1

Transgenic animals in perspective

NORMAN MACLEAN

1.1 Introduction

1.1.1 What are transgenic animals?

Since the earliest days of animal domestication, people have sought to improve their flocks and herds and companion animals by selective breeding. Although the results of such breeding are impressive, the procedure is by nature slow and to some extent imprecise. So the dream existed that perhaps one could take a more direct hand in stock improvement by recovering the genetic factors involved and adding these selectively to individual animals. The first attempts to do this took place some thirty years ago, but since the methods of DNA purification were suboptimal and methods had not been developed for gene isolation and cloning, complete genomic DNA preparations were used. These were simply mixed with animal eggs or embryos in the hope that transfer might occur, and then the resulting adults were screened for phenotypic features previously present only in the donors. Many of us can remember such experiments being undertaken with DNA from pigmented *Ambystoma* (the Mexican axolotl), such DNA being added to or injected into albino embryos of this species in the hope that the pigment gene would be transferred. Since that gene was likely to be swamped by the DNA for at least 1 million alternative sequences, one can see with the benefit of hindsight that the odds against success were enormous.

The developments which have now tipped the balance in our favour are chiefly threefold. The first, along with improvements in methods of DNA extraction and purification, was the discovery of restriction enzymes, which made it possible to dissect DNA into manageable fragments of precise length and sequence. Such fragments are easily separated and recovered from agarose gels.

The second development was the discovery of bacterial plasmids, which, together with bacterial viruses, provided easy vector systems into which genes could be inserted for bulk replication and subsequent recovery. These two discoveries together made gene cloning feasible and even straightforward.

However, they did not offer a means of recognizing and identifying gene sequences. Hence, the third development was nucleic acid hybridization, with the possibility of using a tagged probe sequence to search out and hybridize with any other molecule having partial sequence homology.

With these three developments, the old dreams and experimental visions suddenly became a reality. By purifying and cloning individual gene sequences, it became possible to add many copies of a single gene sequence to eggs or embryos, and then to use DNA hybridization with a radio-tagged probe DNA to search out any persisting or progeny molecules from the initial introduction now present in the adult animals.

So it is that the science of transgenic induction in animals has arisen within the past decade, and in the pages that follow, we endeavour to provide a clear account of what has been achieved in this period of intense experimental effort. We also attempt some predictions of where the future lies. What began as a pipe dream or an attempt to improve stock characteristics by selective breeding has now generated other and sometimes unexpected rewards, such as gene therapy for the treatment of inherited disease, new insights into gene regulatory mechanisms and the use of farm animals as production units for the synthesis of desirable pharmaceutical products.

Transgenic animals can be defined as individuals to which copies of a gene sequence have been artificially added. The novel DNA is referred to as the *transgene,* since it is delivered by artificial means to the individual animal, usually but not invariably by insertion into the fertilized egg. If the sequence is demonstrated to have been incorporated into the chromosomal DNA of the animal, the individual can be regarded as being stably transgenic with respect to the added sequence.

More frequently, the added DNA sequences are not integrated into the chromosomal DNA but are gradually lost from the embryonic cells over the first few days of development. This is not to say that the genes are not expressed, and indeed such *transiently transgenic* animals often display transient expression of these temporarily resident

gene copies and thus display novel proteins in their temporarily modified phenotype.

When a novel gene has been *integrated* into the chromosomal DNA of an animal, the gene may be active in certain cells and tissues, specific proteins coded by the gene may be produced and consequent changes in the phenotype of the animal may result. Such *expression* of the transgene is frequent but not invariant. Also, the transgene is likely to be *transmitted* in the sexual gametes to progeny. Since the initial transgenic individual is almost invariably heterozygous for the novel sequence, no more than 50% of the initial F1 generation progeny are likely to inherit the transgene. The percentage can be lower if the initial transgenic animal is mosaic, in that only some of its cells have arisen from a transgene-inheriting cell (i.e., if incorporation of the transgene copies has occurred after a few rounds of cell replication in the early embryo). Alternatively, the percentage can be higher if more than one chromosome has inherited transgene copies. The three crucial aspects of transgene biology are therefore *integration, expression* and *transmission*.

Although we are here using the term 'transgenic animal', much of the legislation which has arisen to regulate the experimentation and release of such organisms refers to them as 'genetically modified' or 'genetically engineered'. However, since this terminology is ambiguous, inviting confusion with animals whose genomes have been modified by mutation, polyploid or aneuploid induction or whatever, the term 'transgenic animal' is clearer and more straightforward.

1.1.2 Why produce transgenic animals?

There are at least two quite separate reasons for the vigorous efforts in many laboratories to produce transgenic animals. The first is simply that such animals can be used as 'living test tubes' into which a novel sequence is introduced in order to learn about the mechanisms governing its regulation. In this way problems and questions relating to gene promoter characteristics, enhancer sequences, transcription factor site interactions, mutations, position effects, DNA methylation, gender imprinting, tissue specific expression and so on are readily addressed.

The second reason is the long-cherished desire to develop strains of animals with new and desirable genetic traits which will prove useful in agriculture, aquaculture or the domestic environment. Essentially

the logic here is that breeding by genetic selection is usually slow and imprecise, whereas the development of an animal strain carrying and expressing a specific novel gene might be more rapid and more specifically goal-oriented via the transgenic route.

There can be no question that transgenics have richly fulfilled their promise in the first regard. Whether they will do so in the second remains for the most part to be demonstrated, although there have already been promising successes.

Since this is a book about transgenic animals, we will not dwell on the extensive experimental work on the production of transgenic bacteria, fungi and plants. But it is worth noting that all of these systems have natural plasmids which can if necessary act as vector systems to convey and express new genes in the cells of these organisms. Such natural plasmids are almost unknown in animal cells, even in those of protozoans, and therefore transgenic induction in animals is more difficult than it is in other living systems. The ability to grow entire plants from single somatic cells has also made transgenic induction in the plant kingdom somewhat more straightforward than in the animal kingdom.

Some interesting books and reviews covering transgenic animal biology are the following:

Church, R. B., ed. (1990). *Transgenic Models in Medicine and Agriculture*. New York: Wiley-Liss.
First, N. L. & Haseltine, F. P., eds. (1991). *Transgenic Animals*. London: Butterworth-Heinemann.
Grosveld, F. & Kollias, G., eds. (1992). *Transgenic Animals*. London: Academic Press.
Jaenisch, R. (1988). Transgenic animals. *Science* **240,** 1468–74.
Kingsman, S. M. & Kingsman, A. J. (1988). *Genetic Engineering*. Oxford: Blackwells.
Multiauthor review (1991). Transgenic vertebrates. *Experientia* **47,** 865–935.
Murray, J. A. H., ed. (1992). *Transgenesis: Applications of Gene Transfer*. New York: Wiley.
Peters, P. (1993). *Biotechnology: A Guide to Genetic Engineering*. Dubuque, Iowa: Brown.
Watson, J. D., Witkowski, J., Gilman, M. & Zoller, M. (1992). *Recombinant DNA*, 2d ed. San Francisco: Freeman.

1.2 A brief look at the methodology

In general, the easiest way to achieve transgene induction in an animal is to introduce the DNA as multiple copies of the chosen sequence into the fertilized egg. For example, in the mouse system, eggs are

recovered from a superovulated mouse, and the transgene copies are injected directly into one of the pronuclei (usually the male pronucleus). The early embryos are then reintroduced into the uteri of pseudopregnant surrogate females. Variants of this approach are used with insects, fish and amphibians where the nuclei of the egg are too small to be readily visualized and the transgene copies are injected into the perinuclear cytoplasm. In certain fish species, oocytes rather than eggs are injected and *in vitro* fertilization follows injection.

However, injection is slow and often requires considerable skill; a variety of alternative methods exist. These include electroporation of sperm or fertilized eggs (in some cases after egg dechorionation); lipofection, in which the transgene copies are prepackaged in lipid micelles and fusion of these packages with the eggs follows; and bombardment of the eggs or embryos with gold or tungsten particles previously coated with the DNA. If gametes and zygotes are inaccessible, as they are in birds, infective retroviruses may be used as vectors and the embryos exposed to the infective virus. Integration will, of course, be mosaic and germ-cell transmission rather infrequent.

The exploitation of embryonic stem cells, especially in mammals, has opened up altogether new and beneficial ways of inducing transgenesis. Here cells of embryonic origin (in mammals these are cells originating from the inner cell mass of the early blastula) are cultured *in vitro* in conditions that allow proliferation without differentiation. While in culture, they may be transfected with the transgenes by electroporation or facilitated uptake. What is so good about this system is that the cell population can then be exposed to rigorous selection in culture, but not just in terms of transgenic versus nontransgenic cells; cells can also be selected which have fortuitously incorporated the transgene into a particular locus. This also opens the possibility of gene targeting either by homologous recombination or by selective ablation. Once selected, the desired cells can then be replaced in a suitable embryo and the embryo grown to term. Resulting transgenic individuals will be mosaic, but a percentage will be partially germline-transformed.

It will be obvious to many readers that advances in gene cloning permit the test-tube construction of new combinations of DNA sequences, in which the promoter of one sequence is spliced to the coding sequence of another and so on. Such constructs also make it possible to use transgenic animals as test systems in order to investigate the precise roles of regulatory sequences such as TATA boxes

(consensus sequences found in most promoter regions and recognized by RNA polymerase enzymes and transcription factors) and enhancers, especially in terms of tissue-specific expression. An important part of this methodology is the use of reporter genes. These are protein coding genes which encode enzymes not present in the transgenic animal and for which a sensitive assay procedure exists. Popular reporter genes include *lac Z, CAT* and *LUC,* which are, respectively, bacterial β-galactosidase, bacterial chloramphenicol acetyltransferase and luciferase (commonly of firefly origin).

Transgene expression can be profoundly affected by the characteristics of the construct, the site of integration (so-called position effects), the integration of single copy or concatemer tandem copies and endogenous responses such as transgene methylation.

The presence of the transgene is commonly checked by such techniques as the polymerase chain reaction and Southern blotting; it is of great importance to demonstrate that chromosomal integration of transgene copies has actually occurred. Transient expression of nonintegrated transgene copies is routinely observed, so expression is no indication of integration. Integration into mitochondrial DNA is theoretically possible with such techniques as cytoplasmic injection but has been little investigated.

The method for investigating transmission is to look for integration and/or expression in progeny following interbreeding of transgenics with nontransgenics. Transgenes may be lost as well as gained, and this presumably becomes more likely if incorporation has been facilitated by terminal repeat sequences of a transposable element. Although the phenomenon has been reported, most transgenes have proved to be stably inherited once integrated.

1.3 Some important advances

Since the field of transgenic animals is now so large, it is not possible to discuss all the important advances in one short section. Instead, a few key developments will be highlighted and briefly discussed.

1.3.1 Introducing the transgene to the animal

In this fundamental methodology, injection into eggs or very early embryos by microneedle remains the most reliable choice in all but a very few systems. Lipofection and electroporation are probably the

closest rivals in terms of promise, but fusion of the liposome, containing the packaged DNA, with the fertilized egg cell is often difficult and sometimes impossible. Electroporation requires finding a balance between survival of the eggs after exposure to the electrical field conditions and electrical pulse strength which will transfer the DNA into the eggs. If very few copies of the transgene are transferred, the chances of integration and subsequent transmission remain poor. The characteristics of different electroporation kits are highly variable, and it may yet become clear that the technique has much to offer. (The great benefits of exploiting embryonic stem cells for gene transfer are discussed separately in Section 1.3.5.)

Another interesting possibility is the introduction of transgenes directly into gonads by localized microinjection or ballistic delivery via DNA-coated particles propelled by an explosive charge. If the DNA were incorporated into progenitor cells which would later produce gametes, then both integration and germ-line transmission might be achieved in one step by the use of the gametes after such targeted introduction.

1.3.2 Transgene integration

Although many studies using transgenic animal technology require only transient expression of nonintegrated copies of the transgenes to provide useful data, most experimental programmes attempt the chromosomal integration of one or more transgene copies. Some interesting methods have been invoked in order to facilitate integration. These chiefly involve the attempted use of terminal repeat sequences, as are found in retroviruses and other transposable elements. Such terminal repeats, in the form of a series of short repeats, can be added to the termini of gene constructs in the hope of facilitating integration after transgene delivery.

However, the efficiency of such repeats in achieving integration is probably partly dependent on the existence of similar repeats within the genome being targeted and on specific transferase enzymes which facilitate the actual integration. For example, not only do the P elements used so efficiently to achieve transgene integration in *Drosophila melanogaster* carry their own specific transposase gene, but the splicing of the exon is appropriately achieved only in germ-line tissue, thus preventing normal transposition in somatic cells (see review by Rio 1990). Harnessing the machinery of transposition to help achieve

integration in other systems is clearly an attractive proposition, and already the co-injection of transposase enzyme has been shown to be just as effective as co-injecting the P element in *Drosophila* (Kaufman & Rio 1991).

Retroviruses and some bacteriophages effect their integration into host genomes by using recombinase enzymes. One family of site-specific recombinases are termed *integrases,* requiring as they do short regions of genetic homology between the partner molecules (Argos *et al.* 1986). The yeast *FL* recombinase has been used in mammalian cells: Cells were transfected with an expression construct together with a reporter plasmid that was a substrate for the recombinase (O'Gorman *et al.* 1991). The *FL* recombinase system has also been shown to catalyze site-specific recombination in *Drosophila* (Golic & Lindquist 1989). More recently, the phage P1 *Cre* recombinase has been used in tandem microinjection along with a *loxP*–flanked *β*-galactosidase reporter gene, and the resulting transgenic mice shown to have integrated the reporter transgene at specific sites (Orban *et al.* 1992). Of particular interest is the work of Dale and Ow (1991), who achieved the integration of a luciferase gene in combination with a hygromycin phosphotransferase gene (to act as a selectable marker gene) in tobacco plants. The phage P1 *Cre/lox* recombination system was used to provide transgene flanking sequences, and the *Cre* recombinase was subsequently activated to excise the selectable marker but to leave the reporter gene. Such a system might be desirable for application in transgenic animal systems.

1.3.3 Transgene expression

As mentioned earlier, when transgenic animals are used as 'living test tubes' in which to study gene expression by novel constructs, integration may not be required. When transgene copies are introduced into egg cytoplasm rather than into the nucleus, as they are, for example, in insects, fish and amphibians, a pseudonucleus frequently forms around the exogenous DNA. Replication of the transgene copies may then occur, with or without attendant gene expression. Eventually, most or all of this foreign DNA is degraded or lost, presumably because, since it is not associated with centromeres, it cannot be readily segregated at mitosis. (See discussion of this topic in Chapter 3, Section 3.4.1.) This specific temporary amplification of cytoplasmically located transgene copies is therefore a bonus to transient expres-

sion studies. Temporary replication and expression of nuclearly located transgenes may occur but are apparently less dramatic.

Transient-expression studies often involve reporter genes, and the widespread use of *neo, CAT, β-gal* and *LUC* amply demonstrates their effectiveness. Such genes are particularly useful for studies on regulatory elements. The reporter gene of choice varies with the transgenic system. *neo* is often used as a selection system, since it provides resistance to the toxic drug neomycin or its analogue G418. Unfortunately, this approach will not work if the transgenic animals are mosaic with respect to the transgene, since nontransgenic cells will not be resistant. The sensitivity of the assay for particular reporter genes is often highly variable. In my own laboratory we find the procedure for detecting *β-gal* particularly sensitive and error free. The use of some of these reporter genes is bedevilled by confusion with endogenous expression, either of galactosidase in the foetal gut, or even of contaminating bacteria in the sample.

The use of luciferase enzyme in living organisms has stirred the imagination, so that transgenic plants or fish may glow in the dark when provided with the substrate luciferin. Living zebra fish transgenic for this reporter gene may be placed in vials of water containing a trace of luciferin and be scored as positives by scintillation counting! Maximal transient expression is probably best afforded by the use of viral promoters such as the Rous sarcoma virus long terminal repeat, since viral promoters have presumably evolved to function under these precise conditions.

Turning to integrated expression, the expression of genes is profoundly affected by a host of regulatory factors, one of which is the specific attachment of the transcribing DNA to the chromosomal scaffold so as to best effect its looping out to facilitate transcription. Such matrix-attachment regions (MARs) have been isolated, and there is some evidence that their inclusion in a novel construct will indeed facilitate transcriptional expression. A useful discussion of the role of such MAR sequences, especially in connection with the chicken lysozyme gene, is to be found in Sippel *et al.* (1992) and McKnight *et al.* (1992). Endogenous activity of MAR elements may also explain the finding that co-integration of a poorly expressed transgene in the vicinity of an actively transcribed transgene can improve the expression efficiency of the former (Clark *et al.* 1992b, 1993). The inclusion of tissue-specific locus control regions is also important (see discussion by Dillon 1993), and there is a good deal of evidence that genomic

genes with introns are often better expressed than their intronless counterparts (Clark *et al.* 1993).

The expression levels of integrated transgenes are most profoundly affected by the associated 5′ regulatory region. Until recently it was commonly assumed that if sufficient 5′ promoter sequence was provided, maximal expression would be obtained. Sadly, things are clearly not so simple. As is clear from the work of Wada-Kiyama *et al.* (1992) on human ϵ-genes, silencer elements may also be found in upstream sequences, and rates of expression can be dramatically increased by actually shortening the 5′ regulatory region in some cases. It now looks as though each promoter in turn will have to be dissected and manicured by a substantial series of experiments involving a good reporter gene before maximal expression in any one tissue is obtained. Let us hope that not too much tissue specificity and species specificity will turn out to involve silencer as well as enhancer elements.

1.3.4 Optimizing transmission

A common complication in ensuring a high rate of transgene transmission, provided that a reasonable incidence of integration is being achieved, is mosaicism, and anything that can be done to reduce a mosaic status in the original transgenic animals, or at least to ensure that the germ-cells are themselves at least partially transgenic, will ensure a reasonable incidence of transmission. If the germ-line is set aside very early in development, as it is in determinative embryos, any integration events occurring after the first few rounds of division are bound to result not only in mosaicism but in poor or zero rates of transmission to progeny. Regulative embryos like those of fish are more amenable to transmission but still often lead to mosaic transgenics (perhaps because of the cytoplasmic injection procedure). Transgene mosaicism of germ-cells is commonly found in the gonads of transgenic fish.

In mammals, where gender imprinting occurs, this may pose problems for transmission if transgenes have been integrated into the sequences of other crucial genes. Gender imprinting also precludes another useful genetic trick that may prove beneficial in fish and amphibians, namely the induction of gynogenetic or androgenetic development in the eggs of transgenic individuals. This should ensure