PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Genome Manipulation in Embryonic Stem Cells

J. Rossant, C. Bernelot Moens and A. Nagy

Phil. Trans. R. Soc. Lond. B 1993 **339**, 207-215
doi: 10.1098/rstb.1993.0018

Genome manipulation in embryonic stem cells

J. ROSSANT[1,2], C. BERNELOT MOENS[1,2] AND A. NAGY[2]

[1] *Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Canada M5G 1X5*
[2] *Dept. of Molecular and Medical Genetics, University of Toronto, Toronto, Canada*

SUMMARY

Embryonic stem (ES) cells derived from pluripotent cells of the early mouse embryo provide a powerful tool for genome manipulation in mammals. Dominantly acting effects can be achieved by introducing constructs to misexpress or ectopically express a gene product, express an altered product or express antisense constructs. Use of ES cell chimeras to analyse the effects of such alterations may provide information not readily available from transgenic mice. However, the most important use of ES cells, to date, is in the generation of recessive mutations, either in known genes by targeted mutagenesis or randomly by insertional mutagenesis. Examples of these approaches and possible future strategies are discussed.

1. INTRODUCTION

Successful application of a genetic approach to studying any biological question involves integrating information from the phenotypes of specific mutations with knowledge of the nature of the gene product. In mammals, the mouse has provided a storehouse of useful mutations, accumulated spontaneously or as a result of radiation and other mutagenic strategies. These mutations have included phenotypes indicative of roles for the wild-type genes in development, neurological function, immunobiology as well as models of human genetic diseases. However, few of the genes underlying these existing mutations have yet been cloned. The number of mutations available is still limited, and although progress in this area is accelerating with the advent of large-scale genome characterization, the pace of molecular cloning has meant that there are many more cloned genes for which there is no known mutant phenotype, than there are mutations for which the gene product is known. Recent advances in manipulating the mouse genome have opened up new possibilities for investigating gene function, based on directed mutagenesis or novel insertional mutagenic strategies.

Although introduction of DNA constructs into transgenic mice has been an important advance in genetic manipulation, a major technical breakthrough came with the advent of embryonic stem (ES) cells. In 1981, two groups independently described the derivation of permanent cell lines from the inner cell mass of the blastocyst (Martin 1981; Evans & Kaufman 1981). These cell lines resembled previously described cell lines derived from teratocarcinomas in their ability to proliferate indefinitely in an undifferentiated state in culture, but differentiate into a variety of cell types either in ectopic sites or under altered culture conditions *in vitro*. Most importantly, however, it was shown that ES cells could contribute efficiently to both somatic and germ line tissues after reintroduction into the blastocyst (Bradley *et al.* 1984). Although there had been occasional reports of germ-line colonization from teratocarcinoma cells (Stewart & Mintz 1981), this system had not proved reproducible enough for widespread use. At the same time as the ES cell technology was being developed, other laboratories, most notably those of Oliver Smithies and Mario Capecchi, were showing that targeted mutagenesis via homologous recombination in mammalian cells was an attainable goal (Smithies *et al.* 1985; Thomas *et al.* 1986). In 1987 it was demonstrated that targeted mutations could be introduced into ES cells in culture (Thomas & Capecchi 1987; Doetschman *et al.* 1988). Germ line transmission of such mutations was first reported in 1989 (Thompson *et al.* 1989), and since that time there has been an explosion of interest in applying this technology to the study of the function of many different sorts of gene products. Targeted mutagenesis is not the only new avenue of genetic investigation opened up by the availability of ES cells. Dominant gain-of-function gene constructs can be introduced into ES cells and their effects studied in chimeras. Identification of novel mutations using efficient insertional mutagens is aided by the ability to screen the ES cells in culture for the appropriate insertions. The future of mouse genetics seems sure to involve the continued development of genomic manipulation in ES cells. In this paper, we describe evidence for the pluripotentiality of ES cells, and then review the kinds of mutagenic strategies, both directed and non-directed, that are currently being applied to ES cells. Finally, some possible future genetic approaches using ES cells will be discussed.

Phil. Trans. R. Soc. Lond. B (1993) **339,** 207–215
Printed in Great Britain

2. PROPERTIES OF ES CELLS

ES cells can be maintained in an undifferentiated state in culture indefinitely if maintained on feeder layers of inactivated fibroblasts or in the presence of leukemia inhibitory factor (LIF) (Smith *et al.* 1988; Sanchez *et al.*, 1991). If removed from feeders or LIF, they will differentiate spontaneously into a variety of cell types, including endoderm, muscle and hemopoietic cell types (Doetschman *et al.* 1985; Sanchez *et al.* 1991). They will change their morphology when treated with agents such as retinoic acid, but the full range of differentiation that they can undertake *in vitro* has not yet been well characterized. Most effort has been placed into trying to drive ES cells down the hematopoietic lineage, in the hope of obtaining permanent cell lines capable of repopulating the adult hematopoietic system (Lindenbaum & Grosveld 1990; Wiles & Keller 1991; Chen 1992). Despite this evidence that ES cells *in vitro* have multiple potentialities, these studies cannot prove that the cells are truly pluripotent like the ICM cells from which they derive (Gardner 1968).

The best test of the potential of ES cells is their ability to colonize tissues in chimeras. Despite the widespread use of ES cells, there are actually few studies which provide a detailed analysis of the distribution of ES progeny in chimeras: most studies focus on coat colour chimerism and the ability of cells to colonize the germ line. Existing studies indicate that ES cells can make widespread contributions to adult and embryonic tissues (Bradley *et al.* 1984; Williams *et al.* 1988; Kadokawa *et al.* 1990) and limited contributions to extraembryonic tissues (Beddington & Robertson 1989). A direct comparison of the mosaicism resulting from introduction of ES versus ICM cells into blastocysts has shown that both cell types have equal potential to contribute to all somatic tissues of the resulting mice (Rossant *et al.* 1992). Any deficiencies in ES contributions to certain tissues were shown to be due to effects of the particular genotype contributions used and not to deficiencies in ES potential. The only real difference between ES cells and ICM cells in this assay was the tendency for an increase in mortality with extensive ES contribution to the chimeras.

A clear-cut demonstration of the pluripotentiality of ES cells and the association between high ES controbution and mortality was provided by experiments in which the development of ES- or ICM-tetraploid chimeras was compared (Nagy *et al.* 1990). Tetraploid mouse embryos can develop readily to the blastocyst stage but post-implantation development is poor (Snow 1975; Kaufman & Webb 1990). Embryos rarely proceed beyond early somite stages. However, in combination with diploid cells, tetraploid cells can support development of extraembryonic lineages (Tarkowski *et al.* 1978). When ES cells were aggregated with tetraploid embryos, about 15% of aggregates proceeded to full term and the foetuses recovered were almost all entirely ES-derived. Tetraploid cells made up all of the trophoblast and endoderm layers of the yolk sacs. ICM-tetraploid aggregates developed similarly.

Production of entirely ES-derived fetuses provides a clear demonstration of the full potential of such cells. However, using established ES cell lines such as D3 and AB-1, no entirely ES-derived fetus was able to survive beyond one or two days after birth although entirely ICM-derived offspring were viable. The reasons for this are still unclear, but recent experiments with a newly established cell line, R1, demonstrate that the effect must be due to genetic or epigenetic alterations accumulating with continued passage in culture. Viable entirely ES-derived offspring were produced from about 10% of ES-tetraploid aggregates when early passages (passages 8–15) of this cell line were used (A. Nagy, R. Nagy, W. Abramow-Newerley, J. Roder and J. Rossant, in preparation). However, later passages produced the same non-viable fetuses as observed with other cell lines. At least one sub-clone derived from passage 14 was also capable of generating viable offspring suggesting that some fully pluripotent cells persist within the ES cell population even at late passage.

These chimera experiments demonstrate that it is valid to study the behaviour of genetically manipulated ES cells in chimeras or tetraploid aggregates as well as studying the effects of genetic alterations after germ line transmission.

3. INTRODUCTION OF GENETIC ALTERATIONS IN ES CELLS

(*a*) *Directed strategies*

Once a gene has been molecularly characterized and both cDNA and genomic clones are available it is possible to devise a number of strategies to address its function in the context of the whole organism. Approaches that do not require disruption of the endogenous gene include overexpression or ectopic expression of an activated form of the protein, and introduction of antisense or dominant negative constructs designed to interfere with endogenous gene function. All of these approaches can be achieved via standard transgenesis, although the ES route may have some advantages in situations where the predicted outcome is embryonic death or gross abnormality. Accumulating such information via analysing midgestation transgenic mice can be slow. However, ES cell lines carrying the dominant-acting construct of interest can be used to make large numbers of diploid or tetraploid chimeras which can then be studied for phenotypic effects. This approach has been used to study the effect of overexpression of v-*src* on embryonic development (Boulter *et al.* 1991) but has not yet been fully exploited.

Most attention has been directed towards the use of ES cells to introduce recessive loss-of-function mutations into genes of interest by homologous recombination. The common strategies for targeted mutagenesis (Capecchi 1989) all involve vectors which disrupt the endogenous allele by introduction of a selectable marker, such as the neomycin resistance

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

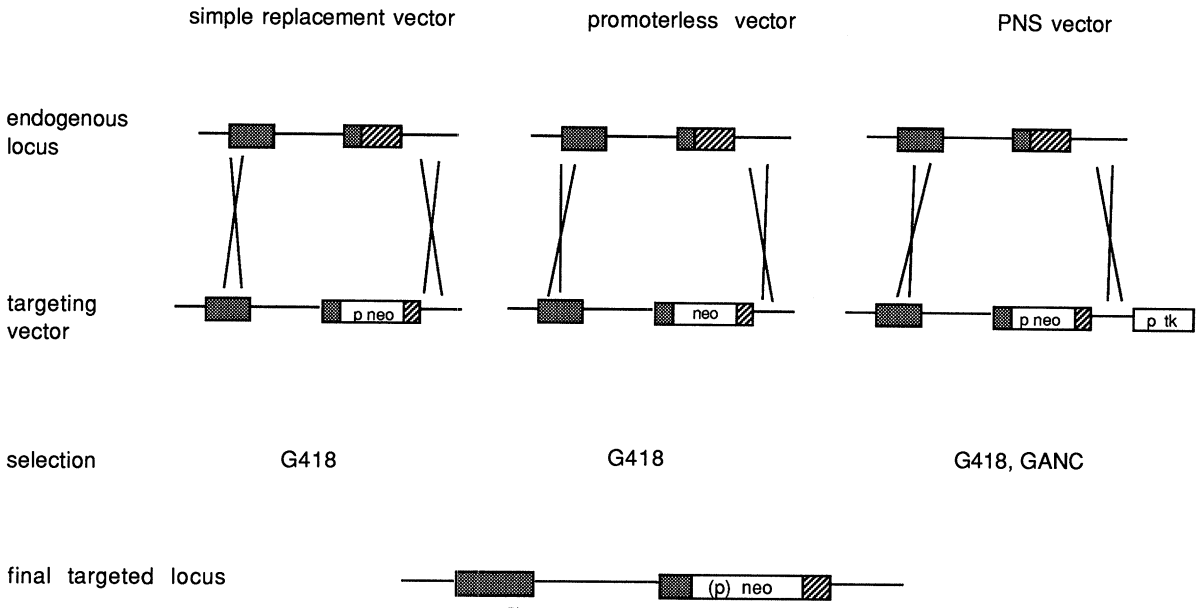

Figure 1. General strategies for targeted mutagenesis using replacement-type vectors. Abbreviations: PNS = positive–negative selection; GANC = gancyclovir; tk = herpes simplex thymidine kinase; neo = neomycin (G418) resistance gene; P = promoter.

gene, into the coding sequence. Identification of a targeted versus non-targeted integration of the targeting vector involves PCR and Southern analysis; enrichment for targeted events can be achieved by positive–negative selection (PNS) (Mansour *et al.* 1988) or promoterless constructs (Charron *et al.* 1990; Te Riele *et al.* 1990) (figure 1). Successful disruption of many genes has now been reported and the phenotypes of such mutations described. It is impossible to review all the results but some general comments can be made.

First, reported targeting efficiencies vary widely, especially when expressed as the number of targeted colonies per *neo*^{r} colonies. Reported efficiencies vary from as high as 1/30 to $<1/5000$ for standard PNS vectors. Using promoterless constructs for genes expressed in ES cells, efficiencies as high as 85% have been reported (Te Riele *et al.* 1990). All the reasons for this variation are not clear but some factors have been identified. A relationship between length of homology in the targeting vector and targeting efficiency has been reported (Thomas & Capecchi 1987). There was also some early indication that expressed genes might be easier to target than non-expressed genes. Successful mutagenesis of several genes apparently not expressed in ES cells has tended to argue against the importance of transcriptional activity (Johnson *et al.* 1989). The use of targeting vectors derived from genomic DNA of the same mouse strain as the ES cells used has had a dramatic effect on targeting efficiencies at particular loci. For example, a 20-fold improvement in targeting efficiency was obtained at the *Rb* locus when a vector derived from 129 rather than BALB/c DNA was used (Te Riele *et al.* 1992). This effect is thought to be due to small sequence differences in non-transcribed DNA between mouse strains. The full impact of the effect of using isogenic DNA in targeting vectors has not yet been felt, but the consistently high efficiencies reported with such isogenic vectors suggests that this may be one of the most important factors required for efficient homologous recombination.

Second, phenotypes of targeted mutations have not always been as predicted from knowledge of the nature of the gene product and its pattern of expression. The production of mice lacking cytotoxic T-cells after disruption of the CD8 gene (Fung-Leung *et al.* 1991) and mice lacking mature lymhocytes after discruption of the RAG-1 and RAG-2 V(D)J recombinase genes (Mombaerts *et al.* 1992; Shinkai *et al.* 1992) provide examples where prediction and actual outcome were very close. In other cases, phenotypes appear to reflect a subset of the expression domains of the gene. For example, loss of *Wnt*-1 causes deletion of part of the brain but has no effect on the developing spinal cord where the gene is also expressed (McMahon & Bradley 1990; Thomas & Capecchi 1990). In extreme cases, gene disruption may result in no obvious phenotype, as has been reported for the p53 (Donehower *et al.* 1992) and *PrP* genes (Lang *et al.* 1992) (although p53^{-} mice are more susceptible to tumours). A full explanation for this rather unpredictable relationship between expression and phenotype cannot yet be provided but is thought to derive from complete or partial redundancies in the gene function either at the single gene level or the genetic pathway level. Sorting out these genetic interactions will require more complete knowledge of genetic pathways and analysis of the effects of many possible combinations of mutations.

All the germ-line mutations described to date have been designed to generate a null mutation in the gene of interest. It is clearly important also to be able to generate more subtle mutations in order to fully dissect gene function. Site-specific mutations within the coding sequence may help elucidate the functions of specific protein domains and mutations in the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

regulatory sequences may be used to alter gene function in subsets of the gene's normal expression domains. Ideally such mutations should introduce the specific sequence changes without the insertion of a selectable marker gene which could, in itself, affect gene expression. A technique that allows such mutations has recently been described using vectors called 'hit-and-run' (Hasty *et al.* 1991) or 'in–out' (Valancius & Smithies 1991). In this approach an insertion vector containing a positively and negatively selectable marker and the small mutation of interest is introduced into cells. Selection against the marker gene favors the retention of cells in which the vector has recombined out of the cells again, leaving the subtle mutation behind in a proportion of cases. This approach has not yet been extensively used and so the efficiency at different loci is still not clear. It may be possible to devise strategies that favour the 'out' event more strongly by incorporating specific recognition sequences for site-specific recombinases, such as the FLP recombinase from yeast (Golic & Lindquist 1989) and *Cre* (Dale & Ow 1991), into the vectors and causing efficient excision of the marker genes by transient expression of the appropriate recombinase. The use of such recombinases in mammalian cells is just in its infancy (O'Gorman *et al.* 1991) and seems likely to have multiple uses for genome manipulation.

The importance of having multiple alleles at any locus for full understanding of gene function is exemplified by many studies in *Drosophila*, and by a few examples in mammals, such as the *W* locus, where alleles with varying severity contain different site-specific mutations (Reith *et al.* 1990). We have recently shown that a leaky mutation targeted to the N-*myc* locus produces a different phenotype to the null mutation when homozygous (Bernelot-Moens *et al.* 1992). N-*myc* is a nuclear protooncogene in the b-HLH *myc* gene family, and is expressed at high levels in a variety of cell lineages during mouse embryogenesis (Zimmerman *et al.* 1986; Mugrauer *et al.* 1988). There is still very little information on the function of *myc* genes in cells and so there has been considerable interest in mutating them to see if this can shed light on their possible roles in tissue-specific transcriptional regulation. We and others (Charron *et al.* 1990; Stanton *et al.* 1990; Sawai *et al.* 1991) have generated null mutations in the N-*myc* allele *in vitro* using a replacement-type vector that disrupts the coding sequence at the start of translation (figure 2). Mice homozygous for such mutations die in mid-gestation (*ca.* day 9–10) at a stage when elucidation of lineage-specific effects is difficult (B. Stanton & L. Parada, personal communication; J. Charron & F. Alt, personal communication).

We generated another targeted cell line in which the expected gene replacement did not occur but instead the *neo*-containing vector inserted into the first intron of N-*myc* (figure 2). This allows the possible generation of some normal N-*myc* transcripts if mRNA splicing bypasses the *neo*-containing exon. Mice homozygous for this potentially leaky mutation die at birth because of an inability to oxygenate their lungs. This is due to a failure of proliferation of the epithelium during branching morphogenesis of the lungs, indicat-

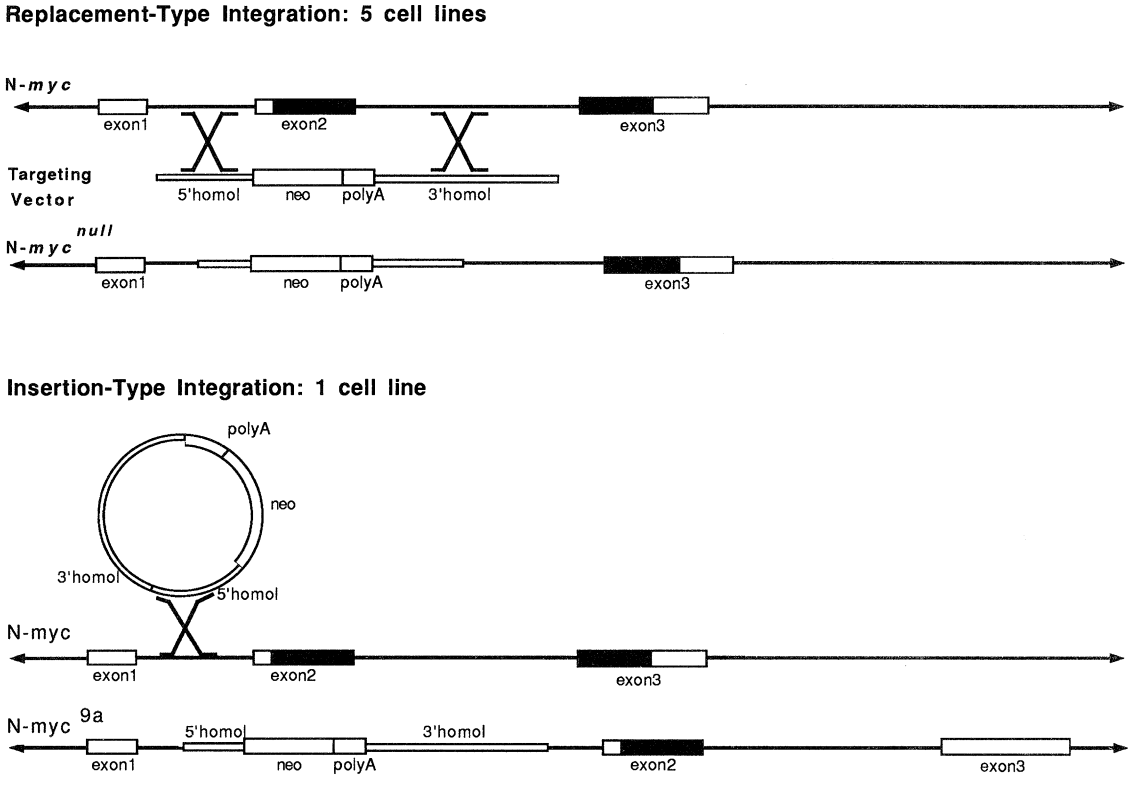

Figure 2. Outcome of gene targeting at N-*myc* locus using a promoterless *neo* vector. In five lines, the targeting vector integrated by the expected double cross-over, replacing exon 2 with the *neo* gene. In one line, the replacement vector recircularized and inserted by a single cross-over in intron 1, resulting in the N-*myc*^{9a} allele shown.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 1. *Phenotypes of N-myc mutations*

	time of death	phenotype
N-*myc*$^{null/null}$	9–10 dpc	failure of general cell proliferation
N-*myc*$^{9a/9a}$	term	failure of epithelial proliferation in lung
N-*myc*$^{9a/null}$	11–13 dpc	lung defect more pronounced: hypoplasia of myocardium

ing a possible role for N-*myc* in the downstream response to the inductive signals from the lung mesenchyme. Analysis of N-*myc* levels showed that there were some normal N-*myc* transcripts in all tissues analysed, varying from 15–50% of wildtype levels.

The effects of N-*myc* deficiency on lung development could not have been determined from the null mutant mice, since they die before lung morphogenesis has occurred. With this in mind, we reasoned that making double heterozygotes with the null and the leaky mutation might generate an intermediate phenotype and indicate other lineages where N-*myc* is required. Preliminary data from such an experiment indicate that such double heterozygotes die around day 12–13 of development (C. Moens, B. Stanton, L. Parada and J. Rossant, unpublished data). The effects of a further reduction of N-*myc* expression are manifest as an even more severe effect on lung branching but also as a defect in heart development. Specifically the ventricular myocardium is extremely thin and the fetuses appear to die because of inefficient blood circulation. This result indicates an unexpected role for N-*myc* in development of cardiac muscle (table 1).

These results indicate that hypomorphic alleles can be very useful in elucidating gene function at different stages of development. However, insertions of targeting vectors into introns is not a guaranteed way of making leaky or hypomorphic mutations. Use of more directed approaches to generating an allelic series of mutations at specific loci would seem to be an important task for the coming years.

(*b*) *Non-directed strategies*

Targeted mutagenesis can provide insights into the roles of already cloned genes but it is unlikely that all the genes involved in any process are currently available as cloned products. It is important therefore, to continue to search for new genes and mutations affecting any particular process of interest. In species such as *Drosophila* large-scale mutagenic screens have been successfully carried out using efficient chemical mutagens, and sets of mutations affecting different stages of development have been characterized (Jurgens *et al.* 1984; Wieschaus *et al.* 1984). In mouse, the size of the animals and the limited number of offspring make such strategies extemely slow and expensive. Attention has turned towards the use of insertional mutagenesis which, although less efficient than chemical mutagenesis, has the advantage that the inserted DNA sequences can act as a tag for cloning the mutated endogenous gene.

There are a variety of approaches for inducing insertional mutagenesis in mice (Gridley *et al.* 1987). Microinjection of DNA into zygotes (Jaenisch 1988) or retroviral infection of embryos (Jaenisch 1976) has proved to be an effective means of generating novel mutations, and, at least for retroviral insertions, cloning the host gene has been relatively easy (Gridley *et al.* 1991; Jaenisch *et al.* 1983; Weiher *et al.* 1990). The insertion of DNA by microinjection tends to cause large rearrangements of the host DNA which can complicate identification of the mutated gene (Covarrubias *et al.* 1987; Wilkie & Palmiter 1987), but successful cloning of the genes mutated by such transgenic insertions has been reported (Maas *et al.* 1990; Lee *et al.* 1992).

Insertional mutagenesis in ES cells offers several advantages over standard approaches. Using standard insertional mutagens, such as retroviruses (figure 3), it is possible to introduce multiple insertions into single ES cells and populate the germline of chimeras with multiple ES cells (Robertson *et al.* 1986). This results in an increased number of mutagenic events per gamete tested. This approach has been pursued in the laboratory of Dr Martin Evans (Cambridge) and has led to the identification of a number of new mutations (see Conlon *et al.* 1991).

More recently different types of vectors have been used which are designed to allow preselection for integration in or near expressed genes, thus increasing the efficiency of mutagenesis. Three general types of vector have been described, all of which use a reporter gene to allow visualization of gene expression (figure 3). Enhancer trap vectors, containing the *E. coli* β-galactosidase (*lacZ*) gene under the control of a weak promoter, have been introduced into ES cells by electroporation, and expression of *lacZ* has been studied in chimeras (Gossler *et al.* 1989). A recent study of 59 enhancer trap lines has shown that this is a very efficient means of observing novel patterns of gene expression: 1/3 of all lines containing the vector showed *lacZ* expression patterns in chimeras during early development (A. Gossler, personal communication). If these ES cell lines are introduced into the germline they can provide a rich source of novel *in situ* markers for many studies. Presumably because the vector is introduced by electroporation rather than microinjection, the integrations are clean and it has proved possible to identify transcripts closely associated with 3/3 integrations studied so far (A. Gossler, personal communication). However, because integrations are not necessarily within host transcribed sequences, expression of the reporter may not reflect the full expression pattern of the host gene and the integration need not be mutagenic.

Gene- or promoter-trap vectors have more stringent requirements for activation of *lacZ* expression, namely that the integration be within the transcribed sequences of a gene. Promoter-trap vectors contain *lacZ* with an initiation codon but no promoter (Reddy *et al.* 1991). Expression requires insertion into an exon of a gene transcribed in ES cells, either in frame in a translated exon, or in any 5′ untranslated exons. Activation is thus a relatively rare event. A higher

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

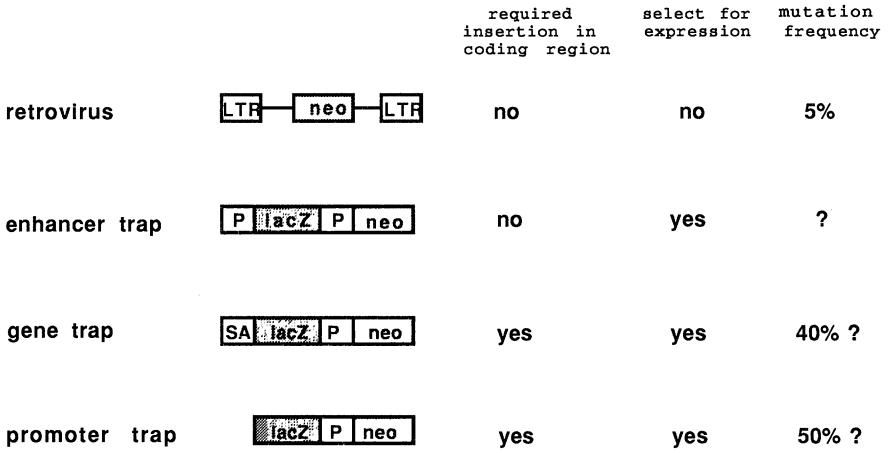

Figure 3. Insertional mutagenesis in ES cells. Abbreviations: LTR = long terminal repeat; P = promoter; SA = splice acceptor.

frequency of activation can be achieved using gene-trap vectors which contain the reporter gene with a splice acceptor sequence upstream (Gossler *et al.* 1989; Brenner *et al.* 1989; Friedrich & Soriano, 1991; Kerr *et al.* 1989). Integrations within introns can allow activation of *lacZ* if there is correct inframe splicing from host exon to the *lacZ* sequence. Both kinds of vectors are likely to cause disruption of host gene transcription and thus be highly mutagenic. This has been confirmed by three recent studies which together suggest that at least 1/3 of all integrations show a readily visible mutant phenotype (von Melchner *et al.* 1992; Friedrich & Soriano 1991; Skarnes *et al.* 1992). Another advantage of these vectors is that expression of the reporter is likely to reflect that of the host gene since *lacZ* is being produced as a fusion transcript with the endogenous gene. Also, cloning the associated exons can be achieved by RACE (Skarnes *et al.* 1992), rather than requiring the cloning of genomic flanking sequences.

All these factors have combined to make the use of these vectors, either introduced by electroporation or in retroviruses, the most efficient means of identifying and mutating novel genes expressed during early

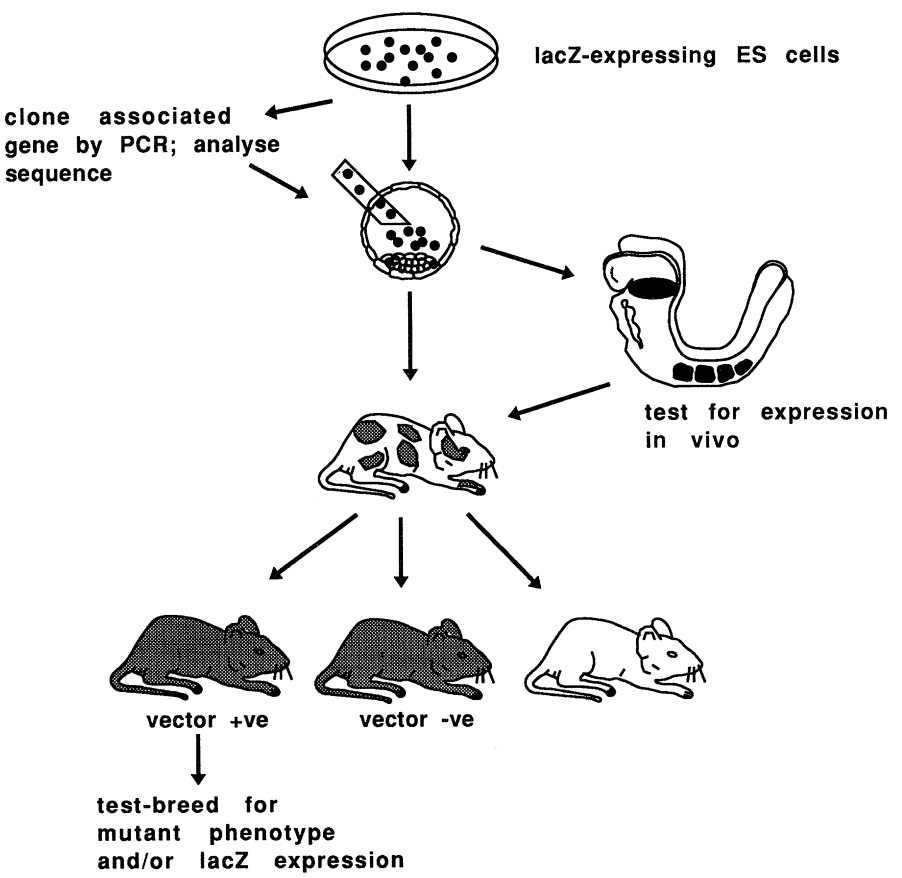

Figure 4. Strategies for large-scale gene-trap screens.

embryogenesis. Various strategies for large-scale screening can be envisaged, using different criteria to preselect for integrations of interest (figure 4). One could clone associated exons and use sequence information as the criterion for further study. Alternatively, expression in specific patterns during embryogenesis could be the method of choice. This is the approach we are using, taking advantage of the widespread contributions of ES cells after blastocyst injection to screen for patterns of expression of individual clones around gastrulation in chimeras. Of 230 cell lines examined to date, 60% showed expression at 8.5 days of development and 1/3 of these showed spatially restricted patterns (S.-L. Ang, A. Auerbach, S. Gasca, F. Guillemot, D. Hill, A. Joyner, J. Rossant and W. Wurst, unpublished data). Alternatively, populations of *lacZ*-expressing ES cells could be introduced into chimeras such that male chimeras could carry several different integrations. Offspring could be scored for expression pattern. Finally, mutant phenotype could be the primary criterion, although the numbers that could be screened in this way are much lower than for a primary expression screen.

A major limitation of the gene trap approach is that the low efficiency of *lacZ* activation makes it not very feasible to screen for genes that are not expressed in ES cells. Although many genes that are involved in early embryonic development may be expressed in ES cells, it is unlikely that *lacZ*-expressing gene trap ES cell lines will be a very useful resource for those interested in identifying and mutating genes involved in later organogenesis and terminal cell differentiation. To explore these areas, the *in vitro* differentiation capacity of ES cells will need to be exploited. ES cells have been shown to differentiate reproducibly into hematopoietic lineages (Wiles & Keller 1991), vascular endothelium (Wang *et al.* 1992) and muscle (Sanchez *et al.* 1991) *in vitro*. More effort needs to be placed on establishing conditions for reproducibly directing the differentiation of ES cells into different pathways, so that one could screen for gene trap integrations that became activated upon differentiation. Other screens using gene trap integrations to visualize genetic responses to particular growth and differentiation factors can also be envisaged.

4. THE FUTURE OF MOUSE GENOME MANIPULATION IN ES CELLS

The currently available technologies in ES cells have opened up many new possibilities in mammalian genetics. Mutations can be generated at will in cloned genes thought to be involved in any biological process, including those involved in human genetic disease. Novel insertional mutations that can be readily cloned can also be generated in large numbers. However, the full extent of genetic manipulation possible with ES cells has certainly not been reached. We can expect to see further refinement of targeting technology to increase targeting efficiencies and allow easy generation of site-specific mutations. Targeting mutations to specific tissues using site-specific recombinases may also be possible. More sophisticated entrapment vectors for random mutagenesis could be envisaged as well.

Methods that rely on the mouse as the final assay system will continue to be slow and expensive. The beauty of ES cells is that they provide the opportunity for establishing genetic screens *in vitro* in cell culture, where there is no limit on numbers of events that can be screened. This fact has not been fully exploited as yet, but one can, for example, envisage generation of sets of ES lines with specific chromosome deletions or truncations, for genomic analysis and phenotype assessment. It should also be possible to devise strategies to screen for dominant suppressors or enhancers *in vitro*. Combined with improvements in directing the differentiation of ES cells *in vitro*, these approaches promise to provide a wealth of genetic information on mammalian development and differentiation.

We thank A. Joyner for continued input and useful discussion. The authors' own work was supported by grants from the Medical Research Council and the National Cancer Institute of Canada, the National Institutes of Health (U.S.A.) and Bristol–Myers Squibb Ltd. J.R. is an International Scholar of the Howard Hughes Medical Institute and a Terry Fox Cancer Research Scientist of the NCIC. C.M. holds an MRC studentship.

REFERENCES

Beddington, R.S.P. & Robertson, E.J. 1989 An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development* **105**, 733–737.

Bernelot-Moens, C., Auerbach, B.A., Conlon, R.A., Joyner, A.L. & Rossant, J. 1992 A targeted mutation reveals a role for N-*myc* in branching morphogenesis in the embryonic mouse lung. *Genes Dev.* **6**, 691–704.

Bradley, A., Evans, M. & Kaufman, M.H. 1984 Formation of germline chimaeras from embryo-derived teratocarcinoma cell lines. *Nature, Lond.* **309**, 255–256.

Brenner, D.G., Lin-Chao, S. & Cohen, S.N. 1989 Analysis of mammalian cell genetic regulation *in situ* by using retrovirus-derived "portable exons" carrying the *Escherichia coli LacZ* gene. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5517–5521.

Boulter, C.A., Aguzzi, A., Williams, R.L., Wagner, E.F., Evans, M.J. & Beddington, R. 1991 Aberrant development and twinning in chimaeric mice. *Development* **111**, 357–366.

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

Charron, J., Malynn, B.A., Robertson, E.J., Goff, S.P. & Alt, F.W. 1990 High-frequency disruption of the N-*myc* gene in embryonic stem and pre-B cell lines by homologous recombination. *Molec. Cell Biol.* **10**, 1799–1804.

Chen, U. 1992 Differentiation of mouse embryonic stem cells to lympho-hematopoietic lineages *in vitro*. *Devl. Immunol.* **2**, 29–50.

Conlon, F.L., Barth, K.S. & Robertson, E.J. 1991 A novel retrovirally induced embryonic lethal mutation in the mouse: Assessment of the developmental fate of embryonic stem cells homozygous for the 413.d proviral integration. *Development* **111**, 969–981.

Covarrubias, L., Nashida, Y., Terao, M., D'Eustachio, P. & Mintz, B. 1987 Cellular DNA rearrangements and early developmental arrest caused by DNA insertion in transgenic mouse embryos. *Molec. Cell Biol.* **7**, 2243–2247.

Dale, E.C. & Ow, D.W. 1991 Gene transfer with subsequent removal of the selection gene from the host genome. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10558–10562.

Doetschman, T., Maeda, N. & Smithies, O. 1988 Targeted mutation of the Hprt gene in mouse embryonic stem cells. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8583–8587.

Doetschman, T.G., Eistetter, M., Katz, M., Schmidt, W. & Kemler, R. 1985 *In vitro* development of blastocyst-derived embryonic stem cell lines: Formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. exp. Morph.* **87**, 27–45.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Jr, Butel, J.S. *et al.* 1992 Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature, Lond.* **356**, 215–222.

Evans, M. & Kaufman, M.H. 1981 Establishment in culture of pluripotential cells from mouse embryos. *Nature, Lond.* **292**, 154–155.

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.

Fung-Leung, W.-P., Schilham, A., Rahemtulla, A.M., Kundig, M., Vollenweider, J., Potter, J. *et al.* 1991 CD8 is needed for development of cytotoxic T cells but not helper T cells. *Cell* **65**, 443–449.

Gardner, R.L. 1968 Mouse chimaeras obtained by the injection of cells into the blastocyst. *Nature, Lond.* **220**, 596–597.

Golic, K.G. & Lindquist, S. 1989 The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. *Cell* **59**, 499–509.

Gossler, A., Joyner, A.L., Rossant, J. & Skarnes, W.C. 1989 Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science, Wash.* **244**, 463–465.

Gridley, T., Soriano, P. & Jaenisch, R. 1987 Insertional mutagenesis in mice. *Trends Genet.* **3**, 162–166.

Gridley, T., Gray, D.A., Orr-Weaver, T., Soriano, P., Barton, D.E., Francke, U. *et al.* 1991 Molecular analysis of the *Mov 34* mutation: transcript disrupted by proviral integration in mice is conserved in *Drosophila*. *Development* **109**, 235–242.

Hasty, P., Ramirez-Solis, R., Kramlauf, R. & Bradley, A. 1991 Introduction of a subtle mutation into the *Hox-2.6* locus in embryonic stem cells. *Nature, Lond.* **350**, 243–246.

Jaenisch, R. 1976 Germ line integration and Mendelian transmission of the exogenous Moloney murine leukemia virus. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1260–1264.

Jaenisch, R., Harbers, K., Schnicke, A., Lohler, J., Chumakov, I., Jahner, D. *et al.* 1983 Germline integration of Moloney murine leukemia virus at the *Mov 13* locus leads to recessive mutational and early embryonic death. *Cell* **32**, 209–216.

Jaenisch, R. 1988 Transgenic animals. *Science, Wash.* **240**, 1468–1674.

Johnson, R.S., Sheng, M., Greenberg, M.E., Kolodner, R.D., Papaioannou, V.E. & Spiegelman, B.M. 1989 Targeting of nonexpressed genes in embryonic stem cells via homologous recombination. *Science, Wash.* **245**, 1234–1236.

Jurgens, G., Wieschaus, E., Nusslein-Volhard, C. & Kluding, H. 1984 Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster II.* Zygotic loci on the third chromosome. *Wilhelm Roux's Arch.* **193**, 283–295.

Kadokawa, Y., Suemori, H. & Nakatsuji, N. 1990 Cell lineage analyses of epithelia and blood vessels in chimeric mouse embryos by use of an embryonic stem cell line expressing the β-galactosidase gene. *Cell Diff. Dev.* **29**, 187–194.

Kaufman, M.H. & Webb, S. 1990 Postimplantation development of tetraploid mouse embryos produced by electrofusion. *Development* **110**, 1121–1132.

Kerr, W.G., Nolan, G.P., Serafini, A.T. & Herzenberg, L.A. 1989 Transcriptionally defective retroviruses containing *lacZ* for the *in situ* detection of endogenous genes and developmentally regulated chromatin. *Cold Spring Harbor Symp. Quant. Biol.* **54**, 767–776.

Lang, Y., Bluethmann, H., Lipp, H.-P., DeArmond, S.J., Prusiner, S.B., Aguet, M. *et al.* 1992 Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature, Lond.* **356**, 577–582.

Lee, J.J., Radice, G., Perkins, C.P. & Costantini, F. 1992 Identification and characterization of a novel, evolutionarily conserved gene disrupted by the murine H*β*58 embryonic lethal transgene insertion. *Development* **115**, 277–288.

Lindenbaum, M.H. & Grosveld, F. 1990 An *in vitro* globin gene switching model based on differentiated embryonic stem cells. *Genes Dev.* **4**, 2075–2085.

Maas, R.L., Zeller, R., Woychik, R.P., Vogt, T.F. & Leder, P. 1990 Disruption of formin-encoding transcripts in two mutant *limb deformity* alleles. *Nature, Lond.* **346**, 853–855.

Mansour, S.L., Thomas, K.R. & Capecchi, M.R. 1988 Disruption of the proto-oncogene *int*-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature, Lond.* **336**, 348–352.

Martin, G.R. 1981 Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7634–7638.

McMahon, A.P. & Bradley, A. 1990 The *Wnt-1* (*int-1*) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**, 1073–1085.

Mombaerts, P., Iacomini, J., Johnson, R.S., Herrup, K., Tonegawa, S. & Papaioannou, V.E. 1992 RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **68**, 869–878.

Mugrauer, G., Alt, F.W. & Ekblom, P. 1988 N-*myc* proto-oncogene expression during organogenesis in the developing mouse as revealed by *in situ* hybridization. *J. Cell Biol.* **107**, 1325–1335.

Nagy, A., Gocza, E., Diaz, E.M., Prideaux, V.R., Ivanyi, E., Markkula, M. *et al.* 1990 Embryonic stem cells alone are able to support fetal development in the mouse. *Development* **110**, 815–822.

O'Gorman, S., Fox, D.T. & Wahl, G.M. 1991 Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science, Wash.* **251**, 1351–1355.

Reddy, S., DeGregori, J.V., von Melchner, H. & Ruley, H.E. 1991 Retrovirus promoter-trap vector to induce *lacZ* gene fusions in mammalian cells. *J. Virol.* **65**, 1507–1515.

Reith, A.D., Rottapel, R., Giddens, E., Brady, C., Forrester, L. & Bernstein, A. 1990 *W* mutant mice with mild or severe developmental defects contain distinct point mutations in the kinase domain of the c-*kit* receptor. *Genes Dev.* **4**, 390–400.

Robertson, E.J., Bradley, A., Kuehn, M. & Evans, M. 1986 Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature, Lond.* **323**, 445–447.

Rossant, J., Mercentes-Diaz, E., Gocza, E., Ivanyi, E. & Nagy, A. 1992 Developmental potential of mouse embryonic stem cells. In *Serono Symposium on Preimplantation*

Embryo Development (ed. B. Bavister). Springer Verlag. (In the press.)

Sanchez, A., Jones, W.K., Gulick, J., Doetschman, T. & Robbins, J. 1991 Myosin heavy chain gene expression in mouse embryoid bodies. An *in vitro* developmental study. *J. biol. Chem.* **266**, 22419–22426.

Sawai, S., Shimono, A., Hanaoka, K. & Kondoh, H. 1991 Embryonic lethality resulting from disruption of both N-myc alleles in mouse zygotes. *New Biol.* **3**, 861–869.

Shinkai, Y., Rathbun, G., Lam, K.-P., Ottz, E.M., Stewart, V., Mendelsohn, M. *et al.* 1992 RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearangement. *Cell* **68**, 855–868.

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

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M. *et al.* 1988 Inhibition of pluripotential embryonic stem cell differentiation by purified peptides. *Nature, Lond.* **336**, 688–690.

Smithies, O., Gregg, R.G., Boggs, S.S., Kolaweski, M.A. & Kucherlapati, R. 1985 Insertion of DNA sequences into the human chromosome β-globin locus by homologous recombination. *Nature, Lond.* **317**, 230–234.

Snow, M.H.L. 1975 Embryonic development of tetraploid mice during the second half of gestation. *J. Embryol. Exp. Morph.* **34**, 707–721.

Stanton, B.R., Reid, S.W. & Parada, L.F. 1990 Germ line transmission of an inactive N-myc allele generated by homologous recombination in mouse embryonic stem cells. *Molec. Cell Biol.* **10**, 6755–6758.

Stewart, T.A. & Mintz, B. 1981 Successive generations of mice produced from an established culture line of emploid teratocarcinoma cells. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6314–6318.

Tarkowski, A.K., Witkowska, A. & Opas, J. 1978 Development of cytochalasin β-induced tetraploid and diploid-tetraploid mosaic mouse embryos. *J. Embryol. exp. Morph.* **41**, 47–64.

Te Riele, H., Maandag, E.R., Clarke, A., Hooper, M. & Berns, A. 1990 Consecutive inactivation of both alleles of the *pim-1* proto-oncogene by homologous recombination. *Nature, Lond.* **348**, 649–651.

Te Riele, H., Robanus Maandag, E. & Berns, A. 1992 Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5128–5132.

Thomas, K.R., Folger, K.R. & Capecchi, M.R. 1986 High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419–428.

Thomas, K.R. & Capecchi, M.R. 1987 Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512.

Thomas, K.R. & Capecchi, M.R. 1990 Targeted disruption of the murine *int-1* proto-oncogene resulting in severe abnormalities in midbrain and cerebellar development. *Nature, Lond.* **346**, 847–850.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L. & Melton, D.W. 1989 Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**, 313–321.

Valancius, V. & Smithies, O. 1991 Testing an "in–out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. *Molec. Cell Biol.* **11**, 1402–1408.

von Melchner, H., DeGregori, J.V., Rayburn, H., Reddy, S., Friedel, C. & Ruley, H.E. 1992 Selective disruption of genes expressed in totipotent embryonal stem cells. *Genes Dev.* **6**, 919–927.

Wang, R., Clark, R. & Bautch, V.L. 1992 Embryonic stem cell-derived cystic embryoid bodies form vascular channels—an *in vitro* model of blood vessel development. *Development* **114**, 303–316.

Weiher, H., Noda, T., Gray, D.A., Sharpe, A.H. & Jaenisch, R. 1990 Transgenic mouse model of kidney disease: Insertional inactivation of ubiquitously expressed gene leads to nephrotic syndrome. *Cell* **62**, 425–434.

Wieschaus, E., Nusslein-Volhard, C. & Jurgens, G. 1984 Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster III.* Zygotic loci on the X-chromosome and fourth chromosome. *Wilhelm Roux's Arch.* **193**, 296–307.

Wiles, M.V. & Keller, G. 1991 Multiple hematopoietic lineages develop from embryonic stem (ES) cells in culture. *Development* **111**, 259–267.

Wilkie, T.M. & Palmiter, R.D. 1987 Analysis of the integrant of Myk-103 transgenic mice in which males fail to transmit the integrant. *Molec. Cell Biol.* **7**, 1646–1655.

Williams, R.L., Hilton, D.J., Pease, S., Wilson, T.A., Stewart, C.L., Gearing, D.P. *et al.* 1988 Myeloid leukaemia inhibitory factor maintains the development potential of embryonic stem cells. *Nature, Lond.* **336**, 684–687.

Zimmerman, K., Yancopoulos, G.D., Collum, R.G., Smith, R.K., Kohl, N.E., Denis, K.A. *et al.* 1986 Differential expression of myc family genes during murine development. *Nature, Lond.* **319**, 780–783.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES