Dev. 9, 1211–1222.
[24] von Melchner, H., DeGregori, J.V., Rayburn, H., Reddy, S., Friedel, C. and Ruley, H.E. (1992) Genes Dev. 6, 919–927.
[25] DeGregori, J., Russ, A., von Melchner, H., Rayburn, H., Priyaranjan, P., Jenkins, N.A., Copeland, N.G. and Ruley, H.E. (1994) Genes Dev. 8, 265–276.
[26] Natarajan, D. and Boulter, C.A. (1995) Nucleic Acids Res. 23, 4003–4004.
[27] Camus, A., Kress, C., Babinet, C. and Barra, J. (1996) Mol. Reprod. Dev. 45, 255–263.
[28] Serafini, T., Colamarino, S.A., Leonardo, E.D., Wang, H., Beddington, R., Skarnes, W.C. and Tessier-Lavigne, M. (1996) Cell 87, 1001–1014.
[29] Chen, H., Bagri, A., Zupicich, J.A., Zou, Y., Stoeckli, E., Pleasure, S.J., Lowenstein, D.H., Skarnes, W.C., Chedotal, A. and Tessier-Lavigne, M. (2000) Neuron 25, 43–56.
[30] Thorey, I.S., Muth, K., Russ, A.P., Otte, J., Reffelmann, A. and von Melchner, H. (1998) Mol. Cell Biol. 18, 3081–3088.
[31] Wood, S.A., Allen, N.D., Rossant, J., Auerbach, A. and Nagy, A. (1993) Nature 365, 87–89.
[32] Deng, J.M. and Behringer, R.R. (1995) Transgenic Res. 4, 264–269.
[33] Townley, D.J., Avery, B.J., Rosen, B. and Skarnes, W.C. (1997) Genome Res. 7, 293–298.
[34] Steel, M., Moss, J., Clark, K.A., Kearns, I.R., Davies, C.H., Morris, R.G., Skarnes, W.C. and Lathe, R. (1998) Hippocampus 8, 444–457.
[35] Skarnes, W.C. (1993) Curr. Opin. Biotechnol. 4, 684–689.
[36] Cecconi, F., Alvarez-Bolado, G., Meyer, B.I., Roth, K.A. and Gruss, P. (1998) Cell 94, 727–737.
[37] Salminen, M., Meyer, B.I., Bober, E. and Gruss, P. (2000) Development 127, 13–22.
[38] Torres, M., Torres, M., Stoykova, A., Huber, O., Chowdhury, K., Bonaldo, P., Mansouri, A., Butz, S., Kemler, R. and Gruss, P. (1997) Proc. Natl. Acad. Sci. USA 94, 901–906.
[39] Doetschman, T.C., Eistetter, H., Katz, M., Schmidt, W. and Kemler, R. (1985) J. Embryol. Exp. Morphol. 87, 27–45.
[40] Wiles, M.V. and Keller, G. (1991) Development 111, 259–267.
[41] Wang, R., Clark, R. and Bautch, V.L. (1992) Development 114, 303–316.
[42] Miller-Hance, W.C., LaCorbiere, M., Fuller, S.J., Evans, S.M., Lyons, G., Schmidt, C., Robbins, J. and Chien, K.R. (1993) J. Biol. Chem. 268, 25244–25252.
[43] Reddy, S., Rayburn, H., von Melchner, H. and Ruley, H.E. (1992) Proc. Natl. Acad. Sci. USA 89, 6721–6725.
[44] Shirai, M., Miyashita, A., Ishii, N., Itoh, Y., Satokata, I., Watanabe, Y.G. and Kuwano, R. (1996) Zool. Sci. 13, 277–283.
[45] Rijkers, T. and Ruther, U. (1996) Biochim. Biophys. Acta 1307, 294–300.
[46] Baker, R.K., Haendel, M.A., Swanson, B.J., Shambaugh, J.C., Micales, B.K. and Lyons, G.E. (1997) Dev. Biol. 185, 201–214.
[47] Yang, W., Musci, T.S. and Mansour, S.L. (1997) Hear. Res. 114, 53–61.
[48] Kroll, K.L. and Amaya, E. (1996) Development 122, 3173–3183.
[49] Pain, B., Clark, M.E., Shen, M., Nakazawa, H., Sakurai, M., Samarut, J. and Etches, R.J. (1996) Development 122, 2339–2348.
[50] Zambrowicz, B.P. and Friedrich, G.A. (1998) Int. J. Dev. Biol. 42, 1025–1036.
[51] Zambrowicz, B.P., Friedrich, G.A., Buxton, E.C., Lilleberg, S.L., Person, C. and Sands, A.T. (1998) Nature 392, 608–611.
[52] Chen, Z., Friedrich, G.A. and Soriano, P. (1994) Genes Dev. 8, 2293–2301.
[53] Ross, A.J., Waymire, K.G., Moss, J.E., Parlow, A.F., Skinner, M.K., Russell, L.D. and MacGregor, G.R. (1998) Nat. Genet. 18, 251–256.
[54] Cecconi, F. (1999) Cell Death Differ. 6, 1087–1098.
[55] Hengartner, M.O., Ellis, R.E. and Horvitz, H.R. (1992) Nature 356, 494–499.
[56] Yuan, J.Y. and Horvitz, H.R. (1990) Dev. Biol. 138, 33–41.
[57] Yuan, J. and Horvitz, H.R. (1992) Development 116, 309–320.
[58] Zou, H., Henzel, W.J., Liu, X., Lutschg, A. and Wang, X. (1997) Cell 90, 405–413.

[59] Li, P., Nijhawan, D., Budihardjo, I., Srinivasula, S.M., Ahmad, M., Alnemri, E.S. and Wang, X. (1997) Cell 91, 479–489.
[60] Kuida, K., Zheng, T.S., Na, S., Kuan, C., Yang, D., Karasuyama, H., Rakic, P. and Flavell, R.A. (1996) Nature 384, 368–372.
[61] Kuida, K., Kuida, K., Haydar, T.F., Kuan, C.Y., Gu, Y., Taya, C., Karasuyama, H., Su, M.S., Rakic, P. and Flavell, R.A. (1998) Cell 94, 325–337.
[62] Sanchez, I., Xu, C.J., Juo, P., Kakizaka, A., Blenis, J. and Yuan, J. (1999) Neuron 22, 623–633.
[63] Gervais, F.G., Gervais, F.G., Xu, D., Robertson, G.S., Vaillancourt, J.P., Zhu, Y., Huang, J., LeBlanc, A., Smith, D., Rigby, M., Shearman, M.S., Clarke, E.E., Zheng, H., Van Der Ploeg, L.H., Ruffolo, S.C., Thornberry, N.A., Xanthoudakis, S., Zamboni, R.J., Roy, S. and Nicholson, D.W. (1999) Cell 97, 395–406.
[64] Soengas, M.S., Alarcon, R.M., Yoshida, H., Giaccia, A.J., Hakem, R., Mak, T.W. and Lowe, S.W. (1999) Science 284, 156–159.
[65] Kennedy, T.E., Serafini, T., de la Torre, J.R. and Tessier-Lavigne, M. (1994) Cell 78, 425–435.
[66] Serafini, T., Kennedy, T.E., Galko, M.J., Mirzayan, C., Jessell, T.M. and Tessier-Lavigne, M. (1994) Cell 78, 409–424.
[67] Deiner, M.S., Kennedy, T.E., Fazeli, A., Serafini, T., Tessier-Lavigne, M. and Sretavan, D.W. (1997) Neuron 19, 575–589.
[68] Voss, A.K., Thomas, T. and Gruss, P. (2000) Development 127, 1–11.