Animal Transgenesis and Cloning

Louis-Marie Houdebine

WILEY

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

Animal Transgenesis and Cloning

Animal Transgenesis and Cloning

Louis-Marie Houdebine

Institut National de la Recherche Agronomique,
Jouy en Josas, France

Translated by

Louis-Marie Houdebine, Christine Young,
Gail Wagman and Kirsteen Lynch

First published in French as *Transgenèse Animale et Clonage* © 2001 Dunod, Paris

Translated into English by Louis-Marie Houdebine, Christine Young, Gail Wagman and Kirsteen Lynch.

This work has been published with the help of the French Ministère de la Culture-Centre national du livre

The Atrium, Southern Gate,
Chichester, West Sussex, PO19 8SQ,
England

National 01243 779777
International (+44) 1243 779777

e-mail (for orders and customer service enquiries):
cs-books@wiley.co.uk
Visit our Home Page on http://www.wileyeurope.com
or http://www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 111 River Street,
Hoboken, NJ 07030, USA

Wiley-VCH Verlag GmbH, Pappelallee 3,
D-69469 Weinheim, Germany

John Wiley & Sons (Australia) Ltd, 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01,
Jin Xing Distripark, Singapore 0512

John Wiley & Sons (Canada) Ltd, 22 Worcester Road,
Rexdale, Ontario M9W 1L1, Canada

Wiley also publishes in books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloguing-in-Publication Data
applied for

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-84827-8 (Hardback)
0-470-84828-6 (Paperback)

Typeset in 10½/13 pt Times by Kolam Information Services Pvt. Ltd., Pondicherry, India
Printed and bound in Great Britain by TJ International, Padstow, Cornwall
This book is printed on acid-free paper responsibly manufactured from sustainable forestry, in which at least two trees are planted for each one used for paper production.

Contents

Introduction

Since the beginning of time, humans have known how to distinguish living organisms from inanimate objects. Cro-Magnon people and their descendants were no doubt aware that living beings all had the same ability to grow and multiply by respecting the specificity of the species. It probably took them longer to understand that heat destroyed living organisms, whereas the cold, to a certain extent, conserved them.

These very ancient observations have fixed in our minds the notion that living organisms are fundamentally different from inanimate matter. We now know that living beings are also subject to the laws of thermodynamics, that they are no more than very highly organized matter and that they only conserve their wholeness below about 130 °C.

Well before having understood what made up the very essence of living beings, the different human communities learned to make the most of what they had, sometimes without even realizing it. The existence of micro-organisms was unknown until the 19th century and yet fermentation has been carried out for thousands of years in certain foods. Agriculture, farming and medicine benefited from empirical observations that enabled genetic selection and the preparation of medicine, particularly from plant extracts.

The situation changed radically during the 19th century with the discovery of the laws of heredity by Gregor Mendel, the theory of evolution by Charles Darwin and the discovery of cells. The classification of living beings has progressively demonstrated their great similarity in spite of their infinite diversity. Jean-Baptiste Lamarck as well as Charles Darwin accumulated observations supporting the theory of evolution. The two scientists admitted that the surrounding environment had and continued to have a great influence on the evolution of living beings. Darwin was the person who most contributed to establishing the idea that living beings mutated spontaneously by chance and the environment was responsible for conserving only those that were the best adapted to

the conditions at the time. Mendel determined in what conditions the traits were transmitted to the progeny, thus establishing the laws of heredity.

The innumerable observations made possible by the invention of the microscope in the 17th century revealed the universal existence of cells in all living beings. The remarkable properties of living organisms began to be explained: their resemblance, their evolution and their diversity.

We had to wait until the discovery of the principal molecules that constitute living organisms (proteins, nucleic acids, lipids, sugars etc.) to begin to understand the chemical mechanisms that govern their existence. The theories of the 19th century are now confirmed every day at the most intimate level of living beings, and in particular by the observation of the structure of genes and proteins.

It is now acknowledged that the big bang, which must have occurred 15 billion years ago, was followed by an expansion of matter, which, when cooling down, progressively and continuously gave way to particles, atoms, mineral molecules, organic molecules and finally living organisms. Only the present specific conditions on Earth enable the highly organized matter of living organisms to survive, proliferate and evolve.

The discovery of the structure of genes and proteins as well as the identification of the genetic code about 40 years ago enabled us to comprehend for the first time what living organisms are and how they function. Even more, these discoveries have in principle provided humans with new and powerful means to observe and make use of certain living species. This has required mastering a certain number of techniques, which we group together under the term genetic engineering.

From the moment it was known that the structure of DNA directly determines the structure of proteins, it was in principle possible to manipulate one or the other by chemical reactions that determine and modify the structure of genes. This presupposes that the genetic information manipulated in this way can be expressed. In practice this is not possible, and only makes sense if the gene can give rise to the corresponding protein and if the protein can exercise its biochemical properties in the complex context of life. To do so, the isolated and possibly modified gene can be reintroduced into a cell or a whole organism. It is for this reason that gene transfer occupies an essential place in modern biology as well as in biotechnological applications.

In the period of only a few decades, the work of biologists has changed dramatically. For about a century, biologists had worked essentially *in vivo* on whole animals, plants or micro-organisms. This made it possible to define the role of the principal functions of living organisms, to identify a number of hormones etc. The traditional scientific approach is based on systematically dividing up problems to try to simplify them and thus resolve them. Biologists have therefore started to work *in cello* with cultured isolated cells. This promising simplification has been followed by studies conducted *in vitro* using cell extracts or even purified molecules. The huge quantity of information provided by genome mapping and their complete sequencing requires biologists to use other ways to deal with the problems. This information is so vast that it needs to be dealt with *in silico* by powerful computer processing.

The present situation is particularly promising. Biologists have the means of knowing all the genetic information of a living organism through the complete sequencing of its DNA. It is clear that the primary structure of a gene makes it possible to predict that of the corresponding protein. Most often, it only indicates very partially the role of the protein. Proteins, like genes, are derived from each other during evolution. Therefore, it is sometimes possible to determine that a protein, whose structure has been revealed by sequencing its gene, has for example a kinase activity, by simple structure homology with that of other proteins known to possess this type of enzymatic activity. The predictions often stop at this level or never even reach it. The transfer of the isolated gene in a cell or even in a whole organism is likely to reveal the biological properties of the corresponding protein. Thus the oversimplification which the isolation of a gene represents is accompanied by a return to its natural complex context, which is the living organism. Hence, biologists are experiencing a spectacular link between traditional physiology and molecular biology. This is now referred to as postgenomics.

In this context, transgenesis has an increasingly important role despite all its theoretical and technical limits. This is why transgenesis workshops are developing in order to enable researchers to try to determine *in vivo* the role of all the genes that are progressively available to them.

Reproduction has always played an essential role in the life of humans. They themselves reproduce of course and sometimes with more difficulty than they would like or in contrast with an excessive prolificacy. Livestock farming and agriculture are to a great extent based on reproduction. In animals, controlling reproduction has occurred progres-

sively. It involved successively favouring mating or not, carrying out artificial insemination, embryo transfer, *in vitro* fertilization and finally cloning. All these operations aim essentially at increasing the efficiency of reproduction (for breeding animals in large numbers) and at enabling an effective genetic selection. These techniques are receiving increasing back-up from the fundamental study of reproduction mechanisms.

The case of cloning does not escape this rule. Cloning animals began with a biologist's experiment. It was adopted by biotechnologists eager to speed up progress in genetics by introgressing the genomes validated by their very existence as is already the case in plants. In all species, transgenesis depends very much on controlling reproduction. The technique of cloning has shown that it was indeed at the source of a simplification of gene transfer and an extension of its use. Reproductive cloning could, in principle, become a new mode of assisted reproduction for the human species. Therapeutic cloning could in principle help in reprogramming differentiated cells from a patient in order to obtain organ stem cells to regenerate defective tissues.

Cloning and transgenesis and the generation of cells for human transplants are henceforth very closely associated. Cloning is the opposite of sexual reproduction, which is accompanied by the reorganization of genes. The fundamental aim of transgenesis, on the other hand, is to modify the genetic heritage of an individual or even a species. The reprogramming of cells concerns the differentiation mechanisms irrespective of any genetic modification. This book sets out to give a clear picture of recent developments in research and its applications in these three fields. It does not describe the techniques in detail, namely those used to generate transgenic animals. The readers may find this information in other books edited by C.A. Pinkert (2002) and A.R. Clarke (2002).

Acknowledgements

The author wishes to thank Ms Annie Paglino, Christine Young, Gail Wagman, Kirsteen Lynch and Mr Joel Gallé for their help in the preparation of this manuscript.

Abbreviations and Acronyms

AAV	adeno-associated virus
BAC	bacterial artificial chromosome
CHO	chinese hamster ovary
DPE	downstream promoter element
EMCV	encephalomyocarditis virus
EBV	Epstein–Barr virus
ES cells	embryonic stem cells
EST	expressed sequence tag
ENU	ethyl-nitroso-urea
EG cells	embryonic germinal cells
EC cells	embryonic carcinoma cells
GFP	green fluorescent proteins
GPI	glycophosphatidyl inositol
GMO	genetically modified organism
GMP	genetically modified plant
GMA	genetically modified animal
HSV	*Herpes simplex* virus
HAC	human artificial chromosome
HAT	hypoxanthine, aminopterine, thymidine
HPRT	hypoxanthine phospho-ribosyl transferase
IRES	internal ribosome entry site
ITR	inverted terminal repeat
ICSI	intra-cytoplasmic sperm injection
Inr	initiator element
KO	knock-out
LCR	locus control region
LTR	long terminal repeat
MPF	maturation promoting factor
MAR	matrix attached region

mRNA	messenger RNA
NMD	nonsense mediated decay
NLS	nuclear localization signal
OPU	ovum pick-up
PrP	proteinous particle
PCR	polymerase chain reaction
PGK	phosphoglycerate kinase
PTGS	post-transcriptional gene silencing
RNAi	RNA interference
RMCE	recombinase-mediated cassette exchange
rRNA	ribosomal RNA
RDO	ribodeoxyribo-oligonucleotide
REMI	restriction enzyme mediated integration
SA	splicing acceptor
SD	splicing donor
tRNA	transfer RNA
TFO	triplex forming oligonucleotide
TAMERE	targeted meiotic recombination
TM	transmembrane
TGS	transcriptional gene silencing
UTR	untranslated region
5′UTR	5′ untranslated region
3′UTR	3′ untranslated region
YAC	yeast artificial chromosome

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

1 From the Gene to the Transgenic Animal

1.1 Genome Composition

A genome is by definition all the genes that characterize a species and in a more subtle manner each individual. In practice, this word designates all the information stored in DNA. DNA contains genes, which strictly speaking correspond to regions transcribed in RNA (Figure 1.1). Some of the RNAs such as ribosomal RNAs (rRNA) or transfer RNAs (tRNA), which provide amino-acids for protein synthesis, have an intrinsic biological activity. The most numerous RNAs in terms of sequence diversity are messenger RNAs (mRNA), which contain the genetic information capable of directing protein synthesis according to a rule defined as the genetic code (Figure 1.2).

Besides the regions transcribed in RNA, genomes contain multiple sequences with diverse functions or seemingly, for some of them, no

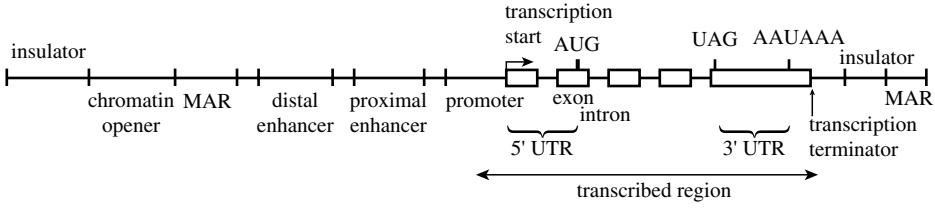


Figure 1.1 Major gene structural elements. L.M. Houdebine, Medecine/Sciences (2000) 16: 1017–1029. © John Libbey Eurotext. Gene expression is controlled by sequences located upstream of the transcribed region. Promoters participate directly in the formation of the preinitiation transcription complex. Enhancers increase the frequency of promoter action. Distal regions, MAR (matrix attached region), chromatin openers and insulators maintain an open chromatin configuration and prevent gene silencing by the surrounding chromatin

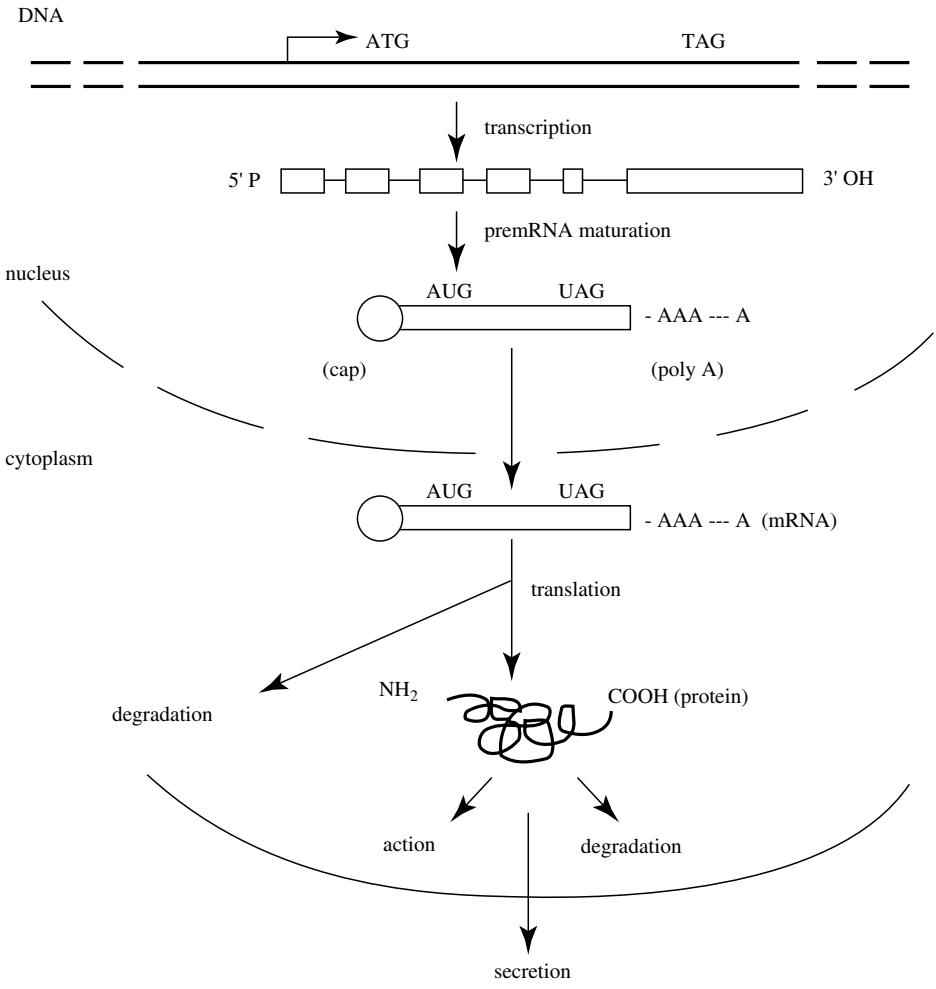


Figure 1.2 Major steps in gene expression. The genetic information in DNA is stable. It is decoded in proteins via synthesis of unstable messenger RNAs. Proteins act inside or outside of the cell and also on the cell membrane. They are unstable and are resynthesized if needed. The regulation of gene expression may occur at all of the steps: transcription, selection of the transcription initiation site, exon splicing, translation and mRNA stability

function. Indeed, DNA must replicate at each cell division. DNA contains regions where DNA replication is induced. DNA is organized in chromosomes which are visible during mitosis. In the other phases of the cell cycle, chromosomes are in euchromatin, which corresponds to the open chromatin regions, where the genes active in a given cell type are

located, or in heterochromatin, which is a condensed form, where the inactive genes are present. The generation of the different forms of chromatin is triggered by the association of regulatory proteins with DNA sequences mostly located outside the transcribed regions.

DNA in eukaryotes contains centromeres formed by long stretches where the cytoskeleton binds during mitosis to dispatch homologous chromosomes in daughter cells. Chromosome ends contain particular repeated sequences, telomeres, which preserve DNA from degradation by cellular exonucleases.

Genomes also contain other DNA sequences whose function is not yet well known. They contain numerous regions that are apparently not useful for the life of the organisms (Comeron, 2001). Some of these sequences seem to alter or even threaten genome integrity. This is the case of sequences from retroviruses that are definitively integrated, more or less randomly, in the genome of infected cells. Transposons are also integrated sequences, which are transcribed, replicate and integrate in multiple sites of the genome without leaving the inside of the cell. Transposons thus spread and tend to invade the genome without any need of infection as is the case for retroviruses. It is well established that transposons have contributed and still contribute to the formation of genomes.

Genomes also contain relics of genes that have become inactivated over time by different mechanisms and which, for this reason, are called pseudogenes.

Very short sequences (microsatellites) or longer sequences (minisatellites) are present in numerous copies in animal and plant genomes. Most of these sequences are very poorly conserved and seem to result from uncorrected errors of transcription.

The vast majority of these sequences seem to have no favourable effect on genome activity. For these reasons, they are sometimes called ‘selfish DNA’, implying that they are programmed to be maintained in genomes. More probably, they are just neutral and are thus not eliminated during evolution as long as they do not hamper genome functioning. Some of these sequences are clearly deleterious for the genomes. Transposons and retroviruses sometimes integrate within genes, which become inactivated. Repeated sequences also modify gene activity when they are in their vicinity or within the genes.

Evolution has endowed cells with mechanisms capable of inactivating parasite DNA sequences and particularly of blocking their propagation, which could severely or completely alter genome functioning.

1.2 Gene Structure

Genes, strictly speaking, vary in size according to species (see Figure 1.1). In eukaryotes, most of the genes are interrupted by non-coding sequences named introns which are eliminated from the native mRNAs to generate the functional mature mRNAs, which then migrate from the nucleus to the cytoplasm to be translated into proteins. Mature mRNAs are thus formed by the exons, which become associated after the introns are eliminated (Figure 1.2).

Both the number and size of the introns have increased during the course of evolution for no clear reason (Comeron, 2001). Introns are mandatory for mRNA maturation in the nucleus and the transfer of the mRNAs to the cytoplasm (Luo and Reed, 1999).

Recent studies have shown that exon splicing requires the action of a ribonucleoprotein complex named spliceosome. After the splicing, a number of the proteins are released from the complex but some of them remain bound to the first 20–24 nucleotides of the upstream exon. This complex plays the role of a shuttle for transferring the mature mRNA to the cytoplasm (Ishigaki *et al.*, 2001).

The spliceosome recognizes the CAG GUA/GAGUA/UGGG consensus sequence in the upstream exon and the CAG G consensus sequence in the downstream exon. After intron elimination and exon splicing, the remaining consensus junction sequence is CAGG. Various splicing enhancer sequences are present in the intron (a pyrimidine rich sequence and the branched point sequence) and in the downstream exon (Wilkinson and Shyu, 2001).

Introns participate in the quality control of mRNAs in the nucleus. It is increasingly acknowledged that a translation of the mature mRNAs occurs in the nucleus to check their functionality. One of the surveillance mechanisms has been recently deciphered. A termination codon followed by an intron at a distance smaller than 50 nucleotides is considered as non-functional and is destroyed in the nucleus by a mechanism that has been named nonsense mediated decay (NMD) (Wilusz *et al.*, 2001).

Some introns are so long that they contain functional genes. The first introns located in the 5′P part of the genes often contain sites for binding transcription factors. Their presence seems important to maintain a local open chromatin and favour transcription.

Some mRNAs have no intron. This is the case for histone and numerous viral mRNAs. These mRNAs contain signals allowing the

mRNA to be transported from the nucleus to the cytoplasm (Luo and Reed, 1999).

Transcription is regulated by mechanisms that are particularly complex. They involve the action of proteins named transcription factors, which recognize short specific DNA sequences (about 12 nucleotides). Some of the transcription factors bind to DNA and control mRNA synthesis only after having been activated by various cellular mechanisms (stimulation by a hormone or a growth factor, modification of the cellular metabolism, cellular stress, contact with another cell or with the extracellular matrix etc.). The total number of transcription factors is not known. There are several hundred (perhaps 2000) in vertebrates. This relatively small number of factors is sufficient to control the transcription of about 40 000 genes in humans. The very complex and diverse actions of the transcription factors are thus a result of their multiple combinations in the different cell types. A given transcription factor may therefore participate in controlling quite different genes as soon as it becomes associated with a set of factors specific to each cell type.

The regulatory regions of the genes are not all completely known. Yet, it is known that, in higher eukaryotes, they can be divided into distinct parts located mostly upstream of the genes and having complementary functions.

Promoters themselves are located in the vicinity of the transcription initiation site. Promoters are no longer than 150–200 nucleotides. The combination of the transcription factors that bind to the promoter determines its potency and its cell specificity. The transcription complex responsible for mRNA synthesis is formed in the promoter region.

The first promoters found in viral genomes and in the most highly expressed cellular genes were shown to contain consensus sequences. An AT rich short region named the TATA box is present in many genes at about −30 bp upstream of the transcription initiation site. Specific factors bind to the TATA box and they are part of the transcription initiation complex. The study of more diverse genes revealed that this concept is far from reflecting the whole truth. A certain number of genes have no TATA box and their promoter is formed by an initiator element (Inr) overlapping the start site. Other genes have their promoter 30 bp downstream of the initiation site. This category of promoters is named downstream promoter elements (DPEs). The three kinds of promoter use different transcription factors and mechanisms to initiate

mRNA synthesis. This is expected to offer a broader diversity and flexibility to the transcription mechanisms (Butler and Kadonaga, 2001).

Upstream of the promoters and at quite variable distances (from a few hundred nucleotides to 10 kb or more) transcription enhancers are found in most if not all animal genes. The name enhancers has been given to these regulatory regions since they increase the global transcription rate. Recent studies have revealed that enhancers do not increase the transcription rate itself but the probability of transcription occurring. Indeed, it appears that the transcription complex is alternatively active and inactive in a cell. Enhancers act essentially by increasing the frequency of the transcription complex being active (Martin, 2001). Enhancers generally contain multiple binding sites for transcription factors. The DNA–transcription factor complex is named an enhancesome. It interacts with the transcription complex from a distance by the formation of a loop which brings the enhancer and the promoter close together.

Much further upstream (up to 30–100 kb), other regulatory regions have been found in a certain number of genes. These sequences have been found at the border between two unrelated genes or groups of genes. Some of these regulatory regions are named locus control regions (LCRs) (Johnson *et al.*, 2001a). They contain different elements. Some of them are enhancers and others are insulators. The insulators seem to be particular silencers, which prevent the action of an enhancer on a neighbour promoter. The insulators and the specific enhancers of the LCR thus render each gene or gene cluster independent of its neighbour (Bell and Felsenfeld, 1999; West, Gaszner and Felsenfeld, 2002). No more than 30 LCRs or insulators have been described so far. Their structure and mechanism of action is only partly known. They seem diverse and no general rule for their exact effect has emerged so far. One of the functions of the LCRs seems to involve keeping locally the chromatin in an open state, leaving the possibility for the transcription factors to stimulate their target genes. It is interesting to note that a gene or a group of genes is or is not in an open configuration depending on the cell type. Hence, the LCR might play an essential role in determining the active chromatin regions in a given cell type during foetal differentiation. The stimuli delivered by hormones and various cellular events in adult organs therefore seem to control gene expression in a finely tuned manner but only after a major decision has been taken during foetal life to put the genes in a position where they can be sensitive to their specific stimuli or not.

The mature mRNAs in cytoplasm contain different regions having distinct and specific functions (Wilkinson and Shyu, 2001). Mutations

in the non-coding region of mRNAs are often responsible for abnormal protein synthesis and human diseases (Mendell and Dietz, 2001).

The region preceding the initiation codon and named the 5′ untranslated region (5′UTR) is sometimes involved in the control of translation (Pesole *et al.*, 2002). Highly structured 5′UTRs (usually rich in GC) do not favour or even inhibit translation. It is known that the scanning of the 5′UTR by ribosomes is considerably slowed down by secondary structures. This reduces the chance of ribosomes reaching the initiation codon. In contrast, the AU rich 5′UTRs favour, or at least do not hamper, translation (Kozak, 1999). Some of the 5′UTRs contain special regulatory regions, which allow an mRNA to be translated or not according to the physiological state of the cell (Houdebine and Attal, 1999).

The region downstream of the termination codon, which is named the 3′ untranslated region (3′UTR), is relatively long in many genes whereas the 5′UTRs are generally short. Some of the 3′UTRs contain sequences to which proteins bind (Pesole *et al.*, 2002). In some cases, the mRNA protein complex stabilizes the mRNA quite significantly. In other cases, AU rich sequences trigger a rapid destruction of the mRNA. These signals are found in mRNAs subjected to a rapid regulation (Mukherjee *et al.*, 2002). The 3′UTRs of some mRNAs contain sequences that form a complex with cytoplasmic proteins, which target the mRNAs to a specific cell compartment (Mendell and Dietz, 2001).

One of the key steps in transgenesis consists of constructing genes that are expected to be expressed in an appropriate manner when transferred to animals. Taking into account the above-described mechanisms is highly recommended in order to have the best chance of obtaining a satisfactory expression of the transgenes. These recommendations have been summarized in a book chapter (Houdebine, Attal and Villotte, 2002). The mechanisms controlling gene expression are not all known and the construction of a gene may eliminate essential signals or combine incompatible signals, leading to disappointing transgene expression.

1.3 The Number of Genes in Genomes

The size of bacterial genomes suggests that they contain 2000–4000 genes. The complete sequencing of more than 200 bacterial genomes has confirmed this point. The yeast *Saccharomyces cerevisiae* has almost 6000 genes.

One of the simplest known and studied animals, *Caenorhabditis elegans*, a worm of the nematode family, has about 19 000 genes. This organism is made up of only 959 cells, but has most of the animal biological functions. Gene transfer is easy and genetics has been studied for years in this species. For these reasons, *C. elegans* is one of the favourite models for biologists.

The *Drosophila* genome has also been completely sequenced. Rather unexpectedly, this genome does not contain more than 15 000 genes, although *Drosophila* appears a more complex animal than *C. elegans*.

It is known that plant genomes contain about 25 000 genes and mammals probably no more that 40 000–45 000 genes. These numbers may be underestimated, especially in mammals, which have long genes and many repeated sequences, which complicate the identification of genes. These data deserve some general comments. As could be expected, the degree of complexity of a living organism is related to how many genes it has. Yet, the number of genes alone cannot account for the difference in complexity between the various species.

It is striking that plants have 25 000 genes although they are devoid of nervous and immunological systems and are controlled by a relatively simple endocrine system in comparison to mammals. Close examination of plant genes has revealed that a large proportion of them are involved in controlling their metabolism. This may be required for organisms that cannot move during their life and that must have a high capacity to adapt to cold, heat, dryness, stress, salt etc.

Another point deserves attention. The number and structure of the genes of the higher primates are quite similar to human genes. The first systematic comparisons of the expression levels revealed that a number of genes are expressed differently in the brains of higher primates and humans. This might be responsible for generating the differences between primates and humans.

It is increasingly considered that the complexity of living organisms is due to a large extent to the number and nature of the interactions between the proteins and the various cell components (Szathmary, Jordan and Pal, 2001). Proteins are larger in animals than in bacteria. They are formed of different domains, which interact in multiple ways with other molecules.

Growing evidence indicates that the genomes contain regions transcribed in non-coding RNA. Some of these RNAs are well known. Ribosomal RNAs and small RNAs involved in forming the ribonucleoprotein complexes that act in exon splicing are examples of

non-translated RNAs. Many of the non-coding RNAs seem to have essentially regulatory roles. They act as antisense RNA, modify chromatin structure, interact with proteins to modulate their activities, etc. (Mattick, 2001). These RNAs might be very numerous and coded by the genome regions considered as containing no genetic information (Ambros, 2001).

It is now commonly observed that a protein has for example a given function in a stage of embryo development and a different function in a differentiated cell of an adult. This diversity of function results from the multiple interactions of proteins with each other and various cell components. One of the most striking examples is the case of transcription factors. No more than 1000 or 2000 transcription factors are sufficient to control the 40 000 human genes, including their own genes. Obviously transcription regulation results from the multiple combinations of the transcription factors.

A gene frequently has several sites of transcription initiation. The same gene can thus generate different mRNAs coding for proteins having different structures and different biological activities.

The elimination of introns from pre-mRNA is followed by splicing the exons surrounding the introns. In a certain number of cases splicing does not occur between the most adjacent exons. Then, several exons and introns may be eliminated and splicing occurs between remote exons. This phenomenon is by no means rare and one-quarter of the pre-mRNAs might be subjected to this mechanism, called alternative splicing. Interestingly, this phenomenon is tightly controlled in different cell types or in a given cell type in various physiological situations. Alternative splicing may lead to the synthesis of different proteins from the same gene. These proteins may have different biological functions.

A mature mRNA may have several initiation codons, which are mostly in the same reading frame. The use of one or other of the initiation codons gives rise to proteins with different lengths. In some cases, essentially in viruses, which have very compact genomes, two coding sequences are superimposed. They use distinct initiation codons, which are not in the same reading frame.

Recent studies have shown that two distinct mRNAs coding for cellular proteins and generated by alternative splicing have different initiation codons. These mRNAs contain 105 overlapping codons. More surprisingly, it has also been observed that the same mRNA codes for two distinct proteins using two different initiation codons and two reading frames (Kozak, 2001a). This genome organization is therefore not

restricted to viruses, which must have compact genomes to replicate rapidly but also to be encapsidated to form infectious particles. It is interesting to note that the two proteins coded by the same mRNA have related biological functions. This observation raises the question of how frequent this phenomenon is in higher organisms. If this mechanism is not an exception, the number of proteins coded by genomes might be higher or even much higher than 40 000 in mammals.

Translation of mRNA is often controlled by specific sequences located in 5′UTR. The most famous example is the case of ferritin mRNA, which is translated only when the hepatic cells are in the presence of iron. This ion binds to a protein linked to a loop in the 5′UTR. In the presence of iron, the protein conformation is modified, allowing the translation of the mRNA. It is interestingly to note that the same loop is present in the 3′UTR of transferrin receptor mRNA. In the presence of iron, the protein bound to the loop stabilizes transferrin receptor mRNA. In this way, the iron metabolism is controlled in a coordinated manner at post-transcriptional levels.

In a certain number of mRNAs, the 5′UTRs contain highly structured GC rich regions that cannot be scanned by ribosomes from the cap. It is believed that these sequences can directly trap ribosomes without any scanning of the 5′UTR. For this reason, they have been named internal ribosome entry sites (IRESs). Experimental data suggest that the IRES might act, at least in some cases, by capturing quite efficiently ribosomes after scanning the 5′UTR. This mechanism implies that ribosomes shunt the IRES very efficiently and pursue its scanning to reach the initiation codon. Many IRESs are active to varying degrees according to the cell type and the physiological state of the cells. IRESs might thus be essentially specific translation regulators, as is the iron binding protein for ferritin mRNA.

After their synthesis, many proteins are biochemically modified in various ways. Some proteins are cleaved to eliminate regions that are inhibitory. The activation of the protein is then dependent on its cleavage. This is the case for most proproteins such as proteases. The fragments generated by cleaved proteins may associate to give rise to the active molecule. This is the case for insulin. Many proteins that are exported out of the cell are glycosylated to varying degrees. This may control their activity but mainly their stability in blood. Proteins may also be phosphorylated, amidated, γ-carboxylated, N-acetylated, myristylated etc. They are often folded in a subtle manner to generate their active sites. Some proteins have several stable or metastable

configurations. One of the most striking cases is that of PrP protein, which plays an essential role in prion diseases. After a folding modification, the PrP protein becomes insoluble and resistant to proteolytic digestion. The deposition of insoluble proteins is found in the brain of patients suffering from prion or Alzheimer diseases. It is known that this phenomenon contributes to inducing these two diseases.

Many proteins, but also some mRNAs, contain targeting signals responsible for their concentration in a given compartment of the cell. Proteins are thus targeted to the nucleus, mitochondria, Golgi apparatus, plasma membrane or outside of the cell according to the signals they contain.

At the gene level, it is well known that DNA methylation on cytosine is responsible for inactivating gene expression. One allele of a given gene may be specifically methylated and thus inactive but not the other. Hence, the allele of paternal origin may be specially inactivated. For another gene, the maternal allele is silenced by methylation. This phenomenon, named gene imprinting, plays an important role in gene expression in vertebrates.

None of these phenomena take place at the DNA level, or at least they do not result from a modification of nucleotide sequence in DNA. For these reasons, they are qualified as epigenetic. These phenomena are reproducible and are genetically programmed.

A gene may therefore generate different proteins (up to three or more) having more or less distinct functions. The importance of epigenesis appears to increase with the emergence of the most evolved living organisms. Obviously, the complexity that characterizes the higher living organisms results from both genetic and epigenetic mechanisms.

A gene may be compared to a microcomputer that has its own program. A cell and, even more so, a living organism may be compared to a network of microcomputers interconnected in a multitude of ways. Genomes are thus data banks and cells are software, which use the data banks each time they need a new protein. The network formed by 40 000 computers interconnected in multiple ways may be highly complex. In this context, transgenesis is somewhat similar to adding a new computer to the network (or to eliminating a computer from the network). Several scenarios may be imagined. The foreign computer may not be compatible with the network. Then, nothing happens. The computer may be compatible with the network and interact with several computers. Adding a single computer may thus enrich the network just

as adding a gene in a living organism results in a higher biodiversity. A third theoretical situation may be encountered: the foreign computer is compatible with the network but disturbs its functioning. This may even lead to completely inactivating the network. Similarly, a foreign gene may alter the health of an animal and even block its development at its first stages. All these situations are observed in transgenic animals.

Another observation is striking in the organization of genomes. The length of DNA is 1 mm in bacteria, 6 mm in yeast, 25 cm–2.5 m in plants, 1.5 m in mammals and 1.8 m in humans. DNA length is therefore related to gene number but not at all strictly. Obviously the bacterial genomes are much more compact than those of higher organisms. This may be due to the fact that genes in animals are longer than in bacteria. Exons but mainly introns and promoter regions occupy a larger space in higher organisms. Introns are much more numerous and longer in mammals than in yeast. Introns may represent up to 90 per cent of the transcribed region of a gene in mammals.

In humans, no more than five per cent of the genome correspond to genes. A major part of the genome is formed by non-functional sequences. A foreign gene added to a genome has thus little chance of being integrated into a host gene. Rather, a foreign gene introduced into a non-functional part of a genome is likely to be silent.

The reason why the genome of higher organisms has kept so many sequences with apparently no function is not known. One may imagine that these sequences are stored and occasionally used to generate new genes. Such events cannot be excluded but appear extremely rare. The intergenic DNA may also have a protective effect. Mutations induced by chemicals or irradiation have more chance of occurring in the non-functional DNA than in a gene. The most likely reason is that the non-functional DNA sequences do not disturb cell functioning in higher organisms. Indeed, in bacteria, yeast and even more so in viruses, DNA must replicate rapidly. Bacteria with a less compact genome divide more slowly and may be eliminated when they are in competition with other bacteria. In most cases, viral genomes must be compact to be integrated into viral particles. On the other hand, many of the viral genomes must replicate as rapidly as possible after infection before the defence mechanisms of the cell start operating to eliminate the virus. The same is not true for the genome of animals. In these organisms, cell divide about once a day and DNA replication takes about two hours. The competition for a rapid DNA replication does not seem a real advantage

for the organism. Extra DNA is therefore not a burden and is not preferentially eliminated.

1.4 The Major Techniques of Genetic Engineering

The aim of this book is not to describe all the techniques of genetic engineering in detail but to consider briefly their potential and their limits.

Most of the messages contained in DNA are linear. This is clearly the case for the genetic messages based on the succession of codons, which define the order of the amino acids in the corresponding proteins. The same is true to some degree for the regulatory regions. The sites that bind the transcription factors are composed of about 12 adjacent nucleotides. The other signals also rely on DNA sequences, each category of signal having its specific language, always based on the four-letter alphabet, ATGC, corresponding to the four bases of DNA.

1.4.1 Gene cloning

To study genes, one step consists of cleaving DNA into fragments, the size of which ranges from a few to hundreds of kilobases. These fragments are introduced into bacterial vectors for cloning. The different available vectors have been designed to harbour different lengths of DNA. Plasmids, cosmids, P1 phage, BACs (bacterial artificial chromosomes) and YACs (yeast artificial chromosomes) can harbour up to 20 kb, 40 kb, 90 kb, 200 kb and 1000 kb of DNA, respectively. Each vector, containing only one DNA fragment, is introduced into a bacterium, which is amplified, forming a clone. Large amounts of each DNA fragment may then be isolated from each clone. The expression 'gene cloning' has been retained by extension of the cloning performed on the bacteria that harbour the DNA fragments.

The direct cloning of a DNA fragment containing a given gene is often not possible. The cloning of the corresponding cDNA is usually an intermediate step. For this purpose, the mRNAs of a cell type are retrotranscribed into DNA by a viral reverse transcriptase. The monostrand DNA obtained in this way is then converted into double-strand DNA by a DNA polymerase. The resulting DNA fragments are cloned in plasmids to generate a cDNA bank. The clone containing the cDNA in question is then identified by the methods described in section 1.5.

1.4.2 DNA sequencing

DNA sequencing consists of determining the order of bases in a DNA fragment. For years, sequencing was performed by slow techniques. It has now been automatized and is carried out on an industrial scale. It is now possible to sequence several thousands of kilobases daily. This is absolutely necessary for the systematic sequencing of genomes. Experimenters also permanently need powerful computers to determine the structure of DNA fragments they have isolated, mutated or assembled.

1.4.3 *In vitro* gene amplification

The technique known as PCR (polymerase chain reaction) for specific amplification of a DNA region is among the most frequently used by molecular biologists. It consists of synthesizing the complementary strand of a DNA region starting from a primer. The primer is an oligonucleotide composed of about 15–20 nucleotides, which is chemically synthesized and specifically recognizes the chosen DNA region. The oligonucleotide is elongated by a bacterial DNA polymerase generating a complementary DNA strand, to which the primer is bound. In most cases, two primers recognizing different sequences of both DNA strands are used simultaneously. This leads to the synthesis of a double-stranded DNA fragment corresponding to the region located between the two primers. DNA regions of 1 kb are commonly used. Up to 20–40 kb may be specifically synthesized under optimized conditions. After about 30 amplification cycles, thousands of copies of the DNA sequence are present in the tube starting from a single copy. This allows the identification of a specific genomic DNA region. This technique is thus used for genome typing but also for identifying individuals. This has become common practice to determine paternity and identify a murderer. PCR is also an essential technique for mutating DNA fragments *in vitro* and for constructing functional genes from various DNA fragments.

1.4.4 Gene construction

Studying genes often requires construction of functional genes starting from various elements. These elements may be regulatory regions but

also transcribed regions. They may be in their native structure or experimentally mutated. This may help identify the regulatory regions that control gene expression. The coding regions may have their native structure. The constructs may then be used to study the effect of the gene in cells or whole organisms. The transcribed regions may contain a reporter gene coding for a protein that can be easy visualized or quantitated by its specific enzymatic activity. This reveals in which cells and at what rate the reporter gene is expressed.

Genetic engineering may also be used on an industrial scale to reprogramme cells or whole organisms to produce recombinants of pharmaceutical interest and to prevent immunological rejection of transplanted cell organs (Figure 1.3).

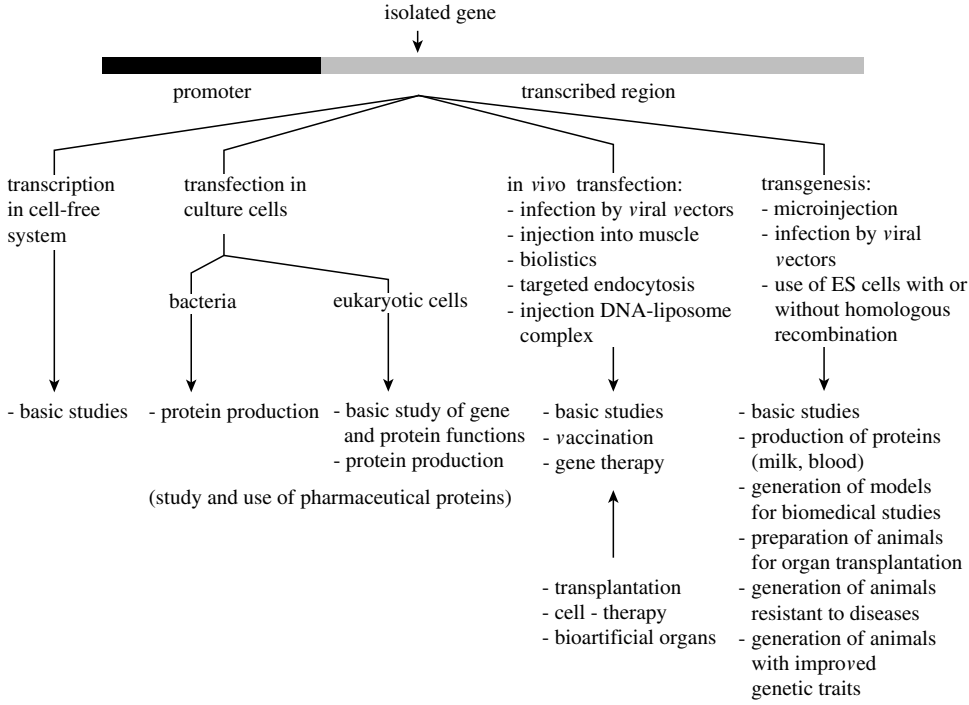


Figure 1.3 Different methods of gene expression. Isolation can be decoded into proteins in cell systems, in bacteria as well as in plant or animal cells. Proteins can be isolated, studied and used as pharmaceuticals. Gene transfer in somatic cells is gene therapy applied to humans. Transgenesis implies foreign DNA transfer and maintenance in the host genome. Genes must be adapted to cell types in which they are expressed

In all cases, genes must be experimentally constructed. Gene constructs contain at least a promoter region, a transcribed region and a transcription terminator. The construct is then an expression vector.

Gene construction implies the use of restriction enzymes, which cleave DNA at specific sites, the chemical synthesis of oligonucleotides, the *in vitro* amplification of DNA fragments by PCR and the covalent association of the different DNA fragments by a ligase. Most of the time, these fragments are added in plasmids, which are transferred into bacteria. The bacterial clones are selected and amplified.

The choice of the elements to be added in a construct depends on the aim of the experiment and particularly on the cell type in which the construct is expected to be expressed. The genetic code is universal even if some codons are used more effectively in a given cell type than others. The code that defines the activity of the regulatory sequences is specific to each type of organism. The promoter from a bacterial gene is not active in a plant or an animal cell and the reverse is generally also true.

1.4.5 Gene transfer into cells

An isolated gene can be transcribed *in vitro* and its mRNA can also be translated in a cell-free system. This provides experimenters with a very small amount of the corresponding protein, which may be sufficient for some biochemical studies. This technique is quite insufficient for a number of studies, such as determining the biological activity of the protein *in vivo* or determining its structure by crystallization.

To be decoded effectively and translated into a protein, a gene must be transferred into cells, which by nature contain all the factors for transcription and translation.

The plasma membrane of the different cell types is a barrier that allows a selective uptake of compounds. In some cases, the molecules enter cells through pores that are open or closed in a controlled manner. Specific carriers may also transport given molecules to be transferred into the cell. In other cases, the molecule recognizes specific receptors on the outside of the plasma membrane and the formed complex modifies the membrane locally, leading to an internalization of the complex and of the membrane surrounding it. This process is called endocytosis.

DNA is a negatively charged and large-sized molecule. It cannot spontaneously cross the plasma membrane. This is a way for cells to protect themselves from foreign DNA that may be present in their

vicinity. Oligonucleotides added to cell culture medium or injected into animals can enter cells on condition that they are present at a relatively high concentration.

Various techniques have been designed to force DNA to enter cells and reach their nucleus. These different ways of transferring gene into cells have been grouped together under the name of transfection. Transfection is different from cell infection, which involves different mechanisms used by viruses to deliver their genomes into cells. The principle of these different transfection techniques is depicted in Figure 1.4. They all rely on various physicochemical phenomena.

1.4.5.1 Cell fusion

A plasmid can be transferred by fusing the protoplast of the bacteria with the cells to be transfected. This method is inefficient and rarely used. Another of the drawbacks is that all the genes of the bacteria are transferred to the cells.

1.4.5.2 Transfer of DNA–chemical complexes

The *in vitro* association of DNA with various molecules forming a complex that enters cells with some efficiency is the most commonly

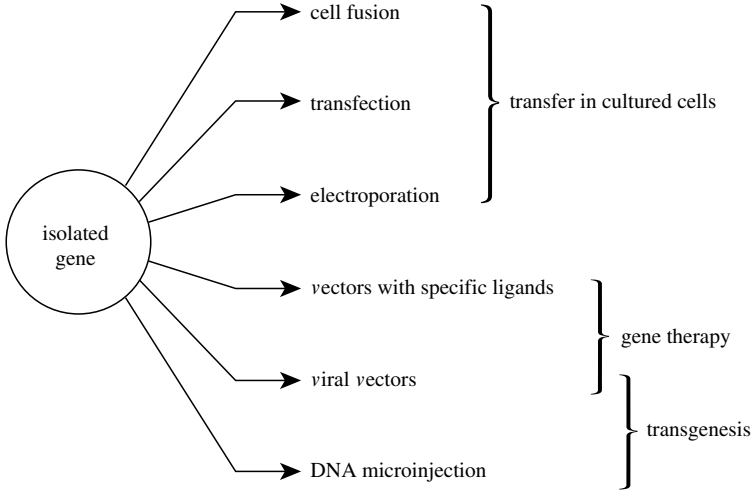


Figure 1.4 Different methods of gene transfer into animals cells

used method. Among these molecules is calcium chloride. The phosphate group of DNA binds calcium to generate an insoluble complex, which precipitates if an excess of calcium and phosphate is added to DNA. The mixture is added to the cell culture medium. A small proportion of the insoluble complex that covers cells is spontaneously endocytosed. DNA is resolubilized in cell cytoplasm. Most of the internalized DNA is degraded and a small percentage reaches the nucleus, where it is transcribed. The endocytosis may be amplified by adding various chemical compounds such as glycerol or dimethyl sulfoxide, which form a complex with water and reduce the cell content in water. This enhances the chance of the cell membrane invaginating and forming vesicles containing the DNA complex, which are internalized.

DNA may also form complexes with polycations (basic proteins or chemical compounds such as polyethylenimine). These polycations may be covalently linked to lipids. The phosphate groups of DNA bind to the polycations, which reduce the negative charge of DNA and spontaneously bind to the negatively charged molecules of the outer plasma membrane of the cell. This association induces the endocytosis of the complex. The presence of lipids in the complex induces a fusion with the plasma membrane and efficient uptake of the DNA by the cell.

DNA endocytosis may be targeted by using ligands that specifically recognize molecules at the surface of the cell. These ligands may be monoclonal antibodies, which can be raised to specifically bind a broad spectrum of molecules at the cell surface. In some cases, the ligands may be hormones, cytokines or molecules such as asialyloproteins, which have specific receptors on the plasma membrane. This approach implies that DNA is strongly associated with the ligands, including by covalent binding.

1.4.5.3 Electroporation

This method consists of subjecting cells to an alternating electric field. This creates transient pores in the plasma membrane. DNA added to the electroporation medium can enter cells through the pores. The electric field also induces DNA mobility and favours its uptake by the cells. This method may be quite efficient and it is being used more especially with the cell types in which the uptake of DNA–chemical complexes does not occur at a sufficient rate. A number of cells are destroyed under the effect of the electric field. Yet, it is a good method for generating clones having stably integrated the foreign DNA. Electroporation is the best method

to transfer genes into ES cells (embryonic stem cells) and replace an endogenous gene by homologous recombination.

1.4.5.4 Infection by viral vectors

Various viral vectors are used to transfer genes into cells. The principle of this method is essentially the same for all the viral vectors. Some of the essential genes are deleted from the viral genome. This generates a viral genome capable or incapable of autoreplicating. This also makes space in the viral genome to introduce foreign genes. These recombined genomes have become incapable of generating functional viral particles, since essential viral proteins are missing. The recombined viral genomes have to be transferred into cells which transiently or stably express the missing viral genes. These cells are called transcomplementing cells. They are capable of synthesizing viral particles containing the foreign genes. The particles, which are secreted in the culture medium, may be used to infect cells and transfer the foreign genes (Figure 1.5).

Several types of viral vector are currently being used and studied. Those containing the adenovirus genome have a high potency to infect cells either *in vivo* or *in vitro*. This genome is rarely integrated into the cell genome. Retroviral vectors infect essentially cultured cells. Their genomes are integrated into the host cell genome. Other vectors that are described in the gene therapy section are also implemented.

The adenoviral and retroviral vectors are tentatively used for gene therapy. They are also designed to transfer genes into cell types for which no other method has proved to be satisfactory. Adenoviral vectors are more and more frequently used by experimenters to transfer genes into given organs of an animal. This makes it possible to evaluate the effects of the gene. This approach is in some ways a prelude to or a substitute for transgenesis. Indeed, infecting an organ by an adenoviral vector is relatively easy and rapid. This may avoid the laborious production of transgenic animals or on the contrary urge researchers to obtain transgenic animals expressing the foreign gene in a stable way.

1.4.5.5 DNA microinjection

DNA in solution can be microinjected directly into the cell cytoplasm or nucleus. This protocol is laborious and requires special equipment (microscope and microinjector) and specific training.

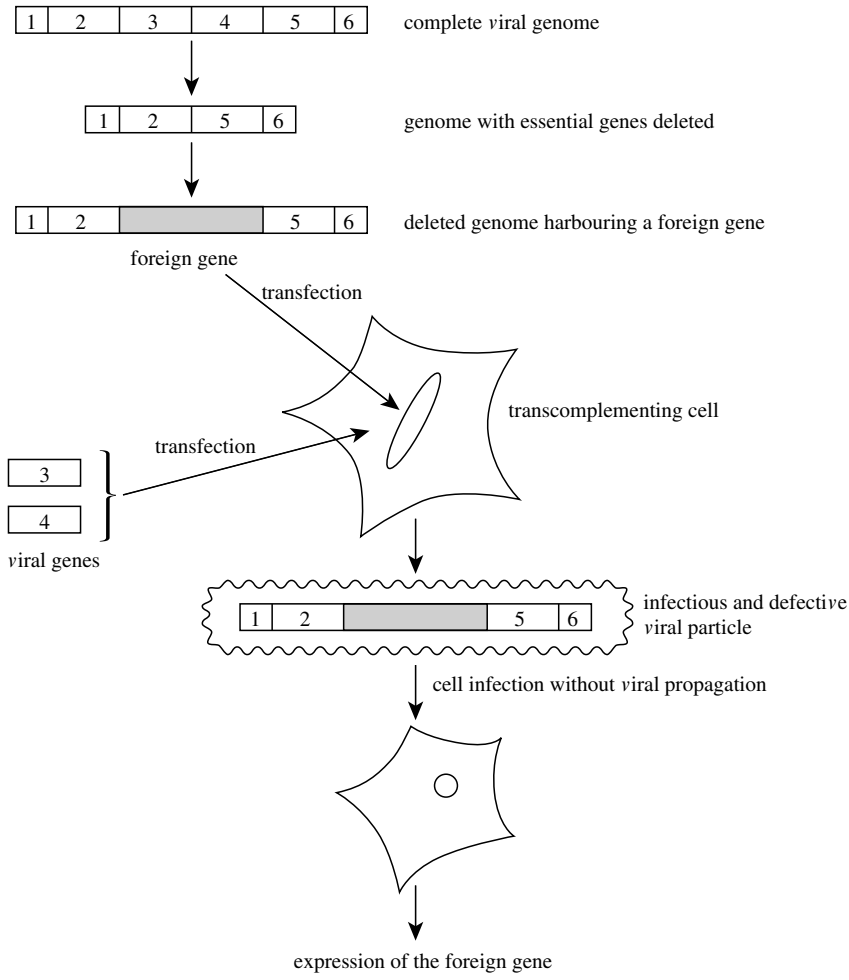


Figure 1.5 Principle of viral vectors. Genes required for virus propagation are removed and replaced by foreign genes of interest. A defective viral genome has to be complemented by a wild virus or by transcomplementing cells that synthesize the proteins coded by the genes deleted from the viral genome. The viral particles produced by the cells may infect cells and transfer their genes without propagating

All these methods of gene transfer are used according to their efficiency and the targeted cell type. Transfection of DNA–chemical complexes and electroporation are generally appropriate for gene transfer into cultured cells. Viral vectors were originally designed for gene therapy and may be quite useful in some cases for gene transfer into cultured cells

or organs *in vivo*. Vectors based on the use of specific ligands are implemented essentially for gene transfer *in vivo*. DNA microinjection into the nucleus is the most effective technique. It is used in some special situations in cultured cells. This is the case with cells for which the other gene transfer methods are ineffective. DNA microinjection is performed in individual cells. It therefore generates cell clones, which can be amplified or observed by non-invasive methods such as microscopy if the gene directs the synthesis of a protein that can be easily visualized, such as the green fluorescent protein.

DNA microinjection is the most frequently used method to generate transgenic animals.

1.5 The Systematic Description of Genomes

The first genetic engineering techniques made it possible to study a limited number of genes. The first problem to solve in most cases was isolating the gene. This is most feasible when the gene is highly expressed. The corresponding mRNA is then abundant and the cloning of its cDNA has every chance of being successful. The cloning of the cDNA provides experimenters with a probe, which may be used to clone the genome fragment containing the native gene. The sequence of the protein may designate the oligonucleotides, which may be used as probes to clone the corresponding cDNA. This method is still being used, specially when the gene of a given species is needed and the same gene is already known in related species. A set of oligonucleotides must be synthesized and tested until the most conserved sequences make it possible to identify the cDNA from a bank or amplify it by PCR.

Cloning a gene that is totally unknown but whose existence is proved by its effects is becoming more and more frequent. This method is mainly based on the use of hyper-variable regions of the genome. Most of the time, these regions are microsatellites. Microsatellites are for example sequences composed of 12–22 GT, which are present in most parts of the genome. The existence of these sequences seems to result from uncorrected errors in DNA replication. The errors are frequent and they generate the hyperdiversity of microsatellites. Microsatellites have no known function and the conservation of their sequence is not subjected to any evolutionary pressure. This also generates hyperdiversity. In some cases, a microsatellite is formed in the functional region of a gene. The action of the gene may be altered by the presence of the microsatellites and it may even generate a

genetic disease in some cases. The positioning of microsatellites in genomes is one of the essential steps of genome mapping.

The multiplication of the animals that have or do not have the genetic trait under study may create families composed of individuals that bear or do not been the unknown mutation. The trait is easily detectable, especially if it is monogenic. In the best cases, it is possible to establish a correlation between the genetic trait in each individual and the existence of microsatellites.

In practice, known microsatellites are amplified by PCR using primers corresponding to the sequences surrounding them. The amplified microsatellites are visualized by electrophoresis. The size of each microsatellite directly reflects its diversity. The correlation between the genetic trait in question and the size of microsatellites results from the fact that during meiosis, the chromosome rearrangement leads to the cosegregation of the gene responsible for the phenotypic effect and the microsatellites. The number of microsatellites to be examined cannot be predicted. A sufficient number of these sequences must be studied to establish a robust correlation with the phenotypic effect.

The establishment of this correlation may generally be used for selecting the animals carrying the genetic trait of interest. This simplifies selection and makes it more precise. Indeed, the microsatellites related to the gene of interest may be examined at any stage in the life of the animals, starting from a few cells, or even a simple cell. The breeding of the animals classically needed to observe the genetic trait for selection is therefore no longer required. This reduces the cost of genetic selection. This method is also more precise. Indeed, if a sufficient number of microsatellites is examined, the selection of the gene of interest may involve a shorter region of the genome than the classical selection based on the observation of the phenotypical property of the animals. The selection by microsatellite markers may thus reduce the number of coselected genes that are not involved in the expression of the genetic trait but have potentially undesired effects. The method of selection may also help reduce the loss of biodiversity that results from coselection of regions of varying length of the chromosome carrying the gene of interest. It is also important to note that selection by microsatellite markers does not imply that the gene of interest has been previously identified. This means that although the selection method is not completely precise, since it is not based on the examination of the gene responsible for the phenotypic effect, it is simple and reliable, once the microsatellites related to the gene of interest have been identified.

The identification of microsatellites related to a gene responsible for the expression of a given genetic trait can be followed by the identification of the gene itself. A correlation between microsatellites and a gene of interest can be established when the distance between the microsatellites is of 10 C Morgans or even more. This corresponds to about 1 million kb or more. The region of the genome defined in this way is too large to directly identify and clone the gene of interest. Other genetic markers located in the same region must be found. These markers may be additional microsatellites but also genes. The growing knowledge of genome structure in humans and several other mammals facilitates the positioning of markers in the region of interest. Indeed, a given gene often has the same neighbouring genes in related species. This is particularly true between mammals and even vertebrates. When the mapping is known in more detail in the region of interest, the techniques of molecular genetics can be implemented. The identified markers can be used to identify the BAC vectors from a genomic bank that harbours the markers and then potentially the gene of interest. A return to the family of animals can reveal which of the markers are the most frequently transmitted to progeny having the expected phenotypic characteristics. These markers are the closest to the gene of interest. This search can finally determine which BAC vector harbours the gene of interest. This vector can be fully sequenced to identify the genes it contains. This approach is named positional cloning. The vector may also be transferred to cultured cells and to mice to determine whether its presence induces properties similar to the gene of interest. To confirm that the identified gene is responsible for the phenotypic property of the animals, the same gene can be knocked out by homologous recombination in mice (see Section 2.3.6). The resulting biological effects may provide additional information on the role of the gene.

This protocol is being applied to plants as well and to humans. In the latter case, the situation is usually far more complicated. Indeed, establishing families showing a phenotypic characteristic is a difficult task, since reproduction in humans is slow and only existing individuals can be solicited for such studies. In practice, the method is implemented in humans to identify unknown genes having a major impact in genetic diseases. Many of the gene mutations involved in a human disease have been and still are identified in this way.

Identifying a gene having a major role in the expression of a genetic trait can be followed by multiple applications (Figure 1.6). In animals and plants, the sequencing of the alleles of the gene allows a direct

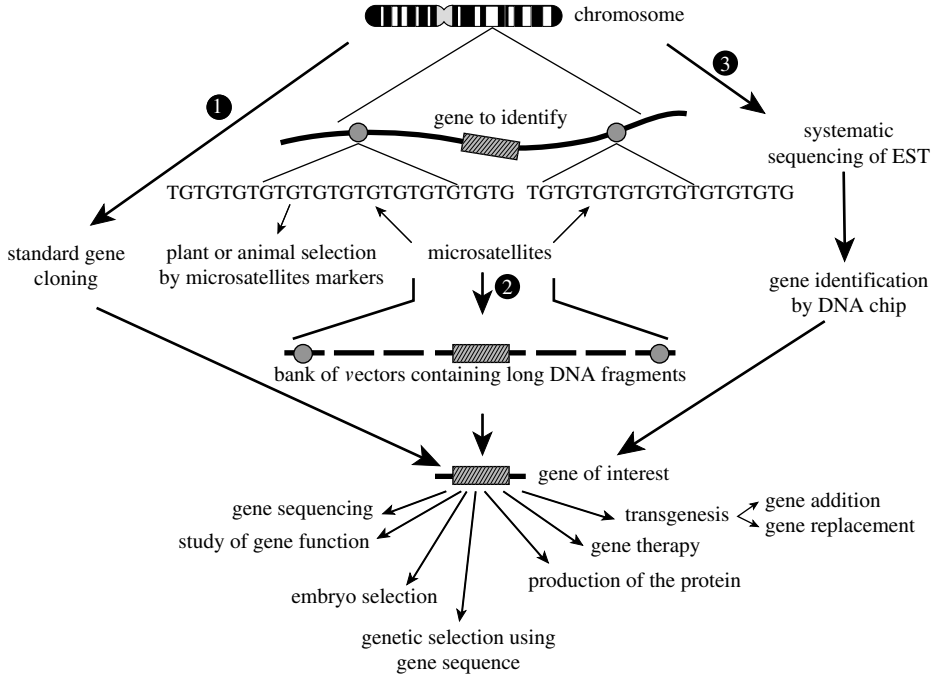


Figure 1.6 Systematic gene study. L.M. Houdebine Medecine/Sciences (2000) 16: 1017–1029 © John Libbey Eurotext. The classical method for gene cloning (1) is now followed by positional cloning based on the presence of microsatellites in the vicinity of the genes (2). Systematic sequencing of EST (expressed sequence tag) and genomes will eventually lead to the identification of all of the genes of a few living organisms (3). The study of gene function and regulation often includes transgenesis

selection of the individuals having the genetic trait of interest. This can be achieved in newborn animals but also potentially in embryos. In humans, the same methods can determine which of the embryos generated by *in vitro* fertilization harbour a mutated gene responsible for a severe disease. One cell is sufficient for this test since the genomic region bearing the mutation can be amplified by PCR and the mutation of the amplified fragment can be determined by restriction mapping or by sequencing.

The isolated gene can be used to study its biological role *in vitro* and *in vivo*. The coding sequence of the gene or of the corresponding cDNA can be introduced into an expression vector to produce the corresponding protein. Small amounts of protein can be obtained from bacteria expressing the gene. This may be sufficient to study the biochemical properties of the protein including via crystallization and X-ray diffraction. Large

amounts of the protein can be prepared on an industrial scale to be used as a pharmaceutical if this appears justified. The coding sequence of the gene can also theoretically be used for gene therapy.

The method described above to identify a gene of interest does not require that any of its elements be known but it implies that one of its major effects has been depicted. A more systematic approach is under way for a certain number of species. It consists of sequencing their whole genome and all their cDNAs which are named the EST (expressed sequence tag) (Figure 1.6). Most of the genes of a genome can be identified in this way. This identification is complicated for higher organisms by the large size of the genomes, which contain many sequences not corresponding to genes, and which are often repeated. The whole transcribed gene sequences of a genome is named the transcriptome.

Gene sequences are therefore established without any prior hypothesis as to their role. Determining the role of these numerous unknown genes will take decades. In some cases, the sequence homology of a newly discovered gene with another or with other already known genes in different organisms may reveal some of the likely functions of this gene. Indeed, the protein coded by the gene may contain typical protease or kinase enzymatic sites and this provides researchers with clues for determining the function of the gene.

The different cDNA sequences identified by their systematic sequencing can be used as probes to determine in which cells the corresponding genes are expressed. For this purpose, several protocols are being used. The oligonucleotides containing a region of each cDNA can be bound to a solid support. The cDNAs obtained by reverse transcription of the whole mRNAs of a given cell type may be labelled by a chemical marker and added to the support containing the oligonucleotides. The cDNAs hybridized to the oligonucleotides can be identified by an automatic system. This gives the pattern of gene expression in a given cell type and between different physiological states of a given cell type. Comparisons between different cell types lead to the identification of the genes potentially responsible for cell differentiation, hormone action or tumour generation. After this systematic search, experimenters have in hand numerous genes that are candidates for a given physiological event. Figure 1.6 indicates how the biological functions of the genes can be determined. When this systematic approach, named reverse genetics, is implemented, gene transfer into experimental animals and gene knockout are particularly important since they are expected to give the first indications on the role of the gene in the organism.

A systematic survey of the proteins present in the different cell types of an organism provides additional information on gene expression. This approach, named proteomics, is complementary to the systematic identification of the mRNAs. It is closer to the biological effects of genes, since the same gene can generate several proteins having distinct biochemical and biological properties.

1.6 Classical Genetic Selection

After the discovery of Mendel's laws of heredity, it has become less empirical and thus easier to select living organisms. Selection is in practice complex since a given genetic trait often does not depend on a single gene having a dominant effect. Selection must therefore be carried out by favouring, through reproduction, the emergence of genes located on different chromosomes.

This approach relies in all cases on the screening of the mutations that occurred spontaneously and have a dominant phenotypic effect transmissible to progeny. Selection of animals is thus based on measuring some of the parameters that characterize the function of interest: size of the animals or milk production for example in domestic animals, behaviour in pets, developmental defects in laboratory animals etc. The identified animals are reproduced to establish stable lines of individuals all exhibiting the genetic trait of interest.

The correlation between the size of microsatellites of the genetic trait may be used to identify the individuals bearing the mutation. The examination of microsatellites is simpler, faster and more precise, as depicted above (Section 1.5).

This method has been applied successfully to identify the gene responsible for the hyperprolificacy in Booroola Merino ewes. The mutated gene is the BMPR-1B gene, which is involved in ovulation (Mulsant *et al.*, 2001). The hyperprolific animals can now be selected by identifying those having the mutated allele of the BCPR-1B gene. Studies are currently being conducted to decipher the mechanism of action of this gene. This may not only provide interesting information on the mechanisms controlling ovulation in mammals; it may also help define new methods to enhance fertility in animals and in humans and to generate new contraceptives. The mutated allele of the BMPR-1B gene may also be transferred to non-hyperprolific sheep and also to goats, cows, pigs and perhaps other domestic species to tentatively enhance their fertility.

Another example of selection performed on the basis of gene structure is that of lactating cows. It has been known for decades that cow milk has variable protein composition and this trait is inheritable. According to the concentration of the different caseins (the major milk proteins), the protein concentration in milk varies as well as the quality of the curd to prepare cheese. A selection was performed for years by identifying the different caseins in milk. This mode of selection was efficient but very slow, specially for bulls, which have to generate females that themselves have to be in lactation before the bull genes coding for caseins can be identified by the structure of the milk proteins. The selection based on the structure of the different casein alleles is now common practice.

The identification of the mutated alleles responsible for a genetic trait of interest is therefore a real progress for animal selection. Yet, this method remains strictly dependent on spontaneous mutations, which occur with a low frequency or not at all during the reproduction cycle.

1.7 Experimental Mutation in Genomes

Spontaneous mutations are rare in each reproduction cycle. This frequency is compatible with an efficient selection of naturally mutated microorganisms. For pluricellular organisms and mainly those having a slow reproduction rate, the experimental induction of mutations is theoretically helpful since, in this way, each reproduction cycle generates a much higher number of mutants. Several techniques can induce mutations in most living organisms.

1.7.1 Chemical mutagenesis

It has been known for decades that a certain number of chemical compounds induce mutations in DNA of various species. These substances are known to be carcinogens. Irradiation by γ- or X-rays also induces mutations in DNA and cancer.

ENU (ethyl-nitroso-urea) is one of the chemical compounds classically used to induce mutations in microorganisms. This reduces markedly the screening of the clones having the expected biological properties.

These methods are also used to generate mutants in plants. This is possible since screening, reproduction and multiplication of the selected individuals is rapid in a number of cultured plants.

A similar approach has been extensively applied for genetic studies in *Drosophila*, which also has a short reproduction cycle and a high prolificacy. This method, which proved particularly fruitful, was considered for years as inapplicable to other species, namely for mammals. A research programme under way aims to use ENU as a mutagenic agent to generate new phenotypes in mice, to establish lines of animals bearing the inheritable mutations responsible for a phenotypic effect and to identify the alleles generating this effect (Brown and Balling, 2001).

Several reasons converged to make it possible to extend this approach to mice. The reproduction cycle in mice is particularly short for a mammal. Mice are hyperprolific animals. Their small size means that their breeding is inexpensive. Genome mapping has been completed in this species and mouse genome sequencing is almost achieved. This makes it possible to identify the genes involved in heritable genetic trait of mutants. Transgenesis including gene addition and replacement is also possible in mice. However, this project is ambitious and requires the collaboration of several institutes. It involves breeding a very large number of animals and observing the animals to identify abnormal biological properties using sophisticated investigation methods. After a few years, it appears that this method is fruitful in mice, as in other species having a simpler genome. It is interesting to note that several genes involved in deafness have been identified in this manner. This should help us understand this type of disease in humans.

Mutagenesis by ENU followed by a systematic screening of the mouse genome may in some way replace the more classical approach of gene knock-out by homologous recombination. Indeed, ENU inactivates genes in a non-targeted manner and the inactivated genes can be identified in a number of cases if they have a phenotypic effect. This systematic approach may make it possible to discover genes that were not suspected to have a given biological effect. In contrast, gene knock-out by homologous recombination implies that the gene is already identified. ENU is highly mutagenic and a number of genes are knocked out in the same genome, as opposed to what occurs with gene replacement by homologous recombination. This may greatly complicate the interpretation of the phenotypic effects observed in the mutagenized mice.

1.7.2 Mutagenesis by integration of foreign DNA

Gene transfer into cells or animals followed by its integration into the host genome is in essence mutagenic. Indeed, the foreign DNA alters genome integrity at integration site and perhaps elsewhere.

Foreign DNA is generally introduced by microinjection in mammal embryos. It is known that the integration rate of foreign DNA is low and it often occurs in non-functional regions of the genome, leading to no phenotypic effect.

The rate of integration may be greatly enhanced by the use of transposons (Figure 1.7). The first transposon sequence is integrated into one or a few sites of the genome. By definition, transposons can replicate and integrate in multiple sites of the genome, generating numerous mutations. This approach is being used successfully in plants. Thousands of arabidopsis clones having insertional mutations generated by a transposon have been prepared in this way and are under study.

A variant of this method relies on the use of retroviral vectors in animals. These vectors infect cells and integrate the host genome very efficiently. This avoids the microinjection step required for transposons.

An even more sophisticated way of generating mutants by foreign DNA integration relies on the use of vectors designed to trap genes. These vectors are described in Section 2.3.10. The principle of these tools

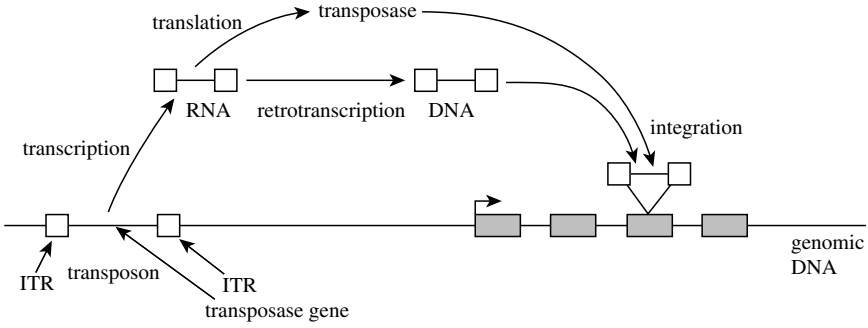


Figure 1.7 Gene mutation by transposon insertion. Transposons are genome DNA sequences that are transcribed in RNA and then retrotranscribed in DNA, which integrates into another site of the genome. Transposons contain sequences for integration known as ITR (inverted terminal repeat) and the gene coding for integrase, which inserts the transposon into the genome. The introduction of a transposon into a genome may inactivate a gene. The targeted gene may be identified by the presence of the transposon sequence

is the following. The vectors contain for example no promoter, a sequence coding for a reporter gene and a transcription terminator. The vector is introduced in mouse embryo stem (ES) cells and the transfected cells are used to generate chimaeric animals. The individuals in which the reporter gene is expressed have integrated the vector within an active gene. If this operation is accompanied by an alteration of the biological properties of the animal, the chance is high that the targeted gene, which is inactivated by the presence of the vector, is involved in the normal expression of the altered function.

The mutations induced by chemical compounds or irradiation are not easily detectable and the method described in figure 1.6 must then be implemented. The mutagenesis induced by foreign DNA integration offers an interesting advantage in this respect. Indeed, the foreign DNA has a known structure, which has no equivalent in the host genome. This foreign DNA can be cloned by conventional methods and the targeted gene sequence is contained in the same clone.

The mutation frequency obtained by these methods remains low and the mutations occur essentially randomly. The study of the mutated gene then usually starts without any hypothesis. This complicates the determination of the targeted gene function. One advantage is that the random process will reveal the existence of genes the function of which is unknown and would not have been subjected to any study without this fortuitous identification.

1.7.3 Mutagenesis by transgenesis

Independently of the host genome alteration that may be induced by the integration of a foreign DNA, transgenesis is essentially a mutation, since it provides the host genome with additional genetic information (or inactivates an endogenous gene).

Transgenesis is one of the essential steps in identifying the function of a newly discovered gene (Figure 1.6). In this case, the effects of the transgene are largely unpredictable.

The same is not true when the biological properties of the transferred gene are known. The gene is then transferred to obtain a known effect in the animals: resistance to a disease, improved growth or fertility, secretion of pharmaceutical proteins in milk etc. In these cases, selection of the animals by transgenesis is more precise and less hazardous than conventional selection. Indeed, the transgene provides essentially one

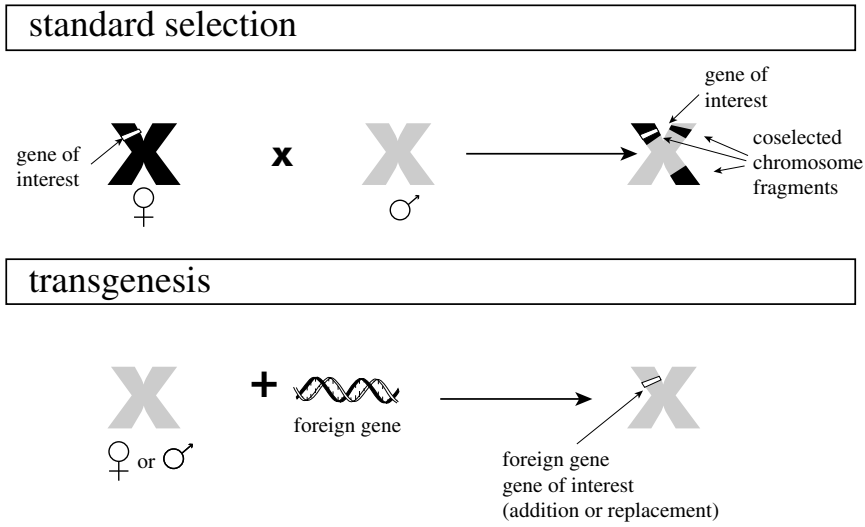


Figure 1.8 Classical selection and transgenesis. Classical selection relies on the evaluation of the biological effects of an unknown gene that has mutated by chance. Selection via sexual reproduction coselects large chromosome regions whose contents and effects are unknown. Transgenesis adds a gene that may be known and that may have known and predictable effects. Transgenesis does not greatly alter the host genome. All of the biological effects of a gene, even of a known gene, cannot be predicted

piece of known genetic information with no or limited alteration of the genome. No deleterious genes are coselected by this process (Figure 1.8). Transgenesis also provides the genome with foreign genetic information, which would never appear after classical selection. All these properties have made transgenesis particularly attractive to create models for studying human diseases and to improve animal production for human consumption. However, the genetic modifications resulting from transgenesis may have side-effects that would not appear after classical selection. This implies that the transgenic animals and their products must be examined with special care before becoming available for human consumption. This point is discussed in Section 4.2.4.

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

2 Techniques for Cloning and Transgenesis

2.1 Cloning

2.1.1 The main steps of differentiation

Cloning is, by definition, the reproduction of a cell and, more generally, of a whole living organism without any modification of its genotype. Basically, non-sexual reproduction is cloning. This phenomenon occurs widely in nature since bacteria and yeast reproduce according to this process. Indeed, the genome of a unicellular organism is replicated without any modification except the errors that have not been corrected. The same situation is encountered with the somatic cells of pluricellular organisms, with the exception of those of the immune system that synthesize antibodies and T receptors. Indeed, the genes coding for immunoglobulins are generated by the random assembly of elements after an immunization. The cells containing genes that synthesize the immunoglobulins that recognize the antigen used for the immunization are saved and amplified, whereas the other cells are eliminated. This genome rearrangement is quite specific to a category of lymphocytes and concerns only a rather limited and specialized region of the genome.

A number of plants can multiply independently of sexual reproduction. Natural layering is a very common strategy of reproduction for many plants. It does not exclude sexual reproduction. Cuttings are also widely used by gardeners to replicate plants without any sexual reproduction. The resulting organisms are clones. The same is true when the bulb of a plant is divided into fragments, each of them being capable of generating a normal plant. Plants have the capacity in some

circumstances to induce root development, which allows the differentiated branch to continue living and even to grow.

Cutting and layering, as well, may generate mutated plants. A plant such as a geranium may have branches with flowers of different colours or shapes. The cutting of these branches results in plants bearing the mutation responsible for the phenotypic modification. This operation corresponds to the cloning of a spontaneous mutant.

No phenomenon of this kind has ever been observed in vertebrates. The development mechanisms in plants and animals are different. This may explain why cloning in plants can be obtained by a simple prolongation of the differentiated state whereas this is not true in animals.

In higher animals, it is recognized that the initial cell of the zygote formed just after fertilization is totipotent. This means that this cell can generate all the cells of the organism. The same is true for the two or four cells of the young embryo. Each of these cells can develop to generate a living organism when introduced into the zona pellucida of an oocyte. These cells are therefore still totipotent. Beyond this stage, the embryonic cells lose their totipotency and become pluripotent. These cells, known as blastomers, are not differentiated, since each of them can participate in the generation of all the organs of the organism. However, these cells lose the capacity to generate a living organism after they are transferred into the zona pellucida.

After a few days, depending on the species, the embryo, known as a blastocyst, is composed of two categories of cells.

1. The cells forming a unicellular layer along the zona pellucida, known as the trophectoderm, which is the precursor of the placenta.

2. The cells forming a compact block, known as the inner cell mass, which is the precursor of the embryo itself (Figure 2.1).

The cells of the inner cell mass are pluripotent. The pluripotent cells can be experimentally transferred into a blastocyst. The transplanted cells participate in the development of the embryo, generating a chimaera. This kind of organism must not be confused with a hybrid. A chimaera results from a mixture of cells that are superimposed, each one keeping its own genome. The gametes and offspring of chimaerae are derived from only one cell type of the original embryo. Their genes are therefore not a combination of those from two embryos. In some cases, chimaeric animals

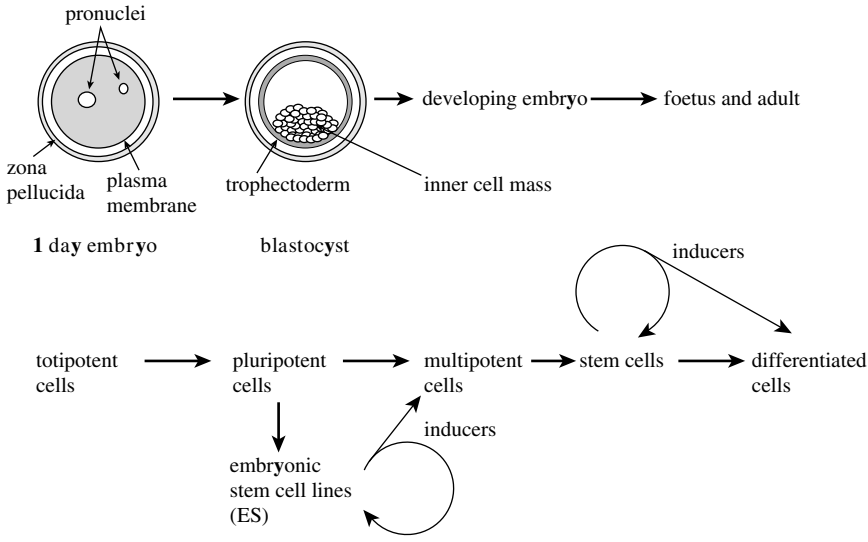


Figure 2.1 Different steps of development from embryo to adult. Cells progressively lose their potency as they differentiate. Stem cells are capable of self-replicating but also of differentiating under the control of specific inducers

may be obtained with cells originating from two different species. The animals are then formed by a mixture of cells from both species. Interspecific chimaerae have thus been obtained with goats and sheep, on the one hand, and with quail and chickens, on the other. The offspring from these chimaerae are sheep or goats but not a mixture of both.

Hybrids, in contrast, are formed by fertilization and thus by a combination of the genomes in all the cells of the organism. Hybrids are very frequently generated within the same species to improve production. Cultured maize is a popular example of a hybrid. Interspecific hybrids can also be obtained in some cases. In plants, this may lead to the creation of a new species. This is the case for triticale, which is a hybrid of corn and rye. Hybrids in higher animals are subfertile and do not generate new species. This is the case for hybrids resulting from the crossing of a horse and a donkey.

Beyond the blastocyst stage, the embryonic cells become progressively specialized. This process is known as embryonic differentiation (Hadjantonakis and Papaioannou, 2001).

During their formation, organs contain moderately differentiated cells, which are multipotent. Each of them can participate in the generation of a restricted number of cell types. The cells of the three major

compartments of the embryo, the endoderm, the ectoderm and the mesoderm, are still moderately committed and are therefore multipotent.

Early in embryo development, a group of cells remains poorly differentiated. These cells migrate to the genital crest to generate sexual organs and, finally, gametes. The precursors of gametes are called primordial germ cells. This process makes it possible to avoid multiple cell multiplications and, thus, mutations. In this way, the specific fate of primordial germ cells preserves the genotype of the species.

All or at least some of the organs in the foetus and in adults contain stem cells known as organ stem cells to avoid any confusion with embryonic stem cells. The organ stem cells are multipotent. They can replicate at a low rate without being subjected to a differentiation process. Organ stem cells can divide rapidly and become differentiated to regenerate an organ. A typical example of organ stem cells is bone marrow cells, which are maintained during the entire life of the organism and give rise to the generation of white blood cells (lymphocytes, mastocytes, macrophages, platelets etc.) and red blood cells. These organ stem cells are therefore multipotent. In contrast, sperm precursors are unipotent organ stem cells, since their differentiation results in only one cell type, the sperm.

Until the last few years, it was widely acknowledged that differentiation was a multistage process, which was essentially irreversible in vertebrates. However, it was known that tumour cells lost their differentiated state and acquired some of the characteristics of foetal cells.

Experiments carried out 50 years ago showed that the situation was different in plants. Indeed, cells from the meristem from some adult plants can completely dedifferentiate *in vitro* and give rise to totipotent embryonic cells. This phenomenon is induced by adding factors to the culture medium of the cells. Each of the differentiated cells can generate a normal plant. Many cloned plants can thus be obtained from a single one. This property of plant cells is extensively exploited in agriculture. It makes it possible to cultivate plants whose genome is the best adapted to food or flower production on a very large scale. The capacity to clone plants by dedifferentiation of somatic cells is also a key element in the generation of transgenic plants. Indeed, foreign genes are transferred by the biolistic method or with the *Agrobacterium tumefaciens* vector into plant cells, which are further used to regenerate cloned transgenic plants.

In vitro dedifferentiation of vertebrate cells by the addition of factors to the culture medium has not yet been obtained. However, animal cells can partially dedifferentiate under particular experimental conditions (see Section 3.1.4).

2.1.2 Cloning by nuclear transfer

About 40 years ago, cloning of a laboratory toad, *Xenopus*, was achieved by a technique that was essentially the same as used to generate the ewe Dolly.

The technique is based on a simple idea. Oocyte cytoplasm must be capable of favouring or inducing cell totipotency. Indeed, the sperm nucleus is transcriptionally inactive. Sperm DNA is covered by basic proteins, the protamines, and chromatin is highly condensed. Soon after fertilization, the sperm nucleus is decondensed. Protamines are replaced by histones, leading to a progressive activation of the genome a few days later, depending on the species. Oocyte cytoplasm thus has a strong capacity to reprogramme a genome for embryo development. Many genes are required for embryo development and massive gene expression is observed in totipotent and pluripotent cells. It is not clear up to what point the expression of all these genes is required in early embryos. It is acknowledged that differentiation is accompanied and probably caused by a selective extinction of specific genes and by the activation of others. This leads to a pattern of gene expression that characterizes each differentiated cell.

The dedifferentiation of a somatic cell to give a totipotent cell implies the reactivation of numerous silent genes. The mechanisms involved in this phenomenon are not known. The phenomenon is referred to as genome reprogramming. Indeed, the gene expression programme of the somatic cell must be erased and replaced by that of a totipotent cell.

This reprogramming occurs in a relatively simple manner in plant cells due to the action of factors added to the culture medium. This phenomenon has never been observed in animals but it has been suggested that it might be possible using the cytoplasm of oocytes.

Isolated nuclei from *Xenopus* embryonic cells were transferred by micromanipulation into the cytoplasm of enucleated oocytes. Enucleation is imperative to eliminate the oocyte genome and to prevent the embryo from becoming triploid after the transfer of a diploid nucleus. Enucleation was carried out by aspiration using a micropipette. These pioneer experiments were not successful. It was postulated that this was due to the utilization of an isolated nucleus, which may have lost essential components during the micromanipulation process.

Another strategy was used (Figure 2.2). It consisted of injecting a diploid embryonic cell between the zona pellucida and the plasma membrane of the enucleated oocyte. This first step was followed by a fusion of

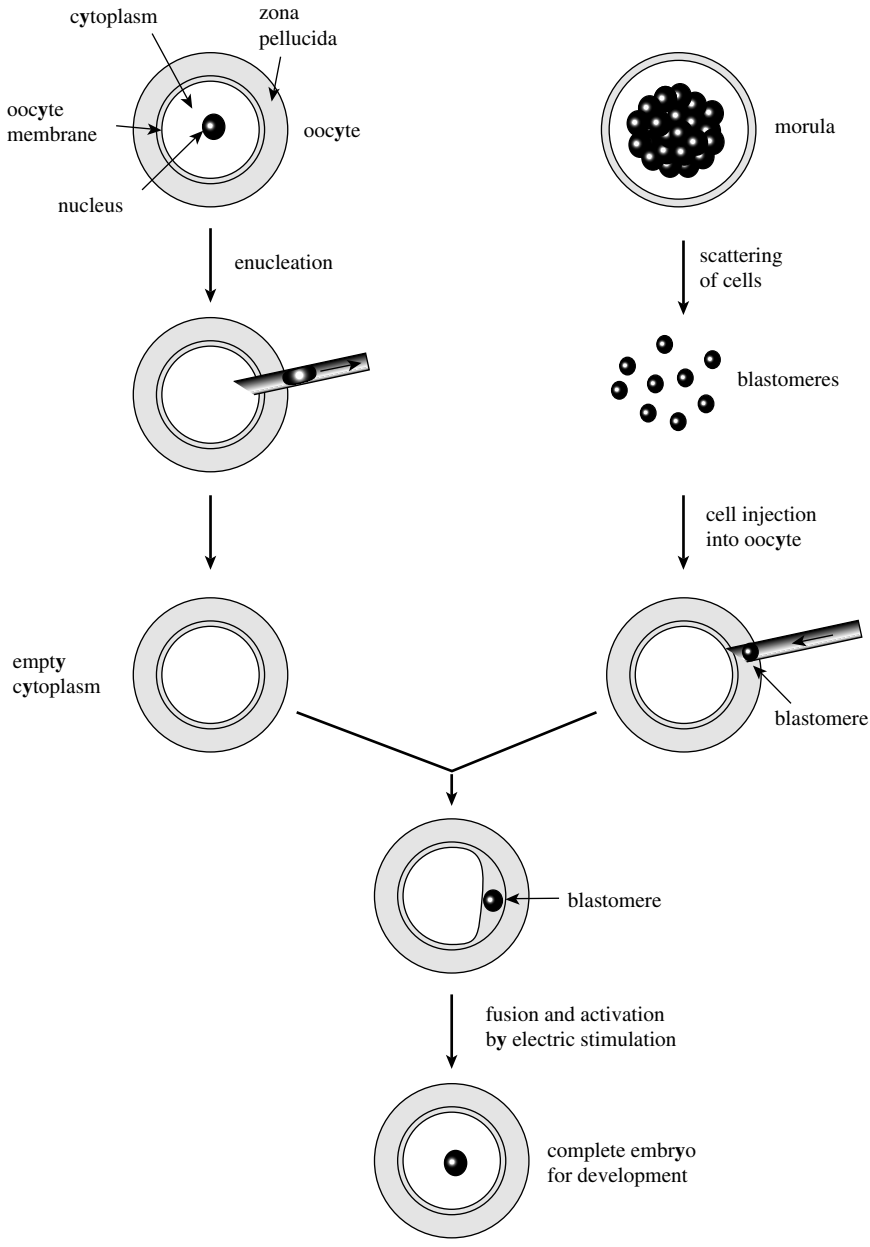


Figure 2.2 Principle of animal cloning by nuclear transfer. L.M. Houdebine. Medecine/Sciences (2000) 16: 1017–1029 © John Libbey Eurotext. A blastomere isolated from a morula or a blastocyst is inserted between the zona pellucida and cell membrane of a previously enucleated oocyte. An electric field induces a fusion

the plasma membranes of the cell and the oocyte. This was achieved by subjecting the material to an alternative electric field. This treatment destabilizes the cell membranes, which spontaneously fuse under these conditions. The nucleus of the cell is thus transferred to the oocyte cytoplasm without any micromanipulation. The electric field has another function, which is essential. Fertilization is not only the transfer of the sperm genome into the oocyte. Sperm also contains factors that induce calcium uptake by the oocyte. This is an essential signal to trigger embryo development. The electric field creates pores in the oocyte membrane, making it possible for the calcium in the medium to enter the cell. This mimics the activation of the oocyte by sperm. The reconstituted embryo can thus start its development. *Xenopus* was cloned in this manner 40 years ago.

This pioneer work was followed by other work done on lower vertebrates and invertebrates (Di Bernardino, 2001). The aim of these studies was the understanding of the mechanisms involved in cloning performed under these conditions.

About 15 years ago, it appeared important to extend the cloning technique to farm animals to accelerate genetic selection as was already the case for plants. The protocol defined for *Xenopus* also worked in sheep and cloned lambs were obtained in this manner. The clones were obtained, in all cases, after a transfer of nuclei from freshly prepared embryonic pluripotent cells. This early stage of development was chosen considering that the cells had just started their differentiation and that the reprogramming of their genome was the easiest in this physiological situation. The experiments carried out in *Xenopus* showed that cloning was not possible when the nucleus donor cells were taken from an adult. The same phenomenon was expected and also observed in mammals.

The success rate of these experiments was and is still low, even if it has been slightly improved after 15 years. About one per cent of the embryos reconstituted by nuclear transfer give rise to normal animals. This is insufficient to accelerate genetic selection. Indeed, the genotype of the embryo from which the donor cells are isolated is not known. Only the genotype and phenotype of its progenitors are known. The reproduction of genetically identical animals of unknown genotype is not justified for

---

of the membranes of the two cells and activates the embryo-like structure generated by the nuclear transfer. The same protocol may be implemented with pluripotent, multipotent and differentiated cells. In mice, the isolated nucleus is microinjected into the cytoplasm of an enucleated oocyte

this purpose and the sheep cloning technique has not been used in breeding. The same observation was made for cows and goats. A few studies carried out on pigs, rabbits, rats and mice revealed that cloning with freshly prepared pluripotent cells was not as easy as in ruminants, for unknown reasons.

Cloning ruminants, and cows in particular, was attempted with cultured pluripotent cells considered to be ES-like cells. These cells can, to some extent, participate in the early development of chimaeric embryos, but they are progressively diluted and do not contribute to the formation of gametes. When these ES-like cow cells were transferred into enucleated oocytes, they resulted in the early development of the embryo *in vitro*. Some of the blastocysts transplanted into recipient females resulted in foetal development, which was interrupted during the first 2–3 months of pregnancy, in all cases. This led us to believe that cultured pluripotent cells had started their differentiation and were no longer capable of generating cloned animals.

To improve cloning results, modified protocols were adopted but with moderate success. One of the reasons why cloning fails so frequently may result from the discordance between the stages of the division cycle of the donor cells and the recipient oocytes. Indeed, it is easily conceivable that a donor cell at the mitotic phase does not deliver its chromosomes to the recipient oocyte in an appropriate manner. The phase of the donor cell is unknown in most cases and nuclear transfer is likely to generate non-viable embryos.

Several laboratories carried out experiments using synchronized cells. These studies were undertaken most notably by I. Wilmut and his colleagues at the Roslin Institute of Edinburgh. This group had the ambitious project of generating transgenic sheep via nuclear transfer (see Section 2.3.4). This was conceivable only if cloning results were greatly improved.

One of the classical ways to synchronize cell multiplication consists of depleting the culture medium of growth factors and of serum. The cells stop their division and enter the G0 phase. Most of the cells cannot survive for a long time under these conditions and are threatened by apoptosis. The addition of serum and growth factors to the medium reinitiates the division cell cycle in a synchronized manner. The cells can be used as nucleus donors at the different phases of the cell cycle over a period of at least one day. This was expected to identify the most appropriate phase to transfer the nucleus into metaphase II enucleated oocytes.

This methodological study initially revealed that sheep pluripotent cells cultured for several weeks were able to generate living lambs after nuclear transfer (Campbell *et al.*, 1996). Essentially, the same technique was used to generate lambs from foetal and adult cells. Cells used for this study were maintained at the G0 phase of the division cycle. This experiment led to the birth of several lambs from foetal cells and one, Dolly, from adult cells. These data demonstrated, for the first time in animals, that not only the genome of pluripotent cells but also that of fully differentiated cells had retained the capacity to generate normal animals (Wilmut *et al.*, 1997).

The success of this cloning was initially attributed to the fact that the donor cells were in the G0 phase. It is conceivable, although not easily understandable, that chromatin has a special configuration in G0 which favours genome reprogramming. It turned out not to be as simple as that.

The pioneer experiment of Wilmut *et al.* strongly stimulated the experimenters working in this field. A cloned cow was born using the same protocol about a year later (Vignon *et al.*, 1998). Other cloned cows, sheep and goats were obtained in the following months and years. Unexpectedly, the different groups involved in the cloning experiments were all successful, although each of them had its own protocol. It became clear that the G0 phase offered, at best, a slight advantage for cloning. A recent study even indicated that cow cloning was more efficient with cells in G1 rather than in G0 (Kasinathan *et al.*, 2001). It is interesting to note that, before Dolly's birth, experimenters were convinced that cloning was possible only using pluripotent cells and regularly failed with differentiated cells. They all had success with multipotent or differentiated cells without any fundamental modification of their protocol. It seems that the inhibition in the minds of the experimenters was replaced by confidence after Dolly's birth and that this change was an essential reason for their success. It appears, in fact, that cloning using differentiated cells such as nucleus donors has a very low success rate and therefore requires many nuclear transfers to generate a few cloned animals.

The birth of the first cloned sheep by the transfer of nuclei from somatic cells was successfully repeated and extended to two other ruminants, cows and goats, without any major adaptation. Attempts to clone mice started several decades ago and failed. Essentially, all the different possible techniques were implemented, including the transfer of isolated nuclei. The reasons for this failure are not known but it was attributed to the fact that the transcription activation of the zygote genome in mice

takes place very early (about one day) after fertilization and the activation of the clones. It was postulated that this phenomenon in mice occurred before reprogramming of the genome was achieved, leading to aberrant gene regulation and embryo development. It is interesting to note that genome activation occurs about four days after fertilization in ruminants and after two to three days in rabbits, and that cloning is currently easier in ruminants than in rabbits and especially in mice. It remains unclear whether the delay between embryo activation and genome transcription is really an important parameter for cloning.

Cloning in mice with nuclei from pluripotent or differentiated cells was finally successful about one year after Dolly's birth (Wakayama *et al.*, 1998, 2000). A new protocol was necessary to reach this goal. Nuclei were transferred to enucleated oocytes by direct microinjection of the isolated nuclei, rather than by cell fusion as is the case in other species. A specific method of zygote activation was also required. The activation was achieved by adding ions such as calcium for controlled periods of time to the culture medium after nuclear transfer. This protocol has now been adopted by several laboratories working with mice. Attempts to use it on ruminants led to low success rates.

Cloning in pigs received little attention from researchers for years. Several reasons may explain this fact. The manipulation of pig embryos is different from that of ruminant embryos. The high lipid content of pig embryos may be one of the reasons for this difficulty. The main reason is that the pig is a naturally highly fertile species and improvement of reproduction by embryo manipulation is not as crucial as it is in ruminants. Dolly's birth, in addition to the fact that genetically modified pigs might become the source of cells and organs for humans (see Sections 3.1.6 and 3.2.4), encouraged researchers to re-evaluate pig cloning. Three groups independently and simultaneously succeeded in obtaining cloned pigs starting from differentiated cells as a source of nuclei. It is interesting to observe that the techniques used in the three laboratories were different but essentially led to the same results. One of them used the technique defined for sheep, regardless of whether or not it started from cells at the G0 phase. The second group used the mouse technique, in which isolated nuclei are microinjected into enucleated oocytes. The third group proposed a more sophisticated approach. The genome reprogramming of the somatic nucleus and the development of zygotes may be considered as a quite distinct phenomena, implying different mechanisms. It may therefore be logical to separate these two steps experimentally. To reach this goal, the generation of zygotes was first obtained by the technique defined

for Dolly. After one day, a time which is supposed to be sufficient for genome reprogramming, the nucleus of the zygote was isolated by micromanipulation and reintroduced into the cytoplasm of a normal enucleated embryo. The nucleus of the cloned zygote was assumed to be in the best possible condition for its development. Indeed, zygote activation is considered to be a crucial step in development and it is assumed that this phenomenon is more easily achieved through fertilization than electric stimulation. It is not currently known whether this logical but laborious protocol offers real advantages over those used by other laboratories.

Cloning in rabbits was obtained years ago, starting with fresh pluripotent cells. However, for unknown reasons, success was inconsistent. Studies in progress in a few laboratories indicate that cloning in rabbits might be successful in the near future. Indeed, high development rates of cloned zygotes *in vitro* up to the blastocyst stage are more and more regularly observed.

A few modifications of the protocol finally led to success. The cell cycle is particularly short in rabbit embryos and the window of time for implantation is narrow. Taking these facts into account gave satisfactory results (Chesné *et al.*, 2002).

The cloning techniques depicted above are being extended to other species for different purposes: rats, dogs, cats, horses, laboratory fish and primates, including humans (Cibelli *et al.*, 2001).

One of the major characteristics of the cloning techniques presently used is their considerably low efficiency. Cloning is a completely artificial physiological situation and this may explain why it works so poorly. It is still remarkable that the genome of somatic cells has a much higher plasticity than anticipated. Differentiation must now be considered as a reversible process and the genotype appears essentially unchanged from embryo to somatic cells.

A growing number of studies are in progress to describe the phenomena which are regularly observed after cloning. One of the most striking findings is the discrepancy between the number of blastocysts obtained after nuclear transfer followed by *in vitro* development and the number of newborns that survive. Another surprising finding is the large number of late abortions that occur after nuclear transfer and the death of about 40 per cent of the foetuses and newborns. Several abnormalities are commonly observed, such as large offspring with an underdeveloped placenta. Other observations have been reported: thymic aplasia, kidney atrophy, fluctuations of body temperature, liver or heart hypertrophy, high leptin concentration, partial immunosuppression etc.

All of these syndromes were described years ago. They were observed when the first cloned animals were obtained by nuclear transfer using fresh pluripotent cells as donors. Generally speaking, foetal abnormalities tend to occur when normal embryos are kept in a culture medium for several days. Hence, these syndromes seem to result from embryo manipulation and are amplified by cloning and long-term cultures. For unknown reasons, goats seem immune to these phenomena. Indeed, a recent study showed that cloned goats were normal in all cases, even when the recipient oocytes were obtained after *in vitro* maturation (Reggio *et al.*, 2001).

The efficiency of cloning is highly dependent on the state of the donor cells. Some cell types give much better cloning results than others. Cells from cumulus are considered to be good nuclear donors cells whereas ES cells, although pluripotent, do not regularly result in a high rate of blastocysts and living mice after nuclear transfer. Foetal skin fibroblasts are the most frequently used cell type to generate transgenic cloned animals. Cells from older animals are not necessarily bad donors. Cells lose part of their capacity to generate cloned animals from the pluripotent to the somatic stage. Unexpectedly, batches of cells from the same foetal sampling but cultured under slightly different conditions show quite different capacities for generating clones after nuclear transfer (Renard *et al.*, 2002).

The examination of DNA from Dolly's cells revealed that her telomeres were abnormally short. It is known that telomeres are sequences located at the end of each chromosome. These sequences protect DNA from degradation by exonucleases. In fact, telomeres are constantly degraded and restored. The balance becomes negative as the cells age, leading to a degradation of chromosomes and to cell death after about 50 multiplications. Gametes that are subjected to a lower number of cell multiplications are protected from telomere shortening. Gamete maturation might also include a lengthening of telomeres to ensure a long life for offspring.

Dolly's telomeres were short but this was also the case for the donor cell line, which was derived from an old sheep and was cultured over a long period of time. It was prematurely suggested that cloning might generate old newborns. In all the cloned animals obtained after Dolly and in which DNA was examined, the length of telomeres was normal or longer than normal. This was also true for clones derived from a 17-year-old bull. It is also interesting to note that the two lambs born after Dolly was naturally fertilized have normal telomeres.

Several hypotheses have been proposed to explain these phenomena. The defect in cells may be of genetic origin. The genotype of each donor cell is unknown. Mutations that do not alter cell life may be deleterious for embryo development. It is conceivable that the number of mutations increases from the pluripotent to the somatic state. It is also possible that culture conditions accelerate the mutation rate. The genotype of each donor cell cannot be examined and the experimenter is forced to operate blindly. During normal reproduction, the primordial germ cells and the gametes multiply a reduced number of times, decreasing their probability of mutating. On the other hand, gametes have to multiply and mature with a haploid genome. Lethal mutations may be eliminated during this period. The same is not true for diploid cells used as donors. The genetic hypothesis is compatible with the late interruption of development in cloned foetuses and newborns. Indeed, the same number of genes is not required for the foetus, which is assisted by its mother *in utero* and must be functional after birth. The putative mutations in the donor cells may thus have only late phenotypic effects. The growing knowledge of genome sequences in several species, particularly in mice, may provide a better understanding of the mutation rate in somatic cells. It remains difficult to evaluate the importance of the mutations in somatic genomes and almost impossible to prevent these gene alterations.

A second hypothesis has recently received strong experimental support. This hypothesis implies that inappropriate genome reprogramming is the essential cause of failure in animal cloning. The defect is then of epigenetic rather than genetic origin. Among the epigenetic mechanisms that regulate gene expression is DNA methylation. It is well established that, in higher organisms, DNA is highly methylated in the cytosines of the CpG motifs. It is also known that methylated genes are generally not active. Selective DNA methylation specifically inhibits gene expression and this mechanism is believed to play an important role in differentiation. During gametogenesis, the genome is almost entirely demethylated and is remethylated in mature gametes. In the days following fertilization, DNA is demethylated in zygote genomes and remethylated at the blastocyst stage before implantation. The remethylation defines which genes can be active after implantation. Part of the remethylation process is a stochastic mechanism. Some parental genes may or may not be demethylated and are thus active or not, depending on the case. This mechanism constitutes the parental gene imprinting. In some cases, one allele is functional whereas the other is not. This mechanism is typically epigenetic since the extinction of an allele that may be phenotypically

similar to a mutation does not imply any mutation. A given allele may, in some cases, be active in an individual but not in its progeny, and the reverse. This mechanism is also dependent upon the sex of the individual in some cases. This was clearly observed for some transgenes as well.

DNA methylation is supposed to play other roles as well. Demethylated DNA in gametes is less sensitive to the mutations that result from a deamination of 5-methylcytosine. DNA methylation might play a role in chromatin organization during gametogenesis after fertilization and throughout development. Methylation of transposons and integrated retroviral genomes is a way to inactivate their transcription. The methylation of these sequences and transposons, in particular, is partly stochastic. This controls the methylation of the region surrounding the transposon to some degree. In this way, the activity of the genes located in the vicinity of transposons may be controlled by stochastic methylation.

A comparison of DNA methylation in somatic cells, ES cells and embryos generated by either fertilization or nuclear transfer has been made (Reik *et al.*, 2001; Rideout *et al.*, 2001; Dean *et al.*, 2001). Striking differences have been observed. DNA in embryos obtained by nuclear transfer is exceedingly methylated as a result of incomplete demethylation and from premature remethylation. This may hamper the expression of genes playing a key role in development. A few genes have been examined in this respect. The gene coding for IGF II, a growth factor important for embryo development, has been found to be abnormally methylated in foetuses whose development has been interrupted.

An examination of mouse ES cell genome also revealed that its DNA is much more methylated than that of the equivalent pluripotent cells of normal embryos. This may explain why cloning with ES cells as donors is less efficient than expected. These observations lead us to consider the possibility that animal development and life is more tolerant of a potent epigenetic phenomenon such as DNA methylation than previously believed.

The fact that DNA methylation is not a strictly defined event might also explain why culture conditions of donor cells have such a strong impact on their capacity to generate clones after nuclear transfer.

This strongly suggests that poor cloning results are mainly due to inappropriate genome reprogramming of donor cells, which results from epigenetic mechanisms and particularly from ill controlled DNA methylation. These hypotheses provide clues for better control of reprogramming. This problem might be remedied by using drugs known to

induce DNA demethylation, which may be added to the culture medium of donor cells. A systematic study of the culture conditions (temperature, composition of the medium, oxygen tension etc.) seems necessary in order to define those conditions which have the lowest negative impact on donor cells.

These observations suggest that cloned animals are genetically identical to the individual from which the donor cells were provided. They are not phenotypically identical, due to epigenetic mechanisms such as those described above. This corresponds to what has been commonly observed. Natural twins and clones show slight visible differences.

Cloned animals may be genetically identical to each other but not strictly to the nuclear donor. Indeed, mitochondria found in clones are those of the recipient oocytes but not of the donor cells (Evans *et al.*, 1999). Clones may be genetically identical only when the recipient oocytes come from the same female and not solely when they are obtained with cells from the same nuclear donor.

Preliminary experiments carried out on several species indicated that some individuals provide quite efficient donor cells or recipient oocytes. This phenomenon seems to be of genetic origin. The selection of animals that are best adapted to the cloning technique is possible, particularly when cloning is used to multiply the genomes in question. This task has just been undertaken and will be long, especially in cows.

Experiments carried out for the first time a decade ago indicated that the cells of cloned embryos could be used to generate greater quantities of other clones. This increases the potency of cloning. It is also conceivable but not clearly proven that a second passage of a nucleus in the cytoplasm of an oocyte might provide more complete genome reprogramming, leading to better cloning results.

The reprogramming of genomes seems to occur without any cell division. Cytoplasmic factors, and probably proteins in particular, are responsible, by interacting with DNA. It is conceivable that these factors, or at least oocyte fractions containing these factors, will be identified in the future. An incubation of these factors with isolated nuclei in a cell-free system might lead to well controlled genome reprogramming. Alternatively, it might be possible to microinject these factors into oocyte cytoplasm to reduce the importance of the phenomena that hamper genome reprogramming. A transfection of these factors or of their genes into somatic cells in culture might also reprogramme their genomes and transform them into embryo-like cells capable of generating a living animal.

2.2 Gene Therapy

Gene transfer can be achieved *in vivo* either in somatic cells or in germ cells. The process is called gene therapy in the first case and transgenesis in the second. The aims of these procedures are different, as are the tools used to carry them out.

The genotype of the organism is not modified in gene therapy, whereas transgenesis is essentially used to generate genetically modified organisms.

2.2.1 The goals of gene therapy

Gene transfer in the somatic cells of an animal or of a human may be carried out for quite different reasons.

The first experimental gene transfer took place many years ago. Indeed, vaccination using a living pathogen consists of transferring the genes of the pathogen with the assumption that the proteins coded by the genes will induce an immune response. When the pathogens are inactivated beforehand, only the preformed proteins will act as the vaccine.

Genetic engineering has made it possible to modify viral genomes in some cases to generate non-pathogenic and poorly viremic mutated viruses, which may be excellent vaccines.

Many experiments carried out over the last decade have shown that naked DNA containing functional genes introduced in an appropriate manner into somatic cells can direct the synthesis of the corresponding proteins, which then act as vaccines.

Synthetic short DNA (oligonucleotides) has been shown to interact with endogenous genes to inactivate or mutate them after systemic injections.

Introducing a gene into a tissue *in vivo* is another way to carry out transfection. Gene transfer in this case is more complicated than in cultured cells, but it has the theoretical advantage of evaluating the gene function and action in its natural context. In some cases, gene transfer in a tissue can be a substitute for transgenesis. It is simpler and faster than transgenesis. It may be a preliminary experiment that encourages researchers to generate transgenic animals to study the effect of the gene in a stably transformed organ.

Strictly speaking, the expression ‘gene therapy’ designates a technique aimed at correcting a genetic disease by introducing a functional gene

into an organ in order to compensate for the presence of an inactive endogenous gene. Gene transfer is indeed used for this purpose.

Gene transfer to somatic cells may also be a possible way to induce tumour cell death. This approach has proven to be effective in some cases.

2.2.2 The tools of gene therapy

The majors tools for transferring genes into somatic cells are the following: naked DNA, naked DNA associated with chemical transfectants, naked DNA associated with molecules capable of targeting DNA transfer to specific cells, synthetic oligonucleotides and viral vectors.

2.2.2.1 Naked DNA transfer

The fact that DNA injected into somatic cells can express the gene it contains and induce an immune response was discovered in a fortuitous manner. Induction of naked plasmids into mouse muscle was followed by the maintenance of the foreign DNA for several months without any degradation and by the formation of antibodies directed against the protein coded by the gene.

Naked DNA was transferred into various somatic cells by injection into organs or by the biolistic method, which consists of projecting small metallic bullets covered by DNA into cells.

Electroporation performed directly *in situ* proved to be highly effective for transferring foreign genes in some cases (Ferber, 2001).

The coinjection of a transposon harbouring a foreign gene and of a plasmid containing the corresponding transposase gene into blood led to the uptake by hepatic cells. The transposons were integrated into the hepatic cell genome and expressed the human factor VIII added to the transposon for several months (Ferber, 2001).

Injection of a huge volume of naked DNA into blood destabilized cell junctions and led to the enhanced uptake of the foreign DNA (Ferber, 2001).

2.2.2.2 Oligonucleotide transfer

Oligonucleotides containing about 20 bases can be synthesized in large amounts. These oligonucleotides can enter cells when added to

their culture medium or after systemic injections. The oligonucleotides contain nucleotides with a modified chemical structure that makes them more stable in blood and in cells (Dagle and Weeks, 2001).

Oligonucleotides having an antisense sequence to a given mRNA can form an RNA–DNA hybrid that is sensitive to RNAse H. The mRNA is specifically destroyed in this way.

The oligonucleotides may be directed towards the functional sites of an mRNA. In this way, they can prevent premRNA splicing, mRNA transfer from nucleus to cytoplasm, translation etc. (Dean, 2001).

A DNA region containing at least 20 uninterrupted pyrimidine nucleotides can form a stable triple helix with an oligonucleotide having an anticomplementary structure. These oligonucleotides, known as triplex forming oligonucleotides (TFOs), can block not only transcription of the targeted gene but also translation of the corresponding mRNA. Triple helices can be formed as well with an RNA. This RNA can be generated by a transgene (Section 2.3.11).

A tail of DNA added to a TFO can induce the mutation of the corresponding genome sequence. In this case, the TFO acts as the carrier, efficiently bringing the mutagene to its site of action (Giovannangeli and Hélène, 2000).

Synthetic oligonucleotides formed by a mixture of ribo- and deoxyribonucleotide sequences and thus called ribodeoxyribooligonucleotides (RDOs), can induce specific mutagenesis in a genome. To be efficient, the RDO must carry the mutation in a deoxyribonucleotide whereas the rest of the sequence is composed mainly of ribonucleotides. Moreover, the RDO must be a double-strand structure where the second strand has a DNA structure. The RDO must also be circularized to be stable (Alexeev *et al.*, 2000). This specific gene modification by RDO is known as chimeraplasty.

2.2.2.3 DNA transfer by chemical transfection

DNA uptake by cells is highly increased when it is associated with various chemical components (Section 1.4.5). Some of the transfectants used for cultured cells are currently being used *in vivo*. They show quite significant efficiency (Ferrari *et al.*, 1997), and one of their major advantages is their low toxicity and capacity to induce rejection.

2.2.2.4 Targeted naked DNA transfer

To avoid direct DNA injection into a given tissue or a tumour which is invasive and may be a difficult task, DNA may be targeted at specific cells by ligands.

Cells have many receptors and other molecules on their surface. Receptors are recognized by their specific ligands and, more generally, monoclonal antibodies may specifically bind to quite different molecules at the surface of the cells. This binding is often followed by an endocytosis of the complex.

Various methods are currently being used to associate a gene, usually in a plasmid, to a specific ligand. These methods are often based on the use of polycations (polylysine, histones, protamines), which form stable complexes with DNA. A covalent binding of the polycations with the ligand generates the final complex capable of targeting the gene into a specific cell type after systemic injection.

Several specific ligands have been successfully used. This is the case for asialyloproteins, in particular, which specifically recognized their receptor in hepatocytes and were able to target DNA uptake in liver. Monoclonal antibodies are a tool of choice for this purpose. This technique is known as antifection (Durbach *et al.*, 1999).

2.2.2.5 DNA transfer by viral vectors

Viruses have the natural capacity to transfer their genome into cells very efficiently. This fact encouraged researchers to use some viral genomes as vectors for foreign DNA. The principle of viral vectors is shown in Figure 1.5.

Several types of viral vector are being used or studied at this time (Ferber, 2001).

Retroviral vectors are currently being used since they have the advantage of integrating the foreign gene into the host genome. Retroviral vectors have intrinsic limitations. Most of them infect only dividing cells. The viral genome can reach chromatin only when the nuclear membrane is disrupted, implying a cell multiplication. Lentiviruses have the capacity to infect non-dividing cells (but not cells kept in G0). This property is due to the fact that this type of virus synthesizes a protein that targets the viral genome at the nucleus through the nuclear membrane. In the lentiviral vectors, this

protein is present in the infectious particles (Amado and Chen, 1999). The lentiviral vectors presently used generally contain the required regions of the HIV genome. Retroviral vectors can recombine easily and replicate in patients. Retroviral vectors have a limited capacity to harbour foreign DNA (no more than 7–8 kb). This may be insufficient for long genes or for constructs requiring long regulatory sequences for transcription. Interferences between the viral promoter located in the long terminal repeat (LTR) and the promoter of the foreign gene may lead to poor or ill controlled expression of the gene. Retroviruses have a relatively low capacity to infect cells *in vivo*. Infection must therefore be carried out in isolated cells from the patient, which are transferred back to the donor after infection and cell selection. Many retroviral sequences are rapidly inactivated after their integration into the host genome. The foreign gene silencing may be quite significantly reduced by adding insulators to the viral genome.

Adenoviral vectors have the advantage of being highly infectious *in vivo* as well as *in vitro* with a low cell specificity. They may harbour up to 10 kb of foreign DNA. These vectors are integrated at a very low frequency and the expression of the foreign gene is essentially transient. Adenoviral vectors also induce immunological reactions, which may severely limit their utilization.

Adeno-associated virus (AAV) vectors integrate themselves into the human genome, carrying about 10 kb of foreign DNA. They infect a broad spectrum of cells at a high rate of efficiency. The generation of the vectors requires the presence of an adenovirus, but they are well tolerated by patients.

The *Herpes simplex* virus (HSV) genome contains 150 kb and it may harbour long foreign DNA. This vector appears more particularly appropriate for infecting nerve cells *in vivo*.

Other viral vectors can be used or are under study. This is the case of vaccine vectors that can accept long foreign DNA and infect many cell types.

2.2.3 The applications of gene therapy

Vaccination with naked DNA is possible although it has a somewhat unpredictable efficiency rate. The methods of transferring DNA *in vivo* into somatic cells for this purpose have not yet been optimized. On the other hand, it has been repeatedly observed that immunization by DNA injection induces an immune cellular response with a low production

of antibodies. It is both interesting and surprising to observe that the GC-rich region of the plasmid carrying the foreign gene plays the role of a potent adjuvant. The injection of the antigen is required to boost the immune response and to induce the massive production of antibodies. It is also surprising to note that DNA applied to skin at the level of hair follicles induces a potent immunological response (Fan, Challah and Watanabe, 1999b). Additional experiments are required to validate this method. Although promising, vaccination by naked DNA is not yet a clinical practice. The efficiency of the method is still too unpredictable and the side effects of this means of vaccination have not yet been fully evaluated. One possible problem is that the foreign DNA might integrate itself into the host genome, alter gene expression and induce the formation of tumours.

Many synthetic oligonucleotides are currently being tested and some of them are used to inhibit the expression of endogenous genes involved in tumour development or of viral genes.

The use of TFOs (triple-helix forming oligonucleotides) is not yet a routine practice and additional studies are required to reach this goal.

Chimeraplasty is a very impressive method although it is not reproducible for a given gene and its efficiency seems highly dependent on the targeted gene (Albuquerque-Silva *et al.*, 2001).

Viral vectors are more and more frequently used by researchers for their high capacity to infect cells *in vitro* and *in vivo*. These tools may be used to study the biological effects of the numerous genes discovered by genome sequencing (Lorens *et al.*, 2001; Janson and During, 2001; Wade-Martins *et al.*, 2001).

The correction of genetic diseases by gene therapy has met with quite limited success so far. This is clearly due to an underestimation of the hurdles involved in efficiently transferring genes into organs but also to obtaining an appropriate expression of the genes.

The most impressive results obtained in 2000 were most certainly related to the therapy of the human form of severe combined immunodeficiency (Cavazzana-Calvo *et al.*, 2000). Children suffering from this disease have a non-functional gene coding for the γc subunit of the receptor cytokines IL 2, 4, 7, 9 and 15. The transfer of the normal gene into haematopoietic cells using a retroviral vector *in vitro* followed by a transplantation of the cells into bone marrow were clearly beneficial from a clinical point of view. The patients showed normal immune reactions for over one year. It must be kept in mind that this success was largely due to the fact that the transfer of a single gene repaired several receptors

and that its effect was strongly amplified by the resulting cell multiplication (Fischer, 2000).

It is also interesting to note that cells from haemophilic patients were transfected with a construct containing the gene for human factor VIII. When these cells were transplanted into the patients, they were able to secrete the coagulation factor in sufficient amounts to prevent bleeding for months.

The use of gene therapy for cancer treatment may seem easier than the correction of a genetic disease. Indeed, tumour cell death must by caused by the foreign gene and not by genetic information to repair a genome. Several killer genes are capable of inducing tumour regression but with unpredictable efficiency. The mechanisms involved in cell death require additional study (Finkel, 1999).

Gene therapy may also contribute to sensitizing tumours to chemicals or to converting prodrugs into active drugs at the tumour level. Gene therapy may also enhance the resistance of haematopoietic cells to the effects of drugs during chemotherapy (Encell, Landis and Loeb, 1999).

Gene therapy *in utero* would be quite useful for foetuses suffering from genetic diseases. Technical problems remain to be solved before this kind of therapy can be put into effect (Billings, 1999; Schneider and Cutelle, 1999).

Although slow, progress in gene therapy is real. The performance of the viral vectors has been more extensively evaluated and improved.

Adeno-associated viral vectors (AAVs) appear to be an increasingly reliable tool, providing efficient integration, no side effects and long term expression of the foreign genes. On the contrary, adenoviral viruses frequently induce acute reactions in patients. Vectors using lentivirus (HIV) and herpes virus (HSV1) genomes are also efficient for transferring genes, the first into many cell types and the latter into the nervous system. Results appear promising for the regeneration of brain tissues.

The non-viral vectors associated with transfectants are also being significantly improved (Brower, 2001).

2.3 Techniques of Animal Transgenesis

2.3.1 The aims and the concept of animal transgenesis

Transgenesis consists of introducing an exogenous DNA sequence into the genome of a pluricellular organism, which then becomes present in

most cells and is transmitted to progeny. The word, transgenesis, is therefore restricted to plants and animals. Yeast, bacterial and cultured cells harbouring a foreign DNA fragment are known as recombinant or transformed cells instead.

Transgenesis is therefore different from gene therapy. Indeed, in the latter case, the germ cells do not harbour the foreign DNA. The expression germinal gene therapy is also used. It designates a therapy that has not yet been attempted in humans and that implies that the foreign gene is contained in the germ cells and transmitted to progeny.

For historical reasons, the term GMO (genetically modified organism) is essentially used to designate transgenic plants cultured for animal and human nutrition. More logically, GMO should refer to all genetically modified living organisms, including microorganisms. The expressions GMP and GMA, which stand for genetically modified plants and animals, respectively, tend to be used.

The DNA fragments used to generate transgenic organisms are, in practice, almost always genes containing a sequence preceded by a promoter driving its expression in an RNA and generally in a protein. However, it is conceivable that the foreign DNA is devoid of the capacity to be transcribed and that its presence is desirable but not its transcript.

The gene transcript may be an RNA not translated into a protein. This is the case for antisense RNA, ribozymes and the genes transcribed by the RNA polymerases I and III.

It should also be mentioned that even if the foreign DNA has always been integrated into the genome of the transgenic organism so far, this is not necessarily the case. The foreign DNA used at this time cannot be maintained in the organism without being integrated into its genome. A free DNA fragment is rapidly eliminated during cell division since it does not have the capacity to replicate and to be transferred into daughter cells. However, it is theoretically possible to maintain a foreign DNA fragment as a minichromosome capable of self-replicating and being present in daughter cells. A number of viral genomes have this property naturally, as in the case of herpes viruses. Some chromosome fragments often found in tumour cells, known as minute chromosomes, contain elements for replication and transfer into daughter cells.

Generally speaking, the aim of transgenesis is to add foreign genetic information to a genome. It is also to suppress an endogenous gene. In some cases, the replacement of a functional gene by another functional gene is desired. The foreign gene may be a mutant of the endogenous gene or a completely different gene.

Gene addition may be performed to provide the organism with a new protein. It may result in the extinction of an endogenous gene as well. Gene addition may be used to study the action mechanism of a promoter in the whole organism. The association of reporter genes with promoters is generally the rule for this procedure.

Gene replacement is mainly used to inactivate a given gene. In practice, it consists of replacing the endogenous gene by an inactivated mutant. This approach is expected to give information about the biological function of the gene, as in the case of gene addition. Indeed, both gene addition and inactivation may induce modifications in transgenic animals, which may be observed or measured. Mutants of the replaced gene may provide information in a subtler manner.

Replacing a gene by another having a quite different function is more rarely achieved, but may be performed to introduce a marker or selection gene into a genome. The site of integration into the genome may be chosen for its potential capacity to express the foreign gene in a reliable manner.

In all cases, it is acknowledged that the integration of a foreign DNA fragment into a genome is mediated by the DNA repair mechanisms of the cell. The proteins involved in these mechanisms recognize abnormal DNA structures, which may be a mismatching of DNA strands, single-strand regions or sites in which the foreign DNA is associated with the host DNA.

When the foreign DNA has no common sequence with the host genome, the recognition between the two DNAs implies only short sequences, which are more or less homologous. This recognition may be sufficient to induce the repair mechanisms. The foreign DNA is then integrated according to a non-homologous recombination process (Figure 2.3). This event is considered to be relatively rare, occurring randomly at quite different sites of the genome, and has been corroborated by our observations.

When the foreign DNA shares a long sequence homologous to a region of the host genome, there is precise recognition of these sequences. The repair mechanisms induce a strictly homologous recombination, which, in practice, consists of replacing the targeted endogenous gene by the foreign DNA. If the latter is mutated, the endogenous gene is replaced by a mutated gene (Figure 2.4). The homologous recombination is, at best, 100 times less frequent than the heterologous recombination. This clearly derives from the fact that the number of sites for an illegitimate recognition is much greater than for the homologous recognition, which is generally unique in each haploid genome.

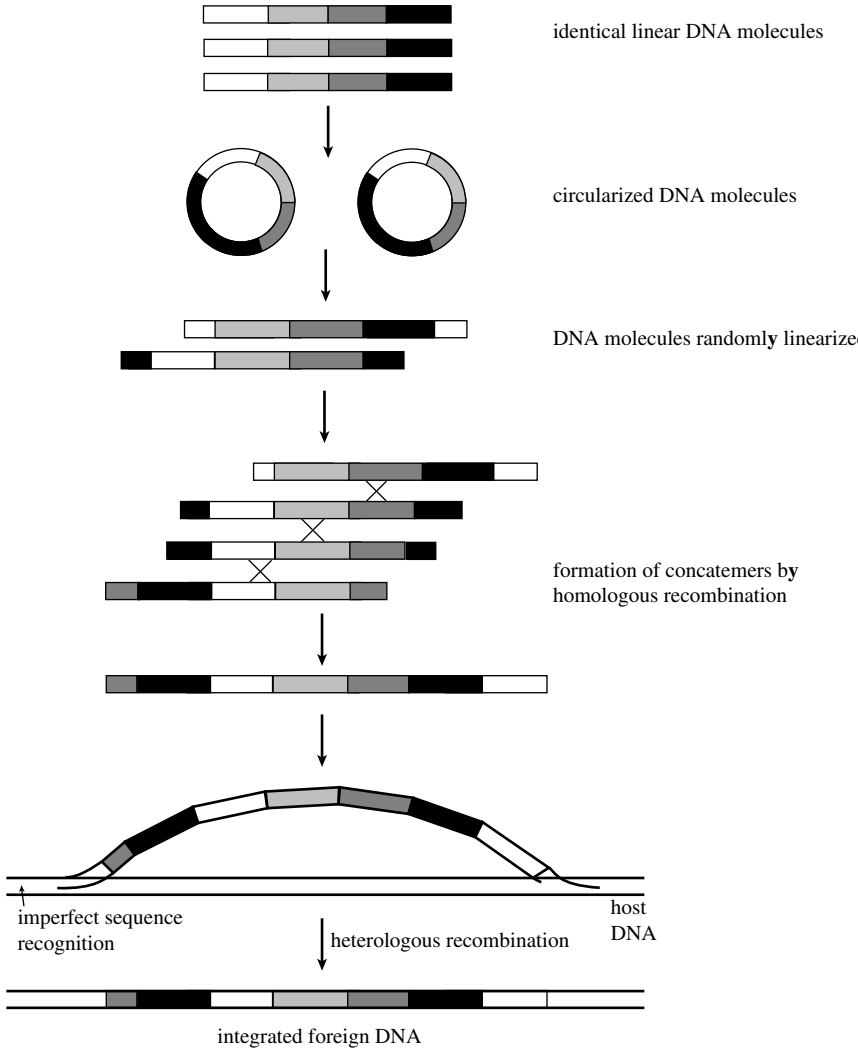


Figure 2.3 Integration mechanisms at a random site of foreign DNA injected into the nucleus of a cell. The injected DNA is randomly cleaved. The fragments are subjected to a homologous recombination process, which generates polymers (concatemers) of the injected gene organised in tandem. The ends of the concatemer are digested by DNAse, generating short single-strand regions, which recognize complementary sites in the genome. During DNA replication, the repair mechanisms integrate the foreign DNA

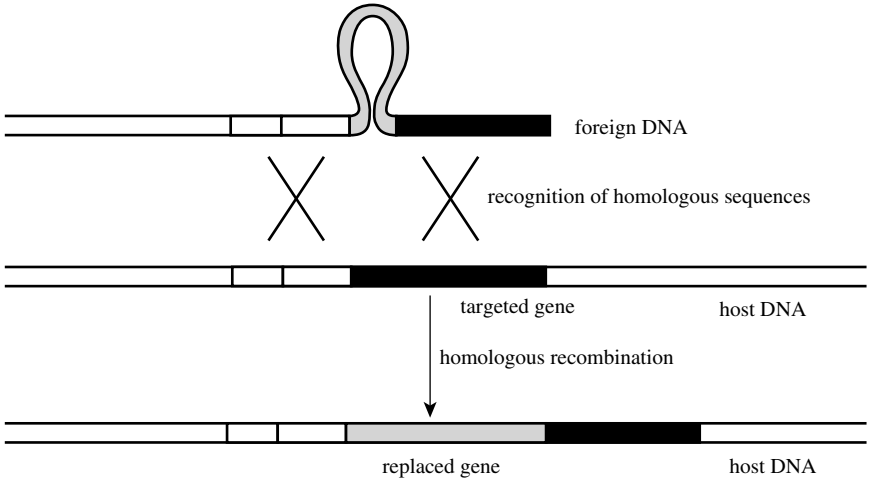


Figure 2.4 Mechanism of gene replacement by homologous recombination. The foreign DNA recognizes strictly homologous sequences in the genome. The repair mechanism of the cell induces specific replacement of the targeted genomic region by the foreign DNA fragments. A sequence located between the two homologous regions is integrated, whereas a sequence outside the homologous regions is eliminated

The foreign DNA must reach the cell nucleus to become integrated into its genome. The fate of the foreign DNA is not the same depending on whether it has been introduced into the cytoplasm or directly into the nucleus.

DNA transfected into cultured cells is generally in the form of a circular plasmid. The plasmid is cleaved by cytoplasmic DNAse at random sites. The major part of the DNA is destroyed in the cytoplasm. A small part migrates to the nucleus, where it can be transcribed. In this form, the foreign DNA is unstable and it is eliminated as cells divide. A small proportion of the foreign DNA becomes integrated into the genome. During the transfer from the cytoplasm to the nucleus, the foreign DNA fragments associate to form polymers called concatemers. In the cytoplasm, the covalent associations of the DNA fragments occur randomly, generating somewhat rearranged genes in a tandem or head-to-tail position.

When DNA is introduced directly into the nucleus, the fragments also form concatemers but through a homologous recombination process. The foreign DNA is cleaved randomly, generating overlapping frag-

ments, which recombine to form a concatemer in which the gene is well reconstituted. The different copies of the foreign DNA fragment are then organized, essentially in tandem form.

When different fragments of DNA are introduced simultaneously into a cell, they generate concatemers, statistically containing several copies of each fragment. The hybrid concatemer is integrated into the genome. Up to four different genes may thus be simultaneously transferred into a cell or an animal.

Hence, the foreign DNA is generally integrated in the form of a concatemer of about 100 kb, usually containing from one to ten copies of the original fragment. Interestingly enough, when large DNA fragments are transferred, the integrated material often contains a smaller number of copies, as though a length of about 100 kb was optimally integrated into the genome.

The foreign DNA microinjected into the nucleus or cytoplasm is linearized beforehand by cleaving the plasmid at a chosen site. This reduces the chance of a cleavage at random sites leading to the generation of concatemers containing truncated genes.

The DNA fragments used for transgenesis are linearized for other reasons. This procedure makes it possible to eliminate the plasmid sequences (namely those rich in GC) that may extinguish the transgenes. It is also well known that a circular DNA injected into the nucleus is integrated at a much lower frequency than a linear fragment.

The fate of the foreign DNA is therefore essentially the same when it is transferred to the cytoplasm by microinjection, transfection with chemical agents or electroporation. The injection of the foreign DNA into the nucleus is more laborious, but the integration rate is much higher as a result and the integrity of the foreign DNA is better preserved. Injection of DNA into the nucleus of the embryo is not always possible, particularly in non-mammalian species.

All these operations imply that (i) a foreign gene has been isolated and possibly modified by genetic engineering; (ii) the construct is present in the cells of the organism including the germ cells (or essentially the germ cells) to make its transmission to progeny possible.

These general remarks can essentially be applied to all species. It is easily conceivable that the techniques for gene transfer will differ according to species. Indeed, the problems are closely linked to animal reproduction. The following section describes the different techniques of gene transfer to generate transgenic animals.

2.3.2 Gene transfer into gametes

It seems logical to transfer genes into gametes to generate transgenic animals. Fertilization then brings the foreign DNA to all the cells of the embryo and the adult. Yet, the first transgenic animals were obtained by transferring DNA into embryos at the first-cell stage.

2.3.2.1 Gene transfer in sperm and precursors

In sperm Mature sperm has an inactive genome which does not replicate. Its DNA is covered by protamines, leaving very little access to foreign DNA in sperm genome. Foreign DNA thus has no chance of becoming integrated into sperm DNA.

Experiments performed about a decade ago, which proved poorly reproducible, showed that sperm incubated in the presence of DNA could carry it to the oocyte during fertilization, leading to the generation of transgenic mice. Biochemical studies showed that DNA binds readily to sperm, but it was not proven that this was followed by its integration. Sperm thus plays the role of a vehicle for the foreign DNA.

After several years, this method was extended to cows, pigs, sheep and medaka (Figure 2.5(A)). In all these cases and for unknown reasons, the experiments appeared to be poorly reproducible. Moreover, in most cases, the foreign DNA integrated into the host genome was highly rearranged. A study revealed that even after careful washing, sperm contain DNAse, which degrades the foreign DNA. DNAse is abundant in seminal plasma and variable amounts of the enzyme remain bound to sperm, explaining why the experiment was not reproducible and why foreign DNA was fragmented. It is also interesting to observe that exogenous DNA seems to induce a DNAse activity in isolated sperm. A possible interpretation of this phenomenon is that DNAse is a means to protect sperm from contamination by exogenous DNA on its way from the epididymis to the oocyte (Baccetti and Spadafora, 2000).

Attempts to enhance DNA trapping by sperm were made. Electroporation or transfection with chemical agents enhanced DNA uptake by sperm. This phenomenon was often accompanied by sperm capacitation, reducing its possibility of fertilizing oocytes.

In *Xenopus*, gene transfer in the embryo by microinjection leads to expression of the foreign gene and its maintenance for days and weeks but not to its integration. To circumvent this difficulty, researchers tried

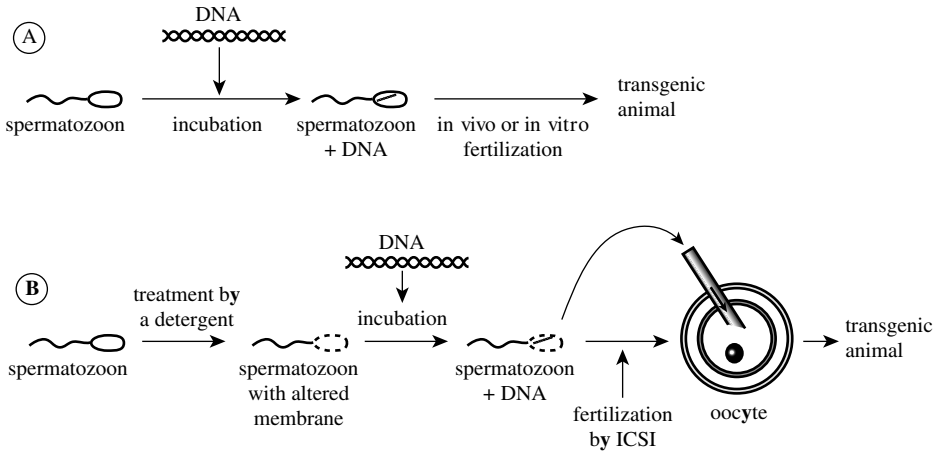


Figure 2.5 Gene transfer via sperm. (A) Washed sperm are incubated with DNA and used for fertilization *in vivo* or *in vitro*. This method is not particularly effective for generating transgenic animals and integrated genes are usually inactivated by DNA rearrangement (B) The sperm membrane is damaged by mild detergent. Foreign DNA freely enters sperm, which is used for *in vitro* fertilization using ICSI (intra-cytoplasmic sperm injection)

to use sperm as the DNA carrier. The protocol defined earlier for mammals proved ineffective in *Xenopus*. To enhance DNA uptake, sperm was treated with Triton, a mild apolar detergent. This led to a destabilization of the sperm membrane, leaving the foreign DNA free access to enter the sperm. The same result was obtained by freezing and thawing sperm. Chromatin was cleaved by a restriction enzyme recognizing a few sites in the sperm chromatin but not in foreign DNA. This induced the repair mechanism and enhanced the chance of the foreign DNA becoming integrated into the embryo genome. This method, known as restriction enzyme mediated integration (REMI), is a general way to promote cell transformation. After these treatments, the sperm lost its capacity to fertilize oocytes. The direct injection of sperm into oocytes, known as ICSI (intra-cytoplasmic sperm injection), which is used in humans for *in vitro* fertilization, was applied to *Xenopus*. This technique led to the generation of transgenic *Xenopus* with acceptable results (Figure 2.5(B)). Transgenesis can now be applied to this species, which is intensively used by biologists to study development, in particular (Marsh-Armstrong *et al*., 1999; Perny *et al*., 2001).

This technique, originally defined for *Xenopus*, proved to be successful in mice. Yet, the yield did not appear to be higher than with conventional

microinjection into pronuclei, which is a simpler method. The ISCI method was used quite successfully in mice to transfer long DNA fragments in BAC and MAC constructs (Perry *et al.*, 2001). The technique described for *Xenopus* might therefore be extended to different species, including large farm animals. The limitation currently arises from the difficulty of using ICSI in these species. It does not seem to have been applied to non-human primates, in which transgenesis is difficult using other means.

A very elegant method has been recently proposed (Qian *et al.*, 2001a, 2001b). DNA bound to a monoclonal antibody that recognizes a protein at the sperm surface was capable of forming a stable complex with sperm. This complex was used to fertilize mice *in vitro*, chickens by artificial insemination and pigs by injection into the uterine horn. In these three cases, about 30 per cent of the newborns were transgenic. The transgene was not rearranged. It was expressed and transmitted to progeny. It is interesting to observe that the same monoclonal antibody recognizes sperm in lower and higher vertebrates, including humans. This approach might be implemented to generate transgenic primates.

In sperm precursors *in vitro* It is known that mature sperm is produced from stem cells through different states of differentiation (Figure 2.6). Sperm stem cells can be isolated, cultured *in vitro* for a short period of time and transplanted into adoptive testis. The transplanted cells follow their differentiation programme, leading to functional sperm. The proportion of sperm resulting from the transplanted stem cells was greatly increased by treating the recipient males with bisulfan, a drug that blocks testis stem cell differentiation. This protocol has been successfully implemented to transfer genes in stem cells during culture. This was achieved using an efficient retroviral vector. The stem cells transplanted into bisulfan treated recipient males generated transgenic mice at a rate as high as 4 per cent (Nagano *et al.*, 2001). This method may be useful to study the biological effect of genes during sperm stem cell maturation and to generate transgenic animals. Most likely, the extrapolation to species larger than mice will be less successful. Indeed, it seems necessary to use more highly transformed cells to increase the chance of colonizing testis at a sufficient rate.

In sperm precursors *in vivo* DNA associated with transfectants can be injected into seminiferous tubules (Figure 2.7). Several groups succeeded in obtaining transgenic mice using sperm from the treated animals (Sato

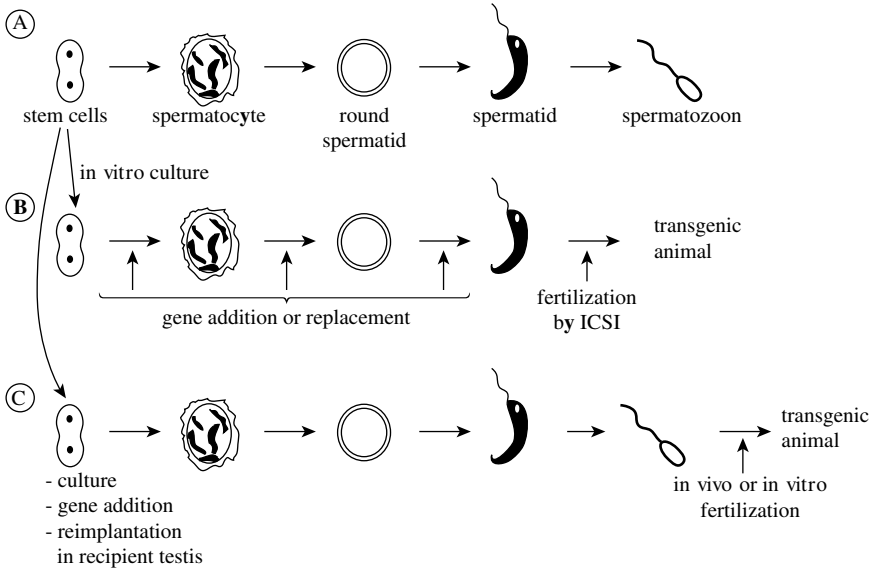


Figure 2.6 Use of sperm precursor cells for gene transfer.
(A) The different steps of stem cell differentiation into sperm
(B) The stem cells or partially differentiated cells are isolated, cultured and transfected by selected foreign DNA. Sperm harbouring the foreign DNA are used for fertilization by ICSI
(C) Stem cells are isolated, cultured under conditions preventing their differentiation, transfected by foreign DNA, selected and reintroduced into a recipient testis, where they differentiate. The resulting sperm are used to fertilize oocytes by conventional methods

et al., 2002a; see articles in *Mol. Reprod. Dev.*, 2000, 56). These encouraging results were moderately efficient and are not yet standardized. They are currently being improved but it is by no means certain that they could be extended to animals larger than mice. Indeed, the structure of seminiferous tubules may be different and the DNA injection difficult. On the other hand, the amount of DNA to be injected may be very large and difficult to manage.

Another point has to be taken into consideration. In mice, gene transfer into sperm precursors *in vivo* occurred independently in different cells, leading to the birth of animals in which different integration events occurred. This is an advantage since the different integrations may give somewhat distinct phenotypes. In this way, researchers can establish multiple lines of transgenic animals from a single experiment.

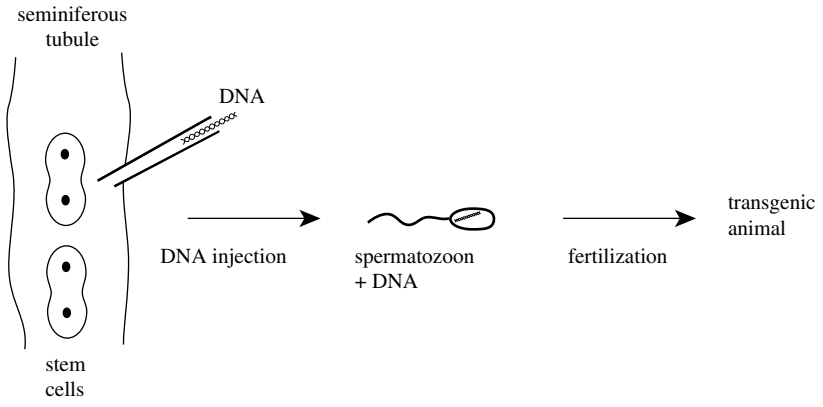


Figure 2.7 Direct gene transfer into sperm precursor cells. Naked DNA or DNA bound to a transfectant is injected directly into seminiferous tubules. The cells that have integrated the foreign DNA will generate transgenic animals after conventional fertilization

2.3.2.2 Gene transfer in oocytes

Microinjection into oocyte nuclei is feasible but did not lead to a sufficiently high birthrate of transgenic animals. DNA does not replicate in oocytes and foreign DNA has little chance of becoming integrated. The injected DNA could have been maintained in the nucleus until the first DNA replication in the embryo. This did not appear to be the case and microinjection was not retained as a means of generating transgenic animals.

Another method based on the use of retroviral vectors met with success in cows (Chan *et al.*, 1998) and in monkeys (Chan *et al.*, 2001). The retroviral vectors used were those routinely implemented in gene therapy, with a modification. Earlier works have shown that the conventional retroviral vectors containing an envelope from a retrovirus were not capable of infecting oocytes. It is believed that this is due to the absence of virus receptors at the surface of the oocyte. To circumvent this problem, the envelope from VSV (vesicular somatitis virus) has been used. This envelope is known to recognize membrane phospholipids in all cell types, allowing for efficient infections, irrespective of cell origin.

The zona pellucida, which surrounds the oocyte, is not permeable to molecular complexes as large as retroviruses. The retroviral particles

were therefore injected between the zona pellucida and the oocyte membrane. These injections were performed when the oocyte was in metaphase II. The nuclear membrane is absent at this stage, leaving the retroviral vector free access to chromatin.

This method led to the generation of transgenic cows, but the rate was not comparable to that achieved with the method based on animal cloning (Figure 2.11 below). The same method has been the only capable of obtaining transgenic primates so far.

This method is rather laborious with the classical limitations of the retroviral vectors: limited efficiency, limited space to introduce foreign genes, poor expression of the transgenes and the necessity of controlling the different steps to avoid dissemination of the vectors.

2.3.3 Gene transfer into embryos

Gene transfer into one-cell embryos is expected to allow transmission of the transgene to all the cells of the organism and to progeny. The various known transfection methods proved inefficient for transferring genes at an acceptable rate in embryos. Direct DNA microinjection was therefore retained.

2.3.3.1 DNA injection into pronuclei

This method has been consistently used in mice since 1980 and in several other mammals since then. About 500–5000 copies of the foreign DNA are injected into a pronucleus in 1–2 pl.

Depending on DNA purity and other uncontrolled factors, one to five transgenic mice can be routinely obtained from 100 injected embryos. This rate is halved in rabbits and rats and still lower in pigs, sheep, goats and cows. The rate is particularly low in cows. The reason for this difference between species is not known. It might reflect different activity of the DNA repair mechanisms involved in DNA integration.

This relatively low yield is due, to a large extent, to the death of embryos after microinjection. This phenomenon is mainly induced by the presence of DNA, which is mutagenic. The process of integration implies a somewhat intense alteration of endogenous DNA at the site of integration (Figure 2.3). In surviving embryos, the integration rate is 1–20 per cent, depending on the species and experiments.

The different steps of this protocol are depicted in Figure 2.8. A large number of embryos is obtained after superovulation. The microinjected embryos are transplanted into recipient females, hormonally prepared to develop embryos by mating with a vasectomized male or by a hormonal treatment.

Essentially, the same protocol has been successfully applied to other mammalian species. Overall success decreases from the mouse to the cow. This is due to the limited number of available embryos in ruminants and their cost but, also, to the low integration rate.

It should also be mentioned that in pigs and ruminants the embryos are opaque due to the presence of lipids. A centrifugation of the embryos for a few minutes at 10 000 *g* transfers the lipids to one part of the embryos. The pronuclei become visible without altering the embryos. To circumvent these difficulties, other protocols have been defined.

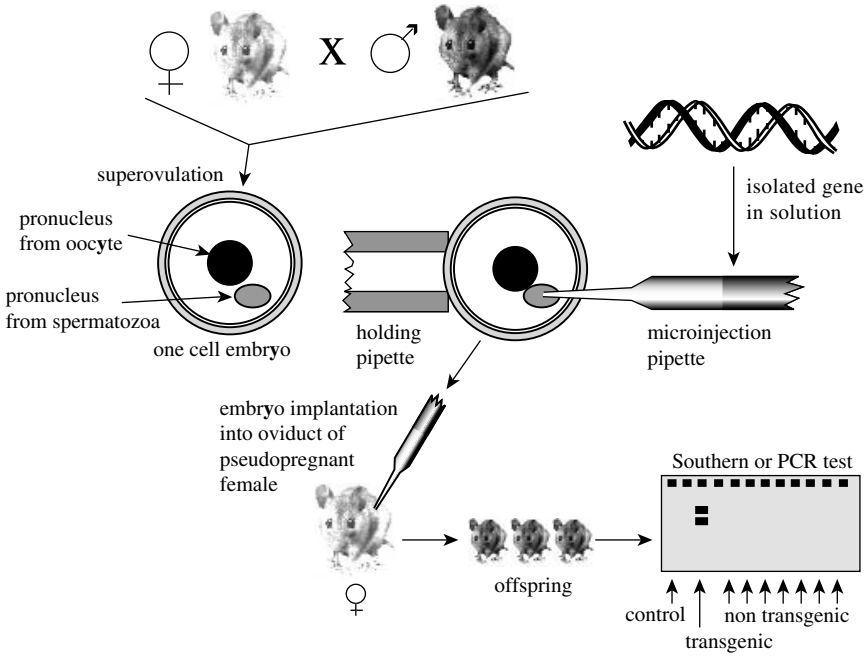


Figure 2.8 Protocol to generate transgenic mice by DNA microinjection. Embryos are obtained after a superovulation and *in vivo* fertilization. The isolated gene is microinjected into one of the pronuclei of a one-cell embryo. Embryos are reimplanted into pseudopregnant recipient females, which have been mated with vasectomized males

One of them, depicted in Figure 2.9, is still based on microinjection. The modifications take place before and after microinjection. Embryos at the one-cell stage are obtained from ovaries collected in slaughterhouses. Oocytes in the process of maturation are isolated and cultured in the presence of hormones to achieve maturation. Fertilization is then performed *in vitro* and microinjection can take place. In this way, up to 300 embryos can be obtained per day at a low cost.

The embryos can start their development *in vitro* until the blastocyst stage. Those that cannot survive until microinjection disappear during the culture. This method makes it possible to avoid the use of donor females and requires a limited number of recipients. It has been used successfully, although it remains laborious (Krimpenfort *et al.*, 1991).

A marker gene whose product is visible under the microscope using a non-invasive test may be injected with the gene of interest. Both genes cointegrate in 70–80 per cent of cases. The GFP gene is an appropriate marker for this purpose. The embryos that appear green under UV light

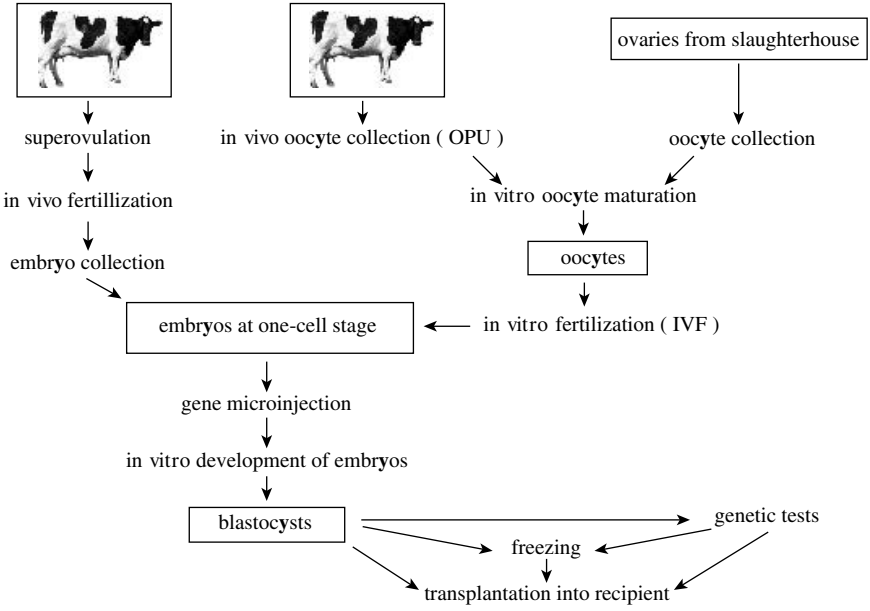


Figure 2.9 Protocol to generate transgenic cows by DNA microinjection. Embryos are obtained from oocytes collected in slaughterhouses, matured and fertilized *in vitro*. The microinjected embryos are cultured until the blastocyst stage. The non-viable embryos are spontaneously eliminated. The transgenic embryos may be selected using a cointegrated transgene and reimplanted into recipient females

harbour the GFP gene and, potentially, the gene in question. This method is simple but has two intrinsic drawbacks. Embryos harbouring only the gene of interest are discarded and some of the embryos have only the GFP gene. On the other hand, the GFP gene remains in the genome, where its presence is no longer required. This marker gene can be theoretically eliminated at subsequent generations if it has been bordered by LoxP sequences (see Section 2.4.9).

Although improved and satisfactory, this method was rapidly abandoned when gene transfer appeared possible by the nuclear transfer technique (see next section).

2.3.3.2 DNA microinjection into cytoplasm

In species other than mammals, pronuclei are not visible and microinjection may only be performed in the cytoplasm of embryos. This is particularly the case for lower vertebrates and invertebrates. Most of them are oviparous. This implies that the development of the embryo occurs outside the mother. A shell is required to protect the embryo. Stored nutrients must be present to allow autonomous development of the embryo until hatching. This hampers microinjection. A small opening must be made in the shell to introduce a micropipette into the cell.

In fish embryos, namely of salmonids, about 10–30 nl of DNA solution containing 20 million gene copies must be injected to obtain the optimal generation rate of transgenic fish. Up to 50 per cent of the animals may be transgenic with a low mortality rate.

The efficiency of this method is quite variable, depending on the species. It cannot be used for the generation of transgenic chickens, *Xenopus* or medaka.

The fate of DNA is different in mammals and lower vertebrate embryos. In chickens and *Xenopus*, the foreign DNA is maintained during early development without being integrated. In salmonids, the injected DNA is intensively replicated irrespective of its sequences. This increases the chance of the foreign DNA being integrated in the host genome. This may contribute to the high generation rate of transgenic salmonids. The drawback is that foreign DNA remains abundant for several days and may be integrated at different stages of embryo development. This leads to the generation of heavily mosaic transgenic animals. Independent and late integrations are responsible for this phenomenon. The founder transgenic animals may be very heterogeneous.

They are all different but the organs of each animal contain cells in which different integrations may have occurred. This is also true in gametes. Homogenous animals are of course obtained at the F1 generation. This complicates the task of experimenters, especially for species having a long reproduction cycle, such as salmonids.

A mosaicism is also regularly encountered in transgenic mammals but to a much lower degree. Up to 30–40 per cent of the transgenic founders in mice (and seemingly in other mammals) are mosaic. The transmission rate of the transgene to the F1 generation suggests that gene integration often occurs at the two-cell stage. A small percentage of the founders are too highly mosaic and do not transmit their transgene to progeny.

The efficiency of gene integration may be enhanced by different tools (see Section 2.3.5).

2.3.4 Gene transfer via cells

To avoid the laborious microinjection process, it may be advantageous to transfer genes into cells and to use them to generate embryos. This approach is required when gene replacement by homologous recombination is needed. Indeed, this event is rare and requires the selection of cells in which homologous recombination has occurred. This implies long cultures to establish cell clones. These cells must have retained their capacity to participate in the development of an embryo. Two different possibilities are presently offered to researchers.

2.3.4.1 The use of pluripotent cells and the generation of chimaerae

Pluripotent cells can generate all the organs of an organism (Figure 2.1). Freshly collected pluripotent cells introduced into an early embryo at the morula or the blastocyst stage can participate in the development of the embryo, giving rise to chimaeric animals. It should be recalled that a chimaera is formed by mixed cells from different animals. Each cell of a chimaera, including the gametes, originates from the donor or the recipient embryo. This is fundamentally different from a hybrid, in which the genome from two organisms is in the same cells after the fertilization process.

Lines of pluripotent cells have been established in mice. The lines derived from early embryos are called ES cells (embryonic stem cells).

The pluripotent cells obtained from foetal gonads are the EG cells (embryonic germinal cells). The first pluripotent cell lines were derived from a teratocarcinoma and were thus called EC (embryonic carcinoma cells). The EC cells can participate in the development of chimaeric embryos but not in the formation of gametes.

Only a small number of ES cell lines are available. They all are of murine origin and obtained from one of two lines of mice. Pluripotency is, by essence, a transient state. The programme of the pluripotent cells is to differentiate. Pluripotent cells have a high capacity to multiply *in vitro* but they spontaneously differentiate to become multipotent cells. The somewhat differentiated pluripotent cells can participate in the development of an embryo to some degree but cannot become gametes.

The ES cells must therefore be maintained in the pluripotent state throughout the entire culture to give generate chimaeric animals capable of transferring their genome to progeny.

The pluripotency of the ES cells is maintained by adding appropriate factors to the culture medium. Numerous attempts to obtain ES cell lines from other mouse lines and from other species have systematically failed. At best, non-germinal chimaeric animals are obtained. It is not clear why ES cell lines can be derived only from two mouse lines. Whatever happens, these lines appear to be the exception and the fact that pluripotent cells cannot be maintained in culture is increasingly considered to be the normal situation. Recent studies, however, have led us to reconsider this point of view. Pluripotent cells capable of transmitting their genes to progeny have been described in chickens and medaka. It is still too early to assume that these cells are equivalent to the mouse ES cell lines.

Foreign DNA can be added to ES cells and clones harbouring the foreign gene can be established using a selection gene. These cells can be used to generate transgenic chimaeric mice. These animals are mosaic for the transgene.

This method is laborious and much less efficient than microinjection for adding genes. It is therefore used only to replace genes by homologous recombination (Figure 2.10).

2.3.4.2 The use of differentiated cells and the generation of cloned animals

Foreign DNA can be added to differentiated cells using transfection methods. Cells harbouring the gene of interest can be cloned by a

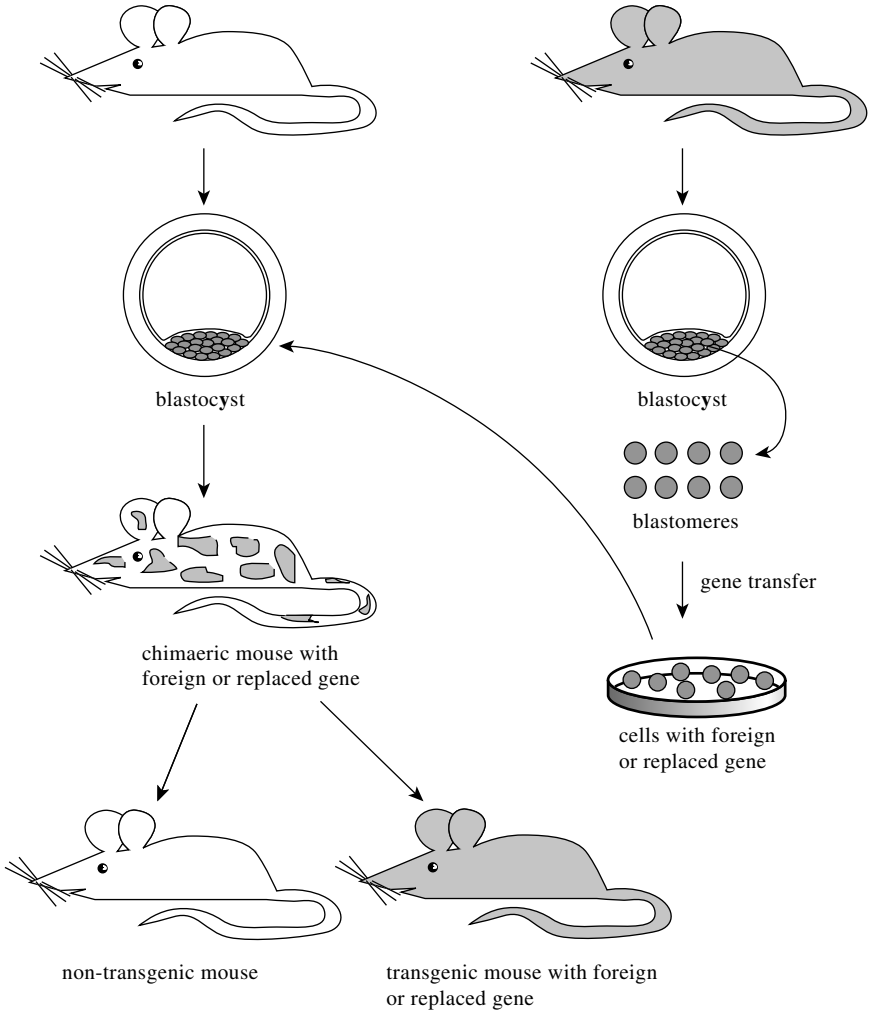
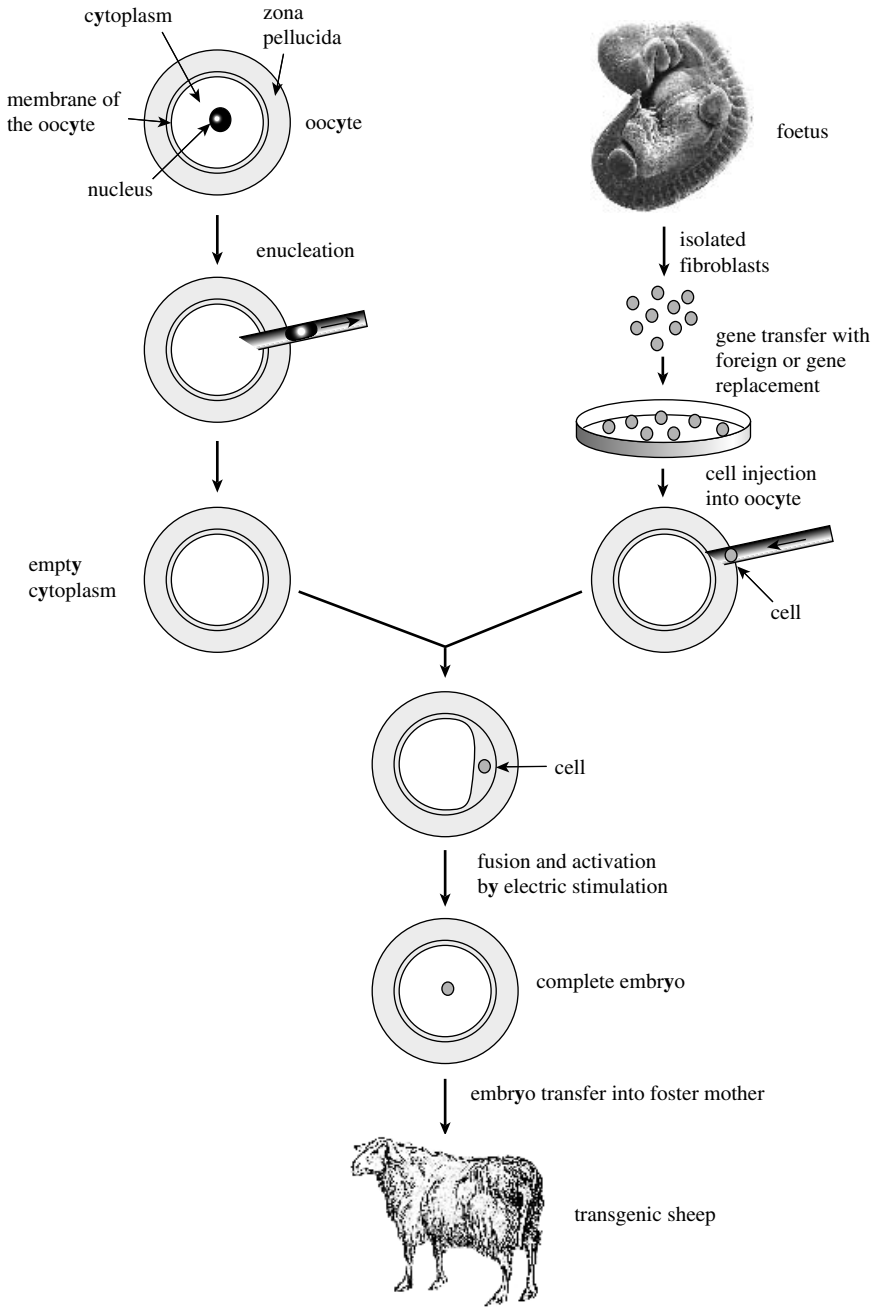


Figure 2.10 Gene replacement using chimaeric animals. L.M. Houdebine *Medecine/Sciences* (2000) 16: 1017–1029. © John Libbey Eurotext. Lines of pluripotent cells known as embryonic stem cells (ES cells) are transfected with a vector allowing homologous recombination with a targeted region of the genome. The genetically modified ES cells are transferred into a normal embryo at the same development stage. The resulting chimaeric animals may transmit the mutation to offspring

selection gene. These cells can then be the source of nuclei for the generation of cloned animals, which will also be transgenic (Figure 2.11).

The birth of Dolly was soon followed by that of a cloned transgenic sheep named Polly (Schnieke et al., 1997).


cytoplasm
zona pellucida
membrane of the oocyte
oocyte
nucleus
enucleation
foetus
isolated fibroblasts
gene transfer with foreign or gene replacement
cell injection into oocyte
empty cytoplasm
cell
cell
fusion and activation by electric stimulation
complete embryo
embryo transfer into foster mother
transgenic sheep

The advantages of this method of adding genes are multiple. About two to five times fewer sheep are needed to generate transgenic sheep than by microinjection. The integrated gene may be examined in cells before nuclear transfer. Cells in which the foreign gene is rearranged or has too many copies may be discarded. The sex and, more generally, the genotype of the nuclear donors may be chosen. The founder animals are never mosaic for the transgene. Several animals having the same genotype, including the same transgene, can be generated simultaneously. Although cloning is a laborious technique, it offers some flexibility to the experimenters. The nuclear donor cells can be kept frozen and used at the most appropriate moment to generate cloned transgenic animals.

Gene replacement was achieved in sheep (McCreath *et al.*, 2000), mice (Rideout *et al.*, 2000) and pigs (Lai *et al.*, 2002; Butler, 2002). This method is very laborious and still poorly controlled. A recent study showed that homologous recombination of two genes could be obtained in sheep cells but that this was followed by the death of the newborn animals obtained by cloning (Denning *et al.*, 2001). This failure may be attributed to the culture of the cells, which is required to select those in which the homologous recombination has occurred. The culture conditions modify the physiology of the cells, which, for unknown reasons, become less capable of generating living cloned animals. A better understanding of these phenomena is necessary before gene replacement in large animals can be considered as a truly viable method.

This is also true for mice. Although gene replacement by the chimaeric approach is laborious, it remains simpler than the cloning method.

2.3.5 Vectors for gene addition

Most of the vectors currently used to generate transgenic animals by gene addition are constructed to be integrated into the host genome. Various methods are being used or studied to enhance the frequency of foreign gene integration or to maintain them as independent minichromosomes.

Figure 2.11 Gene addition and replacement via cloning. L.M. Houdebine *Medecine/Sciences* (2000) 16: 1017–1029. © John Libbey Eurotext. Foetal cells are transfected, allowing gene addition or replacement. The selected cells are used as a source of nuclei to generate transgenic cloned animals

2.3.5.1 Minimum linear vectors

In the great majority of cases, researchers use genome fragments containing one or two genes or prepare functional gene constructs from various elements. The fragments of the vectors containing the transcribed and the regulatory regions are separated from the plasmid beforehand. Indeed, circular vectors integrate at a much lower frequency than linear DNA fragments and plasmid sequences often extinguish the associated transgenes. This is true for the different vectors: plasmid, cosmid, phage, BAC and YAC. However, a few studies have shown that circular BAC vectors integrate as efficiently as their linear counterpart. On the other hand, vectors containing long genomic DNA fragments are often less sensitive to the silencing effect of the prokaryotic sequences. This is most likely due to the presence of insulators in the long genomic fragments or to a simple distance effect.

The DNA fragments containing no particular sequences integrate at a relatively low rate. For unknown reasons, some DNA inserts give a greater number of transgenic animals than others. This may arise from the presence of sequences in the insert that recognize frequent genomic sequences (Figure 2.3). Alternatively, some inserts may contain sequences favouring their replication and their maintenance in the embryos, enhancing the occurrence of an integration.

2.3.5.2 Vectors containing repeated sequences

The mechanism of integration depicted in Figure 2.3 implies recognition between sequences of the insert and of the genome. The integration frequency should be enhanced by the presence at both ends of the insert of sequences that are highly repeated in the host genome, even if they are more or less degenerated. Several experiments have shown that this is indeed the case. In cows, a sequence abundantly present in centromeres and added to inserts enhanced the frequency of integration. In this particular case, the transgenes remained silent. This must be due to the fact that centromeres, which are non-transcribed regions of the genome, also extinguish transgenes.

A similar approach has been taken in mice, using Alu sequences as repeated elements. Alu sequences, which contain 200–300 nucleotides, are abundant in mammalian genomes and particularly in the vicinity of or within transcribed regions. Some of the Alu sequences are transcribed

by RNA polymerase III, giving rise to RNA whose function is not clear and could be non-existent. Experiments conducted by a group of researchers showed that the integration frequency was enhanced for inserts containing Alu sequences. This conclusion was not confirmed by another group, suggesting that the effect might be weak or dependent on the insert sequence.

2.3.5.3 Transposon vectors

Transposons are sequences generally no longer than 2 kb, many copies of which are present in genomes at apparently random sites. Transposon sequences are transcribed into RNAs, which are retrotranscribed into double-strand DNAs, which integrate into the genome with high efficiency. The integration is managed by a transposase coded by the transposon and by inverted repeated sequences (ITRs), which are present at both ends of the transposon (Figure 2.12). This mechanism allows the transposon to rapidly spread and invade a genome, inducing gene inactivation in some cases. Transposon invasion is limited by cellular mechanisms that inactivate the transcription of the transposon.

Transposons are potential vectors for integrating foreign genes into a genome. In order to do this, a large part of the transcribed region of the transposon is deleted. This generates space for a foreign gene and prevents transposons from autonomous and uncontrolled spreading in the genome.

The recombinant DNA containing the foreign gene has no particular capacity to become integrated in the genome by itself. The presence of the transposase is required for this purpose. The coinjection of the transposon containing the foreign gene and of a circular plasmid capable of expressing the transposase gene allows the transposon to be integrated with significant efficiency, in about 1–5 per cent of the embryos (Figure 2.12).

This protocol, originally defined for *Drosophila* using the transposon P, has been extensively used to generate transgenics (Kayser, 1997). Its extension to other species was put on hold until better knowledge of transposons capable of working in different species was developed. The transposon mariner proved effective in medaka, in chickens and in mammalian cells. Various modifications of this transposon might make it possible to obtain an efficient and safe vector for gene therapy (Hackett *et al.*, 2001).

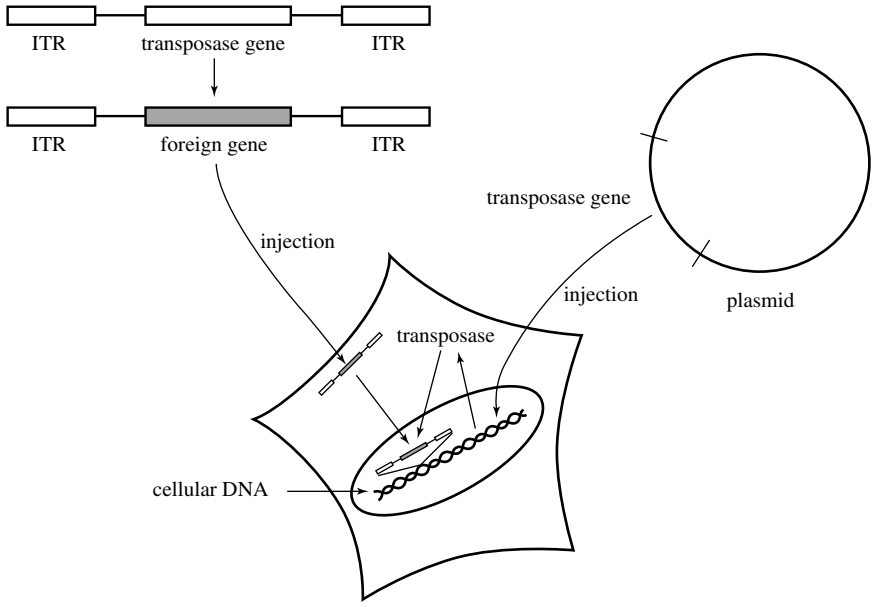


Figure 2.12 Use of transposons as vectors to transfer foreign DNA. The transposon gene is replaced by the gene of interest. The recombinant transposon is injected into the cells. A vector directing the synthesis of integrase is coinjected. The transposon is integrated using its ITR sequences

Other vectors are used to generate transgenic insects such as *Aedes aegypti* or silkworm (Tamura *et al.*, 1999). The transposon called sleeping beauty has recently been adapted to generate transgenic mice (Dupuy *et al.*, 2002).

In all cases, the transposon and the plasmid coding for the transposase must be injected into the embryo cytoplasm under conditions dependent on the species. In this respect, chickens are different from most other species. Indeed, injection can be performed in one-cell embryos, which cannot be reintroduced into foster mothers as is the case for mammals. The injected embryos must be introduced into the yolk of an egg that does not already harbour an embryo. After several weeks of incubation under well controlled conditions, transgenic chickens were born with an acceptable rate of success (Shermann *et al.*, 1998).

Transposon vectors thus allowed the generation of transgenic animals in species in which conventional DNA microinjection was ineffective. This tool is also considered to be safe. The mariner transposon vector, even when it lacks its transposase gene, can replicate and integrate into

the host genome at a low rate. This seems to be due to the presence of an endogenous transposase in the host. This may limit the use of this transposon in some cases. In contrast, the piggy BAC transposon used to generate transgenic silkworms remains perfectly stable after several generations.

Transposons can harbour only foreign DNA fragments of limited length. This is clearly a limit if sophisticated constructs are to be used to express the foreign gene. Moreover, the cellular mechanisms that extinguish the transposons may inhibit expression of the transgenes in some cases.

2.3.5.4 Retroviral vectors

This kind of vector already described in Section 2.2 are essentially those used for gene therapy (Figure 2.13).

Various retroviral vectors have been designed to generate transgenic animals and particularly chickens (Ronfort, Legras and Verdier, 1997). These vectors contain envelopes recognizing embryonic chicken cells. These vectors can be injected into newly laid eggs. At this stage, cells are still pluripotent and may participate in the generation of gametes, leading to the transfer of the transgene to progeny. This approach proved highly inefficient. Indeed, the embryo at this stage contains about 60 000 cells. Only a small proportion of these cells has a chance of being infected by the retroviral vectors. The resulting chickens are highly mosaic at the transgene level, leaving very little chance of transmitting their transgene to progeny. An alternative proved to be more efficient. Vector injections were performed at stage 16 of embryo development in the vicinity of primordial germ cells. These cells were preferentially infected, giving the animals an acceptable chance of transmitting their transgene to progeny (Figure 2.13).

Retroviral vectors have also been implemented to transfer genes into cow and monkey oocytes (Chan *et al.*, 1998, 2001). In order to do this, two particular conditions were required. The envelope of the vector was the G protein of VSV (vesicular somatitis virus), known to recognize membrane phospholipids and thus allowing infection in a wide variety of cell types. The viral particles were injected between the zona pellucida and the oocyte membrane at a time when the nuclear membrane was absent, giving the viral genome the best possible chance of reaching the host genome (Figure 2.13).

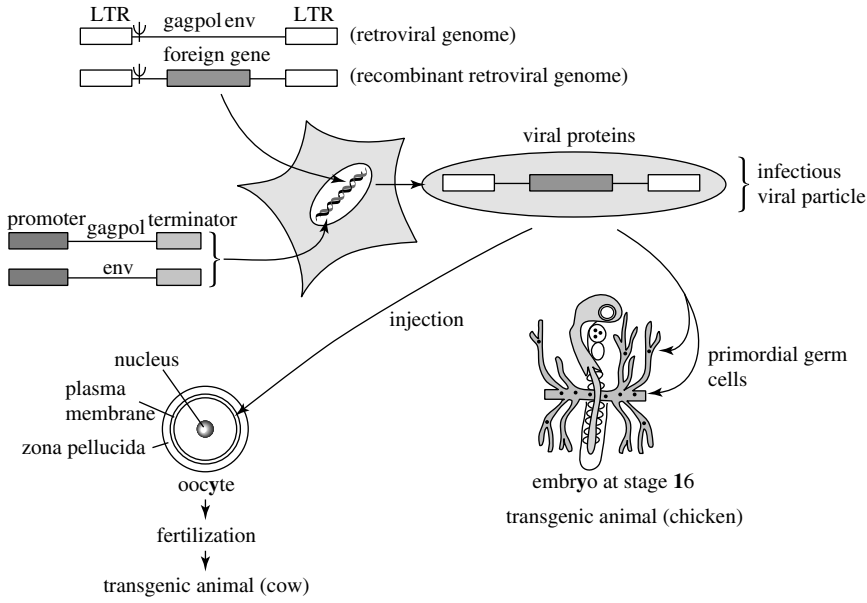


Figure 2.13 Use of retroviral vectors to generate transgenic animals. The viral genome harbouring the foreign gene is introduced into transcomplementing cells, which synthesize the missing viral proteins. The viral particles are used to infect mammal oocytes (cow, monkey), chicken primordial germ cells or mouse one-cell embryos

Lentiviruses are a subpopulation of retroviruses, which have the capacity to integrate into the genome of quiescent cells. This is due to the presence of a signal in one of the viral proteins, which targets the genome of the virus at the nuclear compartment. The AIDS virus (HIV) belongs to the category of lentiviruses. The lentiviral vectors are increasingly studied for their potential use in gene therapy (see Section 2.2).

A recent study indicated that lentiviral vectors injected into one-cell embryos or incubated with embryos from which zona pellucida was withdrawn were particularly efficient for transferring genes into mice (Lois *et al.*, 2002). The vectors used are sophisticated. Their genome originates from a lentivirus that gives the recombined particles the capacity to infect non-dividing or dividing cells. The vectors contain modified LTR, leading to a self-inactivation of the integrated viral genome. This considerably reduces the possibility of a recombination generating an infectious virus carrying the foreign gene in an uncontrolled manner. These types of vector are devoid of their enhancers. This makes it

possible for an endogenous promoter driving the gene of interest to be active without being subjected to the effect of the LTR enhancer. The vectors also contain a post-transcription regulatory element which favours gene expression, and a fragment of HIV virus to allow the viral genome to enter nuclei in non-dividing cells and to integrate into host DNA.

The vectors were able to drive the expression of the EGFP gene in mice, giving a green fluorescence in all cell types or only in muscle when promoters active in all cell types or only in muscle were used, depending on the case.

It should be noted that about 80 per cent of the mice born were transgenic. About 90 per cent of the mice expressed their transgenes without any silencing in several generations. The VSV protein was used as an envelope to allow efficient infection of the embryonic cell.

The authors of this elegant study examined the advantages and the limits of this tool. The major advantages are obviously the high rate of efficiency and the fact that infections can be performed by incubation of embryos devoid of their zona pellucida. A relatively simple manipulation of the embryos is thus sufficient. The method should be extended to other mammals with no particular problem. It would be particularly advantageous in rats, where the microinjection method is difficult to use.

The limits of the tool are multiple as well. Large quantities of viral particles must be prepared with an appropriate biosafety control. This step has become increasingly standardized as a result of the increasing use of these types of vector for gene therapy. DNA fragments not exceeding 7–9 kb can be introduced into the vector. The expression of the reporter gene is rather low in this case. On the other hand, independent integrations of the vectors occur in the same embryo, leading to mosaic transgenic animals. This may somewhat complicate the interpretation of the data.

Clearly, this tool appears able to replace microinjection in a number of cases but not in all of them.

2.3.5.5 Episomal vectors

The advantage of the integration process is that the foreign DNA is stably transmitted to progeny. The major drawback is that the integration events are rare. Another problem, which is discussed in Section 2.3.11, is the fact that the integrated gene is subjected to the action of its

chromatin environment, which frequently alters the expression of the transgene. On the other hand, the number of integrated copies is limited, generally not exceeding 10, and in a number of cases it seems that not all the copies are actively transcribed.

The use of episomal vectors is expected to eliminate most of these drawbacks. Autonomous secondary genomes or episomal genomes are found in a number of living organisms. This is the case for bacteria, which harbour a great variety of plasmids. Chromosome fragments resulting from a partial degradation of a chromosome are sometimes found in cells. This is particularly the case in some tumour cells. These chromosome fragments, known as minute chromosomes, are autonomously replicated and transmitted to daughter cells. Different classes of viruses have genomes that replicate as autonomous entities and are transmitted to daughter cells. This is namely the case for the viruses of the herpes family.

An episomal genome must contain several elements to be stable. Episomal genomes are either circular or linear. When they are circular, their minimum constituents are an origin of replication in which DNA synthesis is initiated as well as a system that makes it possible to dispatch the replicated genome to daughter cells. The linear episomal genomes must contain telomeres at both ends. These structures are present at the ends of each normal chromosome. They are permanently degraded by exonucleases and restored by a specific enzymatic complex. Telomeres protect chromosomes from degradation by exonucleases. When cells become older, telomere synthesis becomes less efficient, leading to a degradation of chromosomes and to cell death.

Different types of episomal vector can be designed to generate transgenic animals. The vectors can be relatively small, containing only known elements for replication and transmission to daughter cells. The other possible approach consists of using chromosome fragments containing all the natural elements to transfer genes to progeny.

The better-known and most frequently used episomal vectors are plasmids, cosmids, phages, BACs and YACs. None of the vectors can be used in higher eukaryotes since their elements are not active in animal cells. These vectors are essentially used to engineer DNA and generate recombined DNA to be studied and transferred into animal cells.

The bacterial vectors may contain only an origin of replication. The number of copies generated after DNA replication is sufficient to allow a statistical transmission of the vector to daughter cells.

In yeast, linear vectors called YACs have been designed (Figure 2.14). They contain telomeres, an origin of replication (ARS 1), a centromere (CEN 4), a selection gene (Ura 3) and cloning sites. This vector is relatively compact and easy to manipulate. This is due to the fact that the origin of replication and the centromere are short. The origin of replication of *Saccharomyces cerevisiae* is concentrated in a short genomic region. This is not the case for *Saccharomyces pombe*, in which the origin of replication is spread over several thousand bases. The YAC vectors can harbour DNA fragments as long as 1000 kb and they can be engineered in yeast by homologous recombination. The major drawback of YAC vectors is their instability (Giraldo and Montoliu, 2001).

BAC vectors have been designed to replicate in *E. coli* and are present in only one copy per cell. They can harbour up to 250 kb of DNA. They can be engineered in bacteria by homologous recombination. This allows both deletion and introduction of foreign DNA fragments in the vectors. BAC vectors are currently the best tool for preparing constructs containing long DNA fragments to be used to generate transgenic animals (Giraldo and Montoliu, 2001).

In animal cells, it remains impossible to generate compact linear vectors containing the natural elements found in chromosomes. The origins of replication in animal genomes are not well defined regions. Some of them have been characterized and used to replicate DNA in culture cells and in transgenic mice. It seems as if genomic DNA fragments containing several thousand base pairs are always required to initiate a DNA replication. A possible approach is to clone genomic fragments into vectors and to select those having the highest capacity to replicate in animal cells. This experiment revealed that at least one

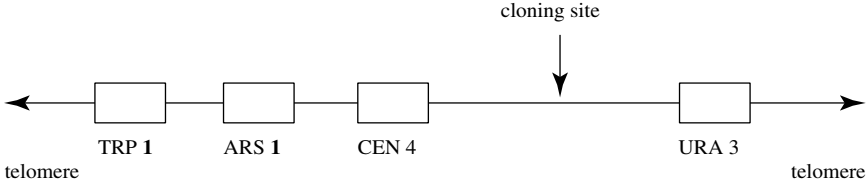


Figure 2.14 Structure of a YAC vector (yeast artificial chromosome). The vector contains an origin of DNA replication (ARS 1), a centromere (CEN 4), which makes it possible to distribute of the replicated vector in daughter cells, telomeres, which protect DNA from exonucleases, two selection genes (TRP 1 and UR 3) and a cloning site in which up to 1000 kb may be introduced

origin of replication is present in an animal DNA fragment of about 40 kb (Kelleher *et al.*, 1998).

The origins of replication are not adequate to allow a circular vector to be efficiently transmitted to daughter cells and animal progeny. A study carried out several years ago showed that a vector containing a mixture of animal origins of replication in a plasmid was highly efficient for generating transgenic mice. This vector could be transmitted from mice to bacteria, from bacteria to pig embryos and back again to bacteria (Attal *et al.*, 1997). However, this vector was not stable when a gene was inserted into it. Moreover, it was not maintained throughout the entire life of the mouse and not transmitted to progeny. This vector thus appeared to contain an origin of replication that was sufficient for a statistical dispatching in rapidly dividing cells but not in slowly dividing cells. This was attributed to the fact the vector did not contain any elements playing the role of a centomere.

Chromosomes in eukaryotes contain repeated sequences in their central section which bind to the cytoskeleton during mitosis, allowing chromosomes to be distributed into daughter cells in an appropriate manner. Centromeric regions have been added to some episomal vectors. Very long centromeric fragments must be used and they appear to be species specific. It is therefore not possible to use them to design compact vectors.

A certain number of viruses have a circular genome, which replicates at a high rate in animal cells and is thus statically transmitted to daughter cells. This is the case for the SV 40 virus.

The origin of replication of SV 40 has been extensively studied. It is short but requires T antigen coded by the SV 40 genome and primate cell components to replicate. SV 40 vectors transferred into cells constitutively synthesising T antigen (Cos cells) are commonly used by experimenters to express foreign genes at a high rate. This vector does not replicate in non-primate cells and, thus, cannot be used to generate transgenic animals.

Herpes viruses remain in a latent state in infected cells. Each cell contains one or a few copies of the viral genome, which are transmitted to daughter cells. Under certain circumstances, and particularly when the infected organism becomes immunodepressed, the viral genome replicates at a high rate, inducing a relatively pathological state. This is the case for the *Herpes simplex* virus, cytomegalovirus, Epstein–Barr virus and others.

Theses viruses have their own origin of replication and a pseudocentromeric system. These two elements have been particularly studied in the Epstein–Barr virus (EBV). The minimum regions required for the replication of the EBV genome and its transmission to daughter cells are the origin of replication region (ori P) and the EBNA 1 gene (Epstein–Barr nuclear antigen). The ori P region contains two regions having distinct functions. Both bind the EBNA 1 protein. One of the ori P regions is the origin of replication proper. The second region is required for the transmission of the viral genome to daughter cells.

Extensive studies have indicated that the EBV origin of replication is active only in primate and dog cells. Experiments carried out in several laboratories have shown that the EBV origin of replication can be deleted and replaced by genome DNA fragments from humans or other mammals. Some of these genome DNA fragments contained an origin of replication, allowing the vector to replicate and to be transmitted to daughter cells, even in transgenic mice (Kelleher *et al.*, 1998). However, it has not yet been proven that these vectors are stable and transmitted to progeny.

In all cases, the EBV vectors must contain the EBNA 1 gene and the ori P region to which this protein binds. Experiments carried out in several laboratories led to the conclusion that the EBV EBNA 1–ori P system is not species specific. It is even perfectly active in yeast. It was originally observed that the vectors containing the EBNA 1–ori P complex bind unspecifically to chromatin in various eukaryotic cells. An elegant study recently showed that EBNA 1 binds to ori P and to a ubiquitous protein called EBP 2, which recognizes unknown components in chromatin (Figure 2.15). A circular vector containing the ori P region binding EBNA 1 and a yeast origin of replication was quite efficiently maintained in yeast expressing both EBNA 1 and human EBP 2 genes (Kapoor, Shire and Frappier, 2001). This suggests that episomal vectors containing an origin of replication active in mouse cells and the genes coding for EBNA 1 and EBP 2 proteins could be maintained as episomal circular vectors throughout the life of the animal and transmitted to progeny. The EBV genome contains 200 kb and EBV vectors could harbour more than 100 kb of foreign DNA.

Vectors containing the SV 40 origin of replication and a MAR sequence are also able to be stably maintained as episomes in cultured cells (Jenke *et al.*, 2002).

These kinds of vector currently studied in laboratories are shuttles since they contain both eukaryotic and prokaryotic replication systems.

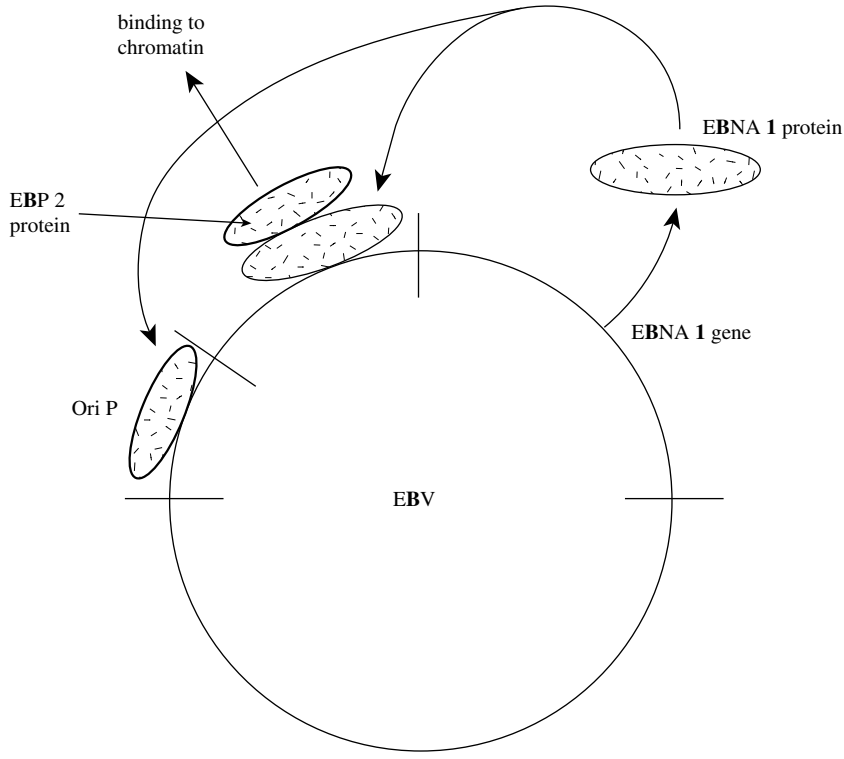


Figure 2.15 Structure of a circular episomal vector based on the Epstein–Barr virus (EBV) genome. The protein EBNA I, coded by the EBV genome, binds to another region of the EBV genome. The EBNA I protein binds to a cellular protein, EBP 2, which is associated with non-specific elements of chromatin. This centromere-like system allows the vector to be transmitted to daughter cells. The origin of replication of EBV is active almost exclusively in primate cells

They might be transmitted to intestine bacteria and disseminated into the environment. This could be easily avoided by removing the prokaryotic origin of replication.

Another possibility for designing autonomous vectors consists of using long genomic DNA fragments containing natural origins of replication, a centromere and telomeres. These vectors, known as human artificial chromosomes (HACs), can only be obtained after random mutations eliminating a large part of a chromosome but keeping the minimum elements to form an autonomous system (Vos, 1998; Voet *et al.*, 2001). These vectors are not easily manipulated and it is particularly difficult to

introduce foreign genes into them. Furthermore, these long genomic fragments contain numerous genes that may interfere with the physiology of the animal or with the gene of interest when it is added to the vector.

A 60 Mb murine chromosome consisting of murine pericentric satellite DNA has been generated in a rodent/human hybrid cell line. This structure was maintained for a long period of time in cultured cells and, interestingly enough, in transgenic mice as well with transmission to progeny (Co *et al.*, 2000).

Several years ago, an elegant work showed that a human chromosome, Chromosome 2, could be transferred into mouse genome, maintained and transmitted to progeny. The human chromosome was transferred from human fibroblasts to mouse ES cells by cell fusion. The ES cells harbouring the human chromosome were used to generate chimaeric mice. These animals expressed the immunoglobulin genes of Chromosome 2. Human monoclonal antibodies can thus be obtained from these mice (Tomizuka *et al.*, 1997).

Satisfactory episomal vectors are not really available. Such tools would be quite useful for researchers for studying genes in transfected cells but also for gene therapy and for generating transgenic animals. Slow but significant progress is being made in this field.

2.3.6 Vectors for gene replacement

A homologous recombination between two DNA fragments can occur in various organisms if they contain a common DNA sequence. A few hundred base pairs are needed in bacteria and at least 20–50 in yeast to trigger homologous recombination. In animal cells, several thousand base pairs are required to induce a homologous recombination. The basis of the mechanism is described in Figure 2.16. If a single homologous sequence is present in the genomic and the exogenous DNA, the recombination leads to a targeted integration of the foreign DNA (Figure 2.17). If two distinct homologous sequences are present in the exogenous DNA, two homologous recombinations occur independently, leading to a specific replacement of a genome domain (Figure 2.16). A foreign sequence may thus be precisely integrated into a given site of the genome. The foreign sequence may interrupt the targeted gene, which becomes inactive. This protocol is called gene knock-out (KO). The foreign sequence may be an active gene, which may or may not be related

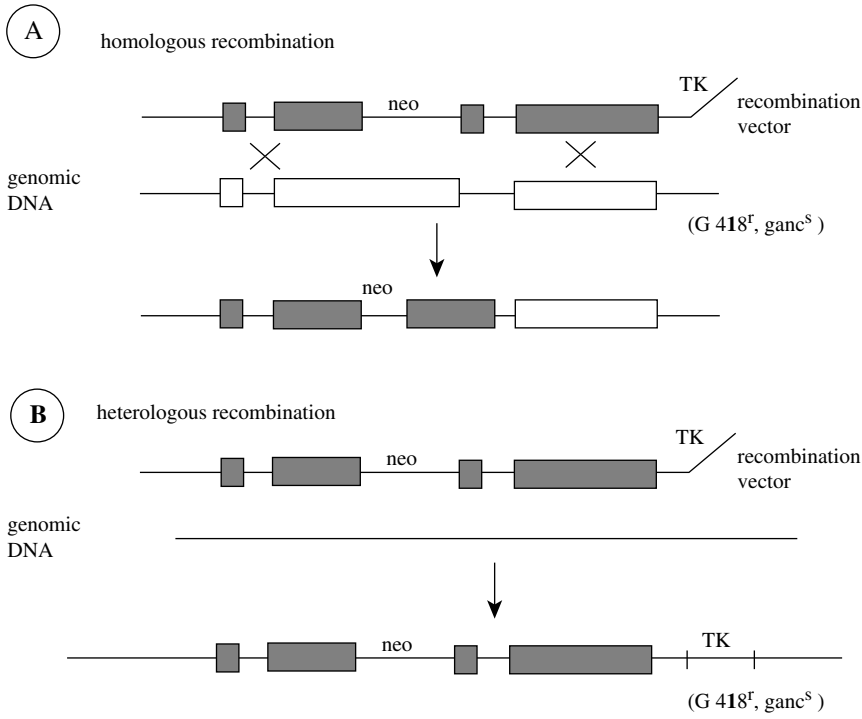


Figure 2.16 Double positive and negative selection to isolate cells in which a homologous recombination has occurred.
(A) Homologous recombination implies the integration of the neo^r gene and the elimination of the TK gene. These cells are resistant to G 418 and gancyclovir
(B) After integration of the vector into a random site, cells are resistant to G 418 but sensitive to gancyclovir

to the targeted gene. This targeted integration of a foreign gene is called gene knock-in.

2.3.6.1 Gene knock–out

Homologous recombination in animal cells is at least 100 times less frequent than random integration. The illustration shown in Figure 2.4 describes the principle of gene replacement but does not reflect reality. In practice, cells in which homologous recombination occurred must be selected. In order to do this, vectors such as those described in Figure 2.16

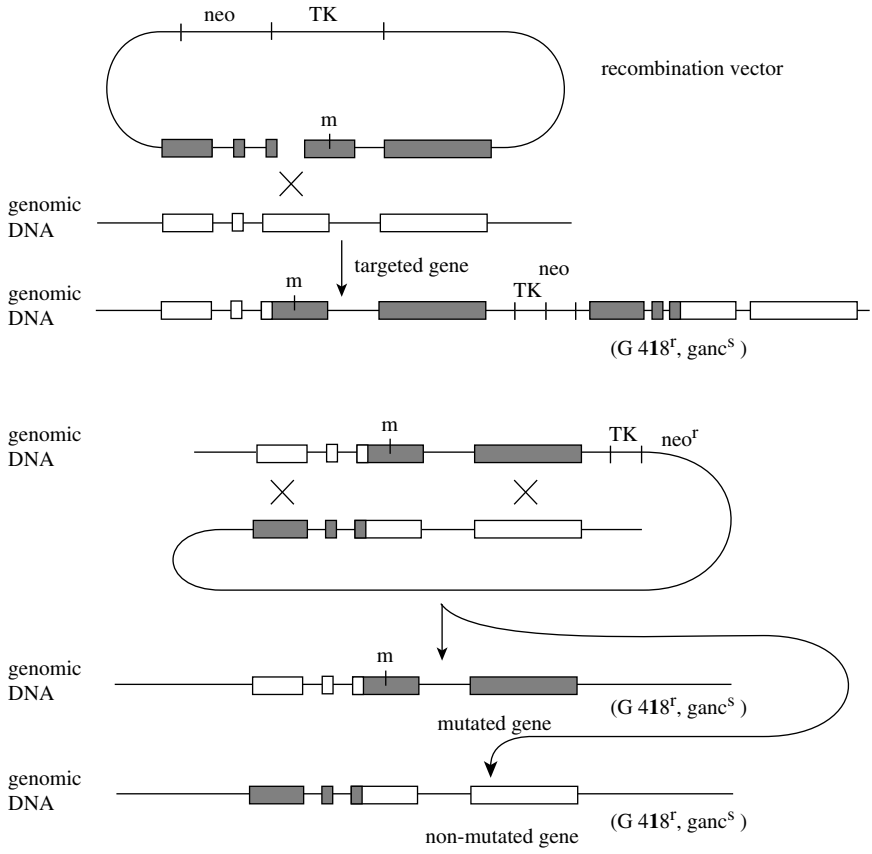


Figure 2.17 Replacement of a gene by a mutated variant using a double homologous recombination

are used. This vector contains the neomycin-resistant gene under the control of a mammalian gene promoter. This gene renders animal cells resistant to G 418 (also known as Geneticin). This gene is added between the two homologous sequences. Another selection gene is added in the periphery of the homologous regions, the thymidine kinase gene from the *Herpes simplex* virus. This kinase is known to transform chemical substances such as gancyclovir into cytotoxic molecules in dividing cells. This double selection generates clones in which the targeted gene was replaced by the sequence of the vector. However, the double selection also keeps clones in which the thymidine kinase gene was unduly eliminated by a rearrangement. Each clone must therefore be carefully studied

using Southern blotting and PCR to make sure than a genuine gene replacement has occurred.

If gene replacement has taken place in ES cells, chimaeric mice can be generated. A transmission of the ES genome to progeny generates animals that are heterozygous for the mutation. Homozygous animals can be further obtained by conventional reproduction (Figure 2.10). If gene replacement has taken place in somatic cells, cloning of animals by nuclear transfer is required to generate animals having the replaced gene (Figure 2.11).

The recombination vectors may contain a gene trap system such as the one described in Section 2.3.10. The vectors are constructed in such a way as to reconstitute a selection gene only when homologous recombination has occurred. This reduces the number of cell clones in which DNA structure at the targeted site must be examined, to a significant degree (Denning *et al.*, 2001).

The frequency of recombination including homologous recombination may be greatly enhanced by DNA cleavage, which induces the repair mechanisms. The introduction of a rare restriction site such as the one corresponding to the enzyme I-Sce 1 can initially be added to a genome by random or targeted integration. The expression of the gene coding for I-Sce 1 in cells cleaves the genomic DNA specifically at the chosen site of integration. The recombination vector then operates at a much higher level of efficiency (Cohen-Tannoudji *et al.*, 1998). This, of course, implies that the I-Sce 1 site has been added to the genome of the cells. On the other hand, it has been observed that the classical homologous recombination occurred at a higher frequency when the vector was in a linear form. In contrast and for no clear reasons, circular vectors appeared more efficient when the protocol involving DNA cleaving by I-Sce 1 was followed.

It is acknowledged that homologous recombination is more frequent in ES cells than in somatic cells. This complicates gene replacement using the cloning technique by nuclear transfer. It is not clear why recombination is more frequent in ES cells. The repair mechanism may be more active in these cells. On the other hand, ES cells replicate at a higher rate than most somatic cells. It is also believed that genes that are actively transcribed are more easily amenable to homologous recombination, although this was not observed in all cases. Many genes are transcribed in ES cells. This may maintain chromatin in an open configuration, which favours homologous recombination.

2.3.6.2 Gene knock–in

There are several reasons for replacing a functional gene by another functional gene. The new gene may be a mutant of the original gene. The knock-in may thus be used to study the activity of the mutant. This technique is also a way to replace an allele by another allele of interest for breeding without using the conventional selection, which co-selects several other genes that do not have a beneficial effect. This protocol can also replace an allele by a mutant never found in nature.

Knock-in may also be used to introduce a foreign gene not related to the targeted gene. This may be done to study the effect of a chromatin environment. The targeted site may have been chosen for its high capacity to express a foreign gene (Kolb *et al.*, 1999).

The gene described in Figure 2.17 allows the knock-in of a given gene by a mutant. A single event of homologous recombination initially occurs. It leads to the integration of the mutant at the targeted site but not to a gene replacement. A second event of homologous recombination between the original gene and the integrated mutant leads to replacement of the targeted gene by the mutant. This occurs spontaneously but at a low frequency. It implies that many recombinants have to be screened.

A more efficient way to replace a functional gene by another functional gene is described in Figure 2.18. It relies on the use of mutant ES cells devoid of the HPRT gene (hypoxanthine phospho-ribosyl transferase). These cells cannot survive in HAT medium, which contains hypoxanthine, aminopterine and thymidine. The HPRT gene is introduced at the targeted site by a double homologous recombination. The cells are selected by the HAT medium supplemented by ganciclovir. A second homologous recombination with a vector containing the exogenous gene is then performed. The recombination takes place when the HPRT gene is eliminated. The cells in which the event occurred are selected with 6-thioguanosine, to which they are resistant, and by HAT medium, to which they are sensitive.

A more universal tool to knock a gene in is described in Figure 2.19. This system does not imply use of HPRT$^{-}/_{-}$ ES cells. The first recombination vector contains the gene to be introduced into the cell genome and the neo^{r} gene bordered by two LoxP sites. The cells in which the first recombination event occurred are selected by their resistance to G 418 and gancyclovir. The neo^{r} gene is further eliminated by adding Cre recombinase (or a gene coding for this enzyme) into the cells. The

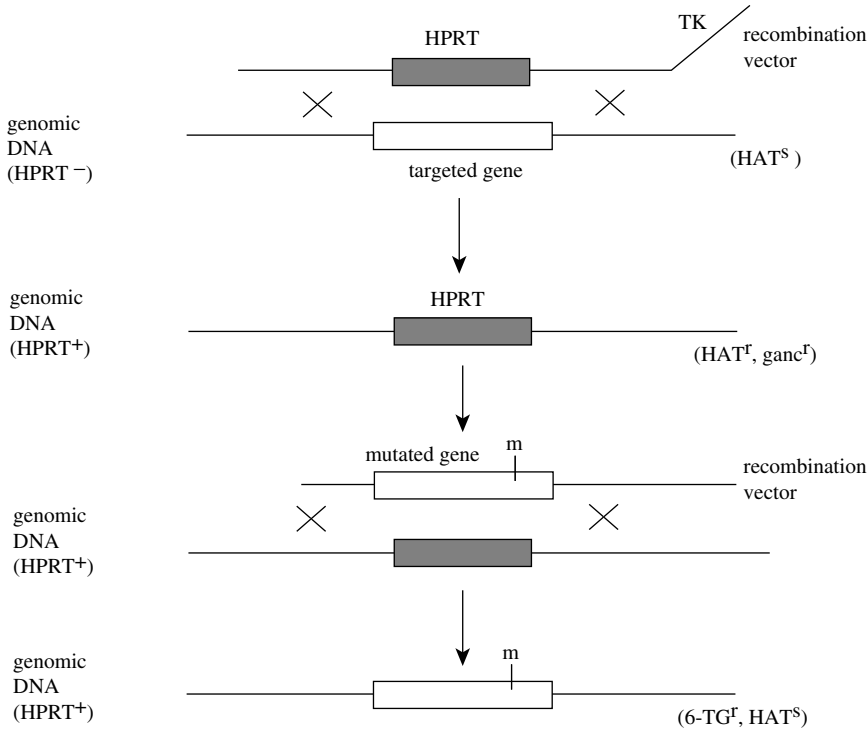


Figure 2.18 Replacement of a gene by a mutated variant using a double homologous recombination and selection by the HPRT gene (hypoxanthine phospho-ribosyl transferase). Cells harbouring the HPRT gene are resistant to the HAT medium and the cells devoid of the gene are sensitive to the HAT but resistant to 6-thioguanosine

recombination of the LoxP sites triggers the elimination of the neo^r gene, leaving only one LoxP site, which contains only 34 nucleotides. The Cre–LoxP system is described in more detail below. The elimination of the neo^r gene can be performed in the ES cells before they are used to generate chimaeric mice. It also may be done in the one-cell stage embryo of the next generation.

2.3.7 Vectors for the rearrangement of targeted genes

Homologous recombination within a given genome is a physiological event that occurs in various situations. The genome rearrangement that generates the functional immunoglobulin genes is an example. It may be

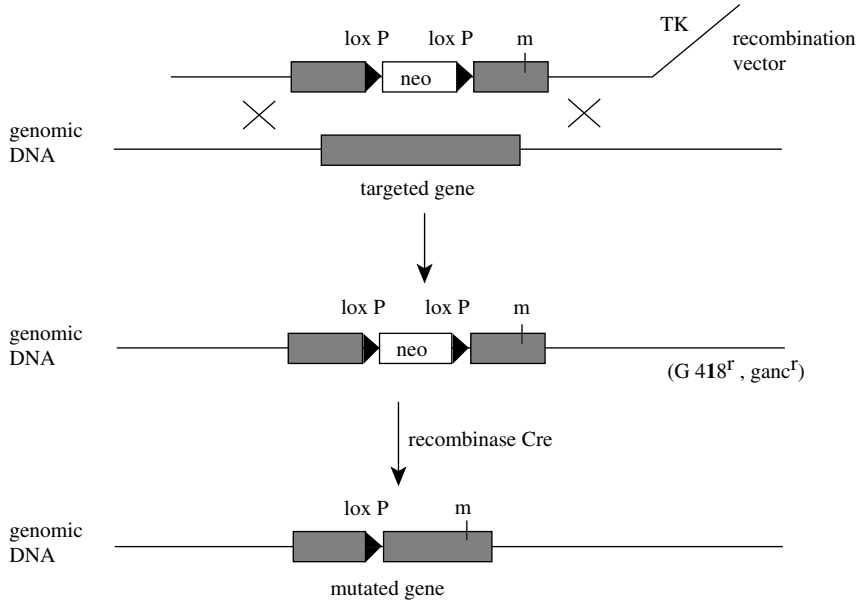


Figure 2.19 Targeted gene mutation by homologous recombination with elimination of the selection gene. The neo^r gene previously bordered by two LoxP sequences is eliminated by the action of the Cre recombinase

important to induce genome rearrangement in sites naturally not able to do so and in a controlled manner.

Two systems not found in the animal kingdom have been implemented in order to do this. One of them derives from the bacteriophage P1. This genome contains LoxP sequences formed by 34 nucleotides, which recombine with high efficiency and specificity only in the presence of the phage Cre recombinase. The other system is fundamentally similar but originater from a yeast. FRT sequences recombine with high efficiency and specificity in the presence of the Flp recombinase. The Cre–LoxP system is the most popular and it is used for many different studies (Nagy, 2000). Various mutants of LoxP, FRT sequences are being used to modulate recombination efficiency or specificity. The two recombinases also exist in the form of different mutants.

In practice, the LoxP or FRT sequences must be added to the gene constructs. A DNA region bordered by LoxP sequences is said to be floxed (for flanked by LoxP).

A major advantage of the Cre–LoxP and FRT–Flp systems is that the recombination events occur only when the recombinases are present. The cells or the animals harbouring the floxed DNA do not have the recombinases. These enzymes may be added to the cells *in vitro* or *in vivo* in the transgenic animals by a direct microinjection into the cytoplasm or by transfection methods for proteins. Alternatively, the recombinase genes can be brought to the cells by adenoviral vectors, which can be injected into a given tissue of the transgenic animal harbouring the floxed DNA. Moreover, plasmids containing the recombinase genes can be introduced into one-cell embryos or into somatic cells. The plasmids have little chance of integrating in their circular form. They rapidly disappear during cell multiplication and the presence of the recombinases is transient.

The recombinases may also be transmitted to the transgenic animals by transgenesis. Transgenic mice harbouring the recombinase genes may be prepared and crossed with the lines harbouring the floxed DNA.

Ideally, the recombinase genes should not be expressed in many situations in all tissues or permanently. If this is required, the recombinase can be put under the control of promoters working in all cell types. In contrast, tissue-specific promoters can restrict LoxP recombination in cells in which the promoters are active. The expression system controlled by tetracycline and derivatives (see Section 2.3.11) allows the recombinase genes to be expressed only in a given cell type and only when tetracycline is given to the animals.

Another regulation step may be added to control Cre recombinase even more closely. Fusion genes containing the Cre recombinase coding sequence and part of the steroid receptors have been constructed. The fusion protein is only active in the presence of the steroid, which induces a modification of the protein conformation.

An NLS (nuclear localization signal) sequence targeting the Cre recombinase to the nucleus may also be added. This allows the recombinase to concentrate in the nucleus and to be efficient even when present in low concentrations. This situation may occur when cell-specific but weak promoters are being used to express the recombinase genes.

All these sophisticated systems aim at inducing LoxP recombination as precisely as possible to mimic natural gene expression as much as possible or to create the most appropriate experimental conditions.

It is also important to avoid a high expression of the recombinase genes. These enzymes show some cytotoxicity at high concentrations,

most likely due to their capacity to recognize LoxP or FRT-like sites in animal genomes.

2.3.7.1 Gene deletion

The method for deleting a gene is described in Figure 2.19. It may be used for several purposes. It may be a way for knocking a gene in. The same method may be implemented to induce a conditional gene knock-out. In order to do this, the gene to knock out must first be floxed by homologous recombination. The expression of the Cre recombinase in a given tissue and at a chosen period induces deletion of the floxed sequence and, thus, the gene knock-out. This method is required when the effect of a gene inactivation is to be obtained only in a controlled manner and not in the early embryo as is generally the case. This may help to create a better model, mimicking a biological function or a human disease. It is necessary when the gene knock-out is lethal in the embryo, preventing a relevant study of the gene inactivation.

The specific deletion of a floxed DNA may also be required to eliminate sequences that were useful at a given step of the experiment but not beyond. This is the case for selection genes added with the gene of interest to transfer genes into cells subsequently used to generate transgenic animals. The selection gene may be deleted in the cell clones or in the embryos of the next generation.

2.3.7.2 Gene activation

A gene may be inactivated by the introduction of an inhibitory sequence between the promoter and the cap site. This sequence may be withdrawn by the action of the Cre recombinase when desired if it was previously floxed (Figure 2.20).

2.3.7.3 Induced gene inactivation

An element essential for the expression of the studied gene may be floxed either in the genome by homologous recombination or during gene construction. The presence of the Cre recombinase may inactivate the gene by removing the essential element in a controlled manner (Figure 2.21).

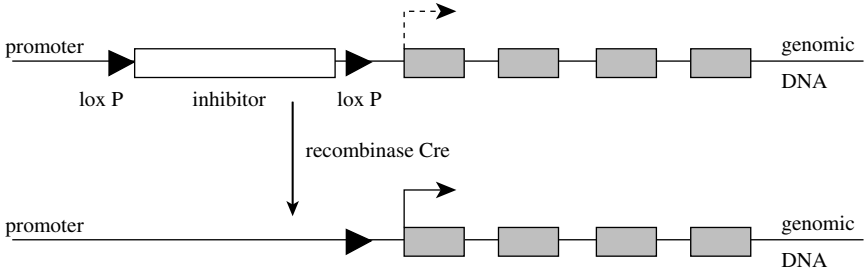


Figure 2.20 Conditional activation of a gene by elimination of an inhibitor sequence. The promoter is separated from the gene by an inhibitory sequence bordered by two LoxP sequences. The Cre enzyme induces a recombination of the LoxP and an elimination of the inhibitor, leading to gene activation

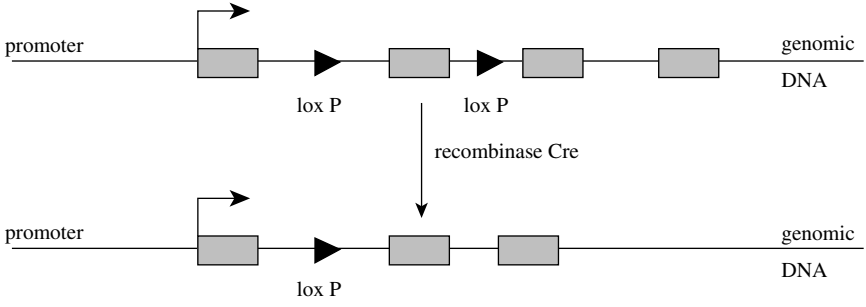


Figure 2.21 Conditional inactivation of a gene by the elimination of an active element. A region of the gene (or of the promoter) is previously bordered by two LoxP sequences. The Cre recombinase induces the elimination of the floxed region, leading to gene inactivation

2.3.7.4 Deletion of genomic fragments

Classical homologous recombination may delete a chosen fragment of a genome. This can be achieved by adding two remote genomic sequences to the recombination vector. The gene replacement will eliminate the genomic region located between the two sequences of the vector (Figure 2.22).

The same goal may be achieved by using the Cre–LoxP system. Two LoxP sequences may be added in the vicinity of remote regions of the genome in a recombination vector. The vector can replace the targeted genomic region by homologous recombination. The deletion of the targeted region is triggered by the presence of the Cre recombinase (Figure 2.23).

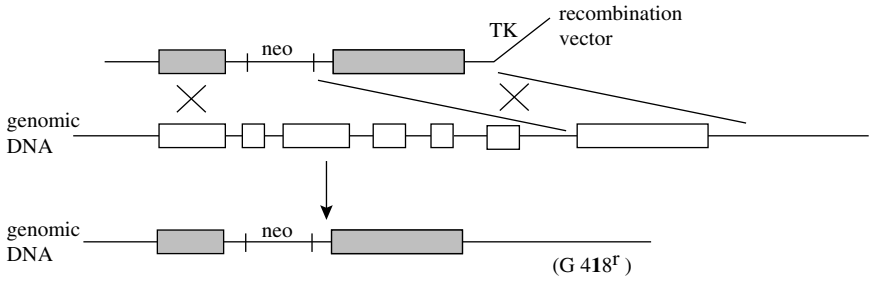


Figure 2.22 Deletion of a chromosome fragment by homologous recombination. The recombination vector contains two homologous and distal regions of a chromosome. The recombination eliminates the region between the two homologous sequences

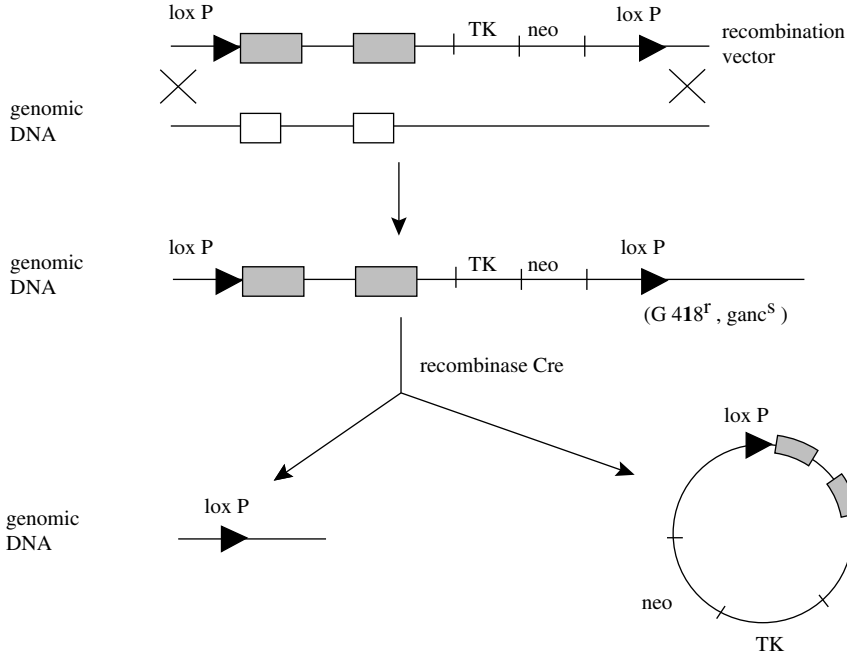


Figure 2.23 Deletion of a chromosome fragment by an induced homologous recombination. A vector is used to introduce two relatively distant LoxP sites. The Cre recombinase induces deletion of the chromosomal region between the two LoxP

Relatively long DNA sequences (up to 50 kb) may be deleted with these tools. Longer genomic regions may be deleted with the Cre–LoxP system. For this purpose, two LoxP sites must be introduced independently into targeted regions of the genome by homologous recombination. The Cre recombinase will induce deletion of the region located between the two LoxP sequences (Figure 2.24). The frequency of the LoxP recombination becomes lower when the two sites are very distant. Sophisticated selection systems are then necessary to keep only the cell clones in which the recombination has occurred. One possibility described in Figure 2.24 consists of adding the LoxP sequences to an intron of the HPRT gene. The first vector carrying the first LoxP to the genome contains the first part of the HPRT gene and the second gene is then reconstituted and functional only if the expected recombination event occurs. The cells in which the genomic deletion has taken place can be selected by the HAT medium.

Variable regions of the genome can be deleted at a given site using a retroviral vector as the second recombination vector carrying the second LoxP site. The retroviral vectors integrate randomly into the genome. The recombination between the two LoxP sequences is then induced in different cell clones previously selected as harbouring the retroviral genome at different genomic sites. Deletions of genomic regions of variable length are thus obtained in the different clones (Figure 2.25).

2.3.7.5 Interchromosomal recombination

The Cre–LoxP system may also be used to induce recombination between two chromosomes at chosen sites. This was achieved by introducing LoxP sequences into predetermined sites of two chromosomes by homologous recombination. The recombination of the LoxP sequences induced a translocation between chromosomes 12 and 15 in mice (Figure 2.26). This experimental protocol can mimic natural chromosome translocation.

If should be kept in mind that the frequency of recombination between two chromosomes is a very rare event and that it has been successful in only a limited number of cases.

A somewhat different approach proved to enhance the frequency of recombination. This method, known as TAMERE (targeted meiotic recombination), is based on the expression of the Cre recombinase, specifically found in sperm precursors (Herault *et al.*, 1998). The

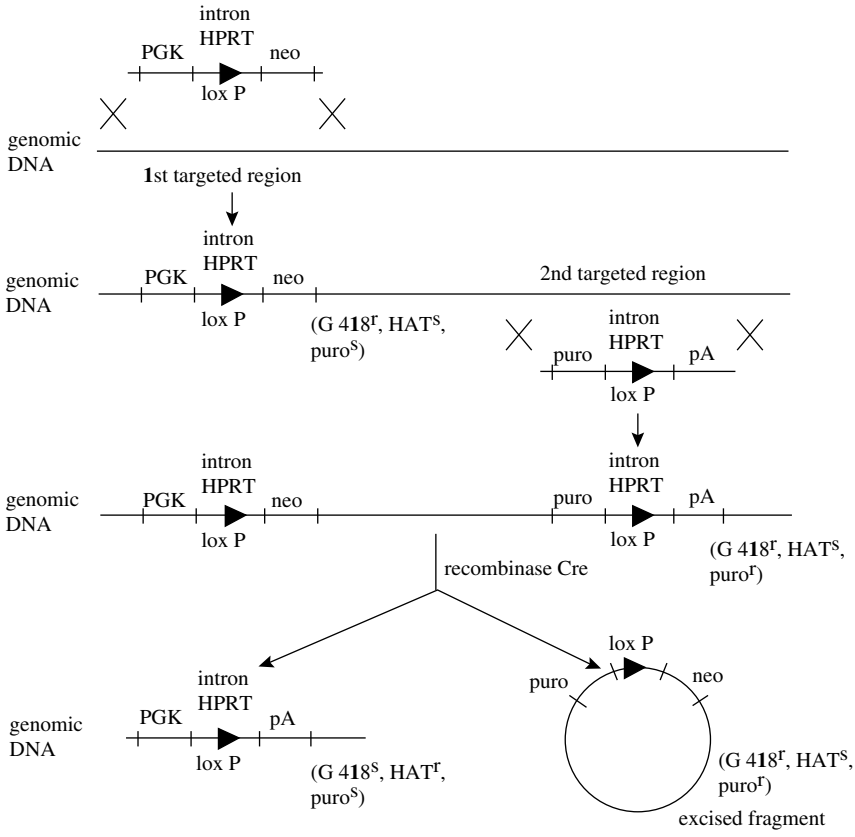


Figure 2.24 Induced deletion of very long chromosome fragments. Two LoxP sites are introduced through targeted integration with by two independent homologous recombinations. The Cre recombinase eliminates the region between the two LoxP sequences. This event is very rare and selection genes are required to eliminate cells in which non-specific recombination has occurred

recombinase is believed to be bound to the LoxP sites before fertilization and to favour recombination in the embryo.

Multiple interchromosomal recombinations may be envisaged with the Cre–LoxP system. The major possibilities are described in Figure 2.27.

2.3.8 Targeted integration of foreign genes

The Cre–LoxP system may be theoretically used for purposes other than the deletion or recombination of genomic regions. It may also induce

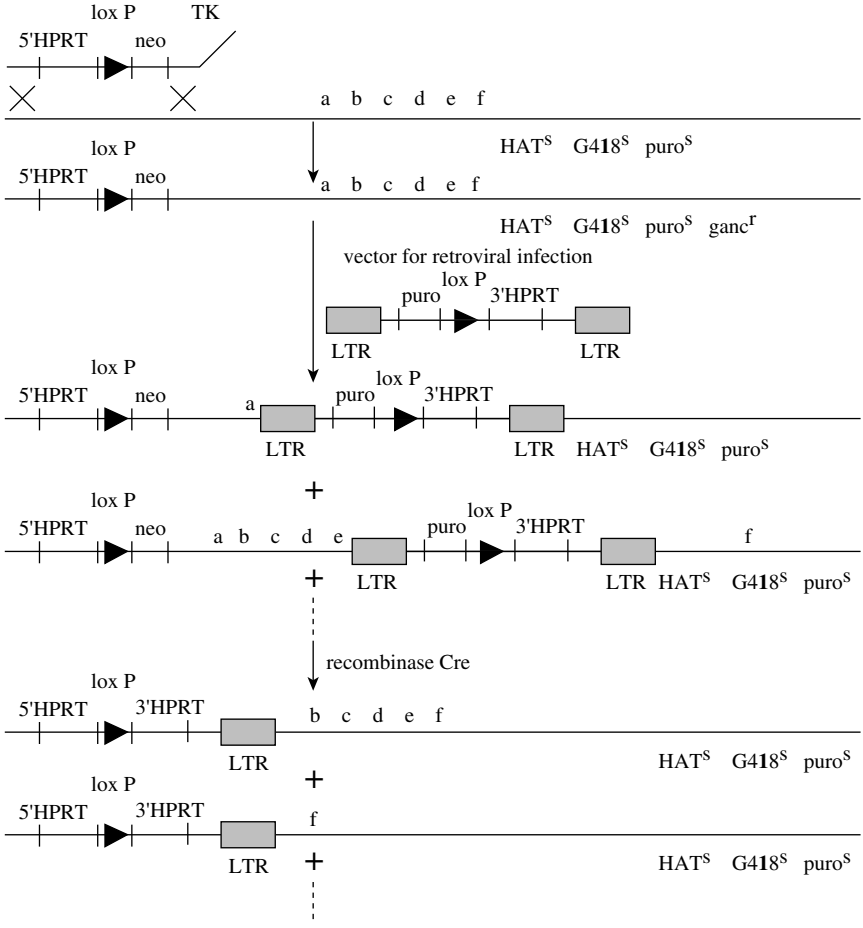


Figure 2.25 Induced deletion of a chromosomal region of variable length. An LoxP sequence is introduced into a given site of the genome by homologous recombination. A second LoxP is randomly introduced by a retroviral vector. The Cre recombinase eliminates the region between the two LoxP. Selection genes are required to keep only cells in which the specific recombination has occurred

integration of a DNA fragment containing a LoxP sequence into a genome site that also contains a LoxP site (Figure 2.28).

The LoxP site may have been introduced randomly into the genome by DNA microinjection (Figure 2.29). It may also be introduced into a chosen site using homologous recombination (Figure 2.28).

The gene of interest may be in a plasmid that also contains a LoxP sequence. The co-introduction of this plasmid and of another plasmid

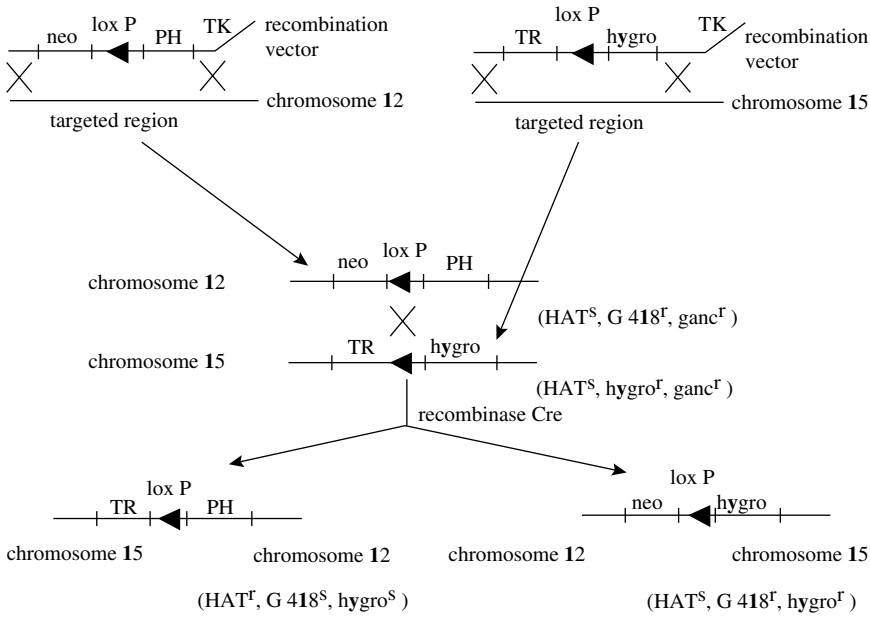


Figure 2.26 Chromosome translocation induced by homologous recombination. Two LoxP sequences are introduced into given sites on two chromosomes. The Cre recombinase induces the association of the two chromosomes. The HPRT gene reconstituted by the specific LoxP recombination is used to select cells in which the chromosome translocation has occurred

containing the Cre recombinase gene makes it possible for the gene to be integrated at the genomic LoxP site (Figure 2.28).

The plasmid sequence of the vector may be removed prior to its introduction into the cell. This prevents the integration of the plasmid sequence in the vicinity of the gene of interest and possibly its extinction.

Another way to favour the expression of the transgene integrated at the genomic LoxP site consists of associating gene insulators with the LoxP sequence (Figure 2.30). Insulators attenuate or suppress extinction of the transgene by chromatin (see Section 2.3.11).

In practice, the integration of a foreign gene at a genomic LoxP site is quite specific but not frequent. Cells in which the targeted integration has occurred must be selected before they are used to generate transgenic animals. This means that the targeted integration by the Cre–LoxP system cannot be achieved with a reasonable chance of success in a one-cell embryo after DNA microinjection.

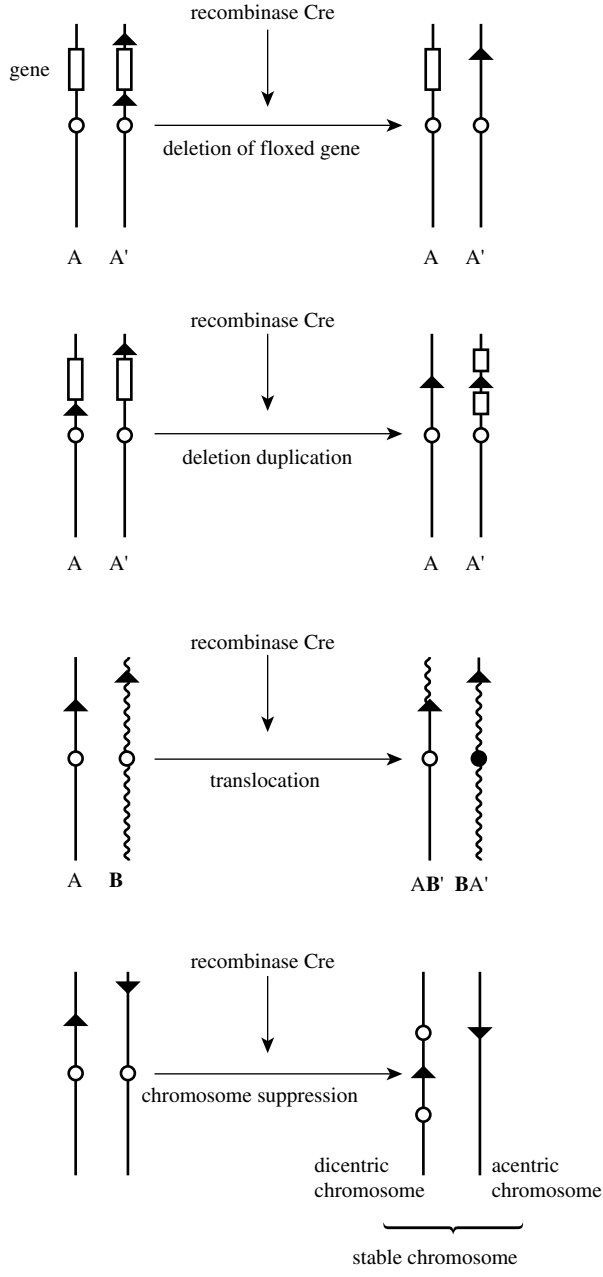


Figure 2.27 Different possible intra- and inter-chromosomal recombinations induced by the Cre–LoxP system

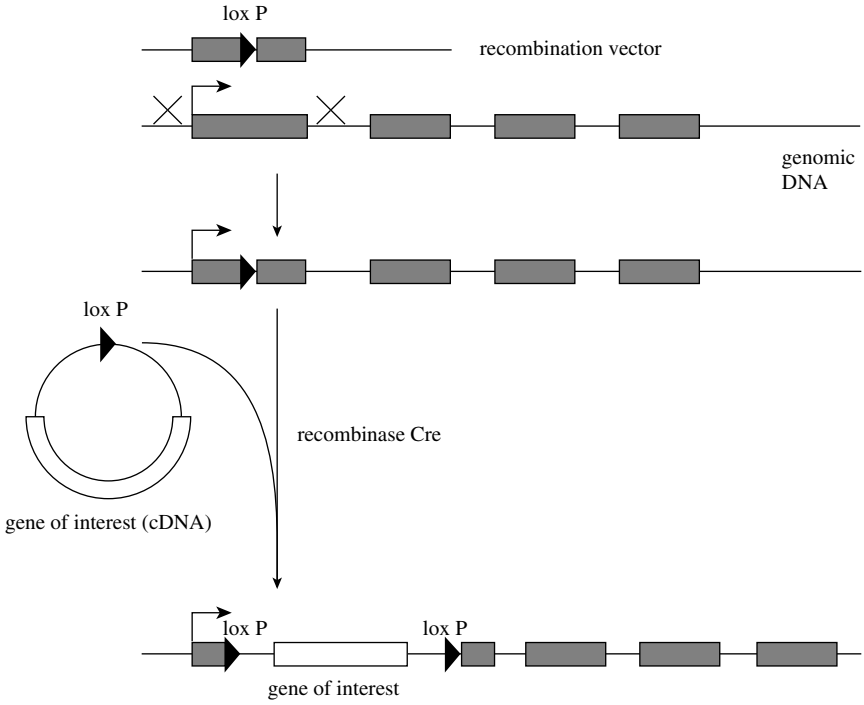


Figure 2.28 Targeted integration of a foreign gene into a genome by the Cre–LoxP system. An LoxP sequence is introduced into a given site by homologous recombination. The foreign gene bordered by a LoxP sequence is introduced into the genomic LoxP site

This low recombination frequency is due, to a large extent, to the fact that the integration site contains two LoxP sequences, one provided by the foreign DNA and the other previously added to the genome. These two LoxP sequences have a greater chance of recombining under the action of the Cre recombinase than those present in the genome and in the vector. Hence, the balance between these two contradictory recombination events does not favour integration (Figure 2.31).

Several methods have been defined to modify the balance in favour of integration. One of these methods is described in Figure 2.31. It is based on the use of two LoxP mutants that can each recombine LoxP sequences containing both mutations that may no longer be processed by the recombinase. This makes integration of the foreign gene possible but not its withdrawal (Araki, Araki and Yamamura, 1997). In practice, this

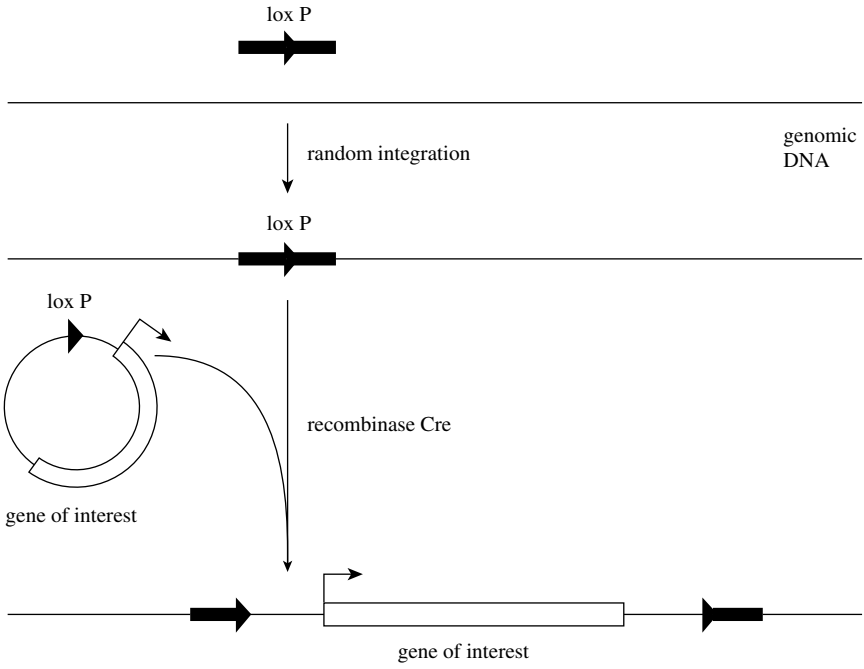


Figure 2.29 Introduction of a foreign gene into a random site of the genome by the Cre–LoxP system. An LoxP sequence is randomly introduced into the genome by microinjection or transfection. A vector containing the foreign gene and a LoxP site targets the integration into the genomic LoxP site. The genomic LoxP sites, which allow a satisfactory expression of the foreign gene, may be used to introduce mutant genes or different genes

approach is of limited interest due to the poor capacity of the two LoxP mutants to be efficiently recombined by the recombinase.

Two other methods have been proposed. One, known as RMCE (recombinase-mediated cassette exchange), uses two different LoxP mutants that can recombine with themselves but not with each other. The targeted genomic region and the gene to be introduced are both floxed with the two LoxP mutants. A double recombination event induces integration of the foreign gene at a relatively high frequency (Feng *et al.*, 1999). The efficiency of the method relies on the normal capacity of the LoxP mutants to be recombined by the Cre recombinase but not to recombine with each other.

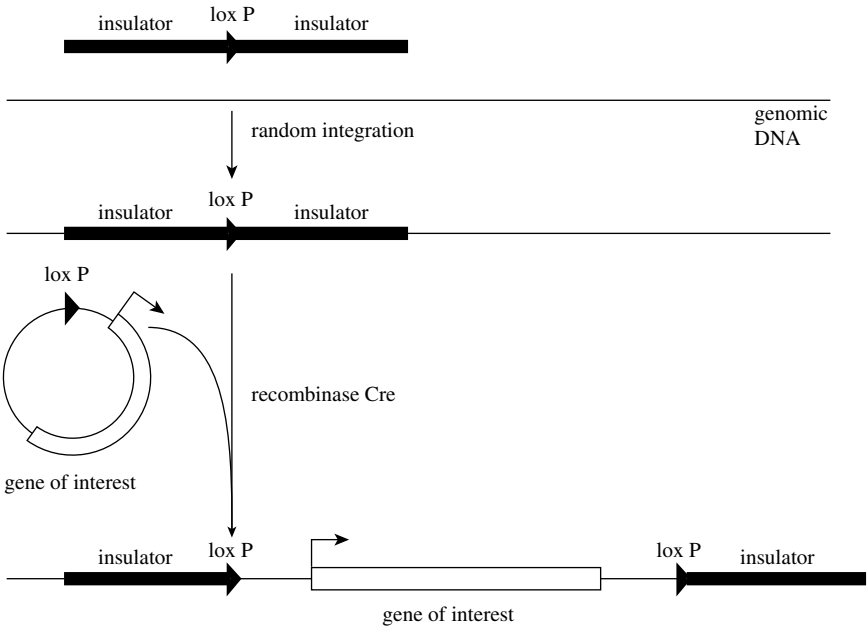


Figure 2.30 Introduction of a foreign gene into a random site previously surrounded by sequences favouring gene expression, such as insulators. The protocol is that of Figure 2.29 but LoxP is associated with insulators

The second method also uses two LoxP or FRT cassettes. The addition in excess of the recombination vector enhances the frequency of the integration event. The selection of the clones in which the targeted integration has occurred using the gene trap system (see Section 2.3.10) allows for a very high yield of targeted integration (Baer and Bode, 2001).

Targeted integration using the Cre–LoxP system has proved efficient for introducing a foreign gene at genomic sites, making reliable transgene expression possible. Its main advantage is to allow researchers to consistently introduce different mutants of a gene at the same genomic site. This considerably reduces the side effects due to random integrations in the genome. A few transgenic mouse lines may give relevant results in this way whereas random integration requires the study of many mouse lines to obtain data reflecting the activity of the different gene constructs rather than position effects (Day *et al.*, 2000).

It is also interesting to mention that the Cre–LoxP system can be implemented to engineer vectors in bacteria. This is particularly useful

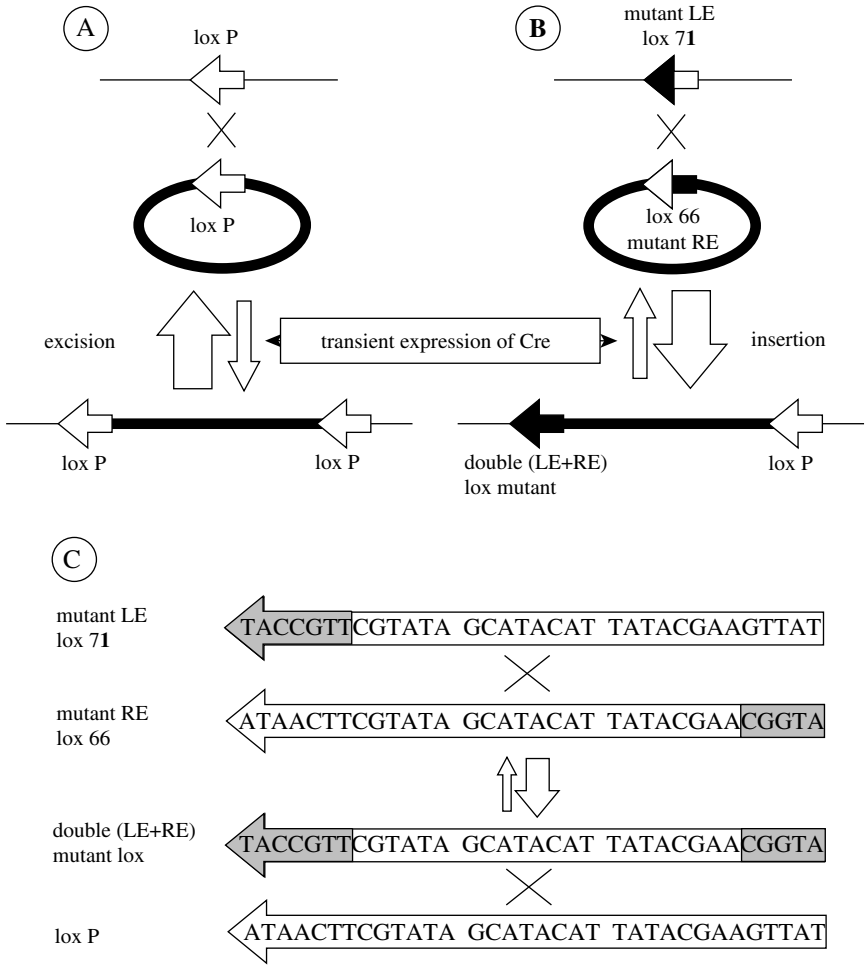


Figure 2.31 Utilization of mutated LoxP sequences.
(A) A plasmid harbouring a normal LoxP has little chance of remaining integrated in a genome at a similar LoxP site since the Cre recombinase will eliminate the integrated DNA fragment very rapidly
(B) Two mutated LoxP sequences make integration possible but not elimination of the foreign DNA fragment
(C) The sequence of the two LoxP mutants before and after recombination

when BAC vectors are manipulated. This method makes it possible to introduce foreign DNA fragments into factors without using restriction sites (Liu *et al.*, 1998).

2.3.9 Non-classical vectors for the recombination of targeted genes

Although efficient, the available methods to replace a gene by homologous recombination are laborious. They all imply that homologous recombination is performed in cultured cells, which have to be selected and used to generate an embryo by chimaerism, nuclear transfer or, potentially, fertilization if sperm precursors are used. Ideally, homologous recombination should occur at a frequency high enough to be directly triggered in one-cell embryos by vector microinjection.

The methods described above and implemented for gene therapy (Section 2.2.2) must be greatly improved if they are to be extended to transgenesis. This is obviously the case of chimeraplasty based on the use of RDOs (Graham *et al.*, 2001) (Figure 2.32) or triple helix forming oligonucleotides (Giovannangeli and Hélène, 2000), capable of inducing point mutation in a targeted gene. The genomic DNA cleavage at a pre-integrated I-Sce 1 site markedly enhances homologous recombination, but it has not yet been proven that the frequency is high enough for this method to be used directly in one-cell embryos. The same is true for the methods based on the use of the Cre–LoxP or Flp–FRT systems described above.

Another possible approach might consist of using bacterial recombinases. A pioneer study indicated that the bacterial Rec A enzyme associated *in vitro* with a mononstrand DNA sequence was able to induce

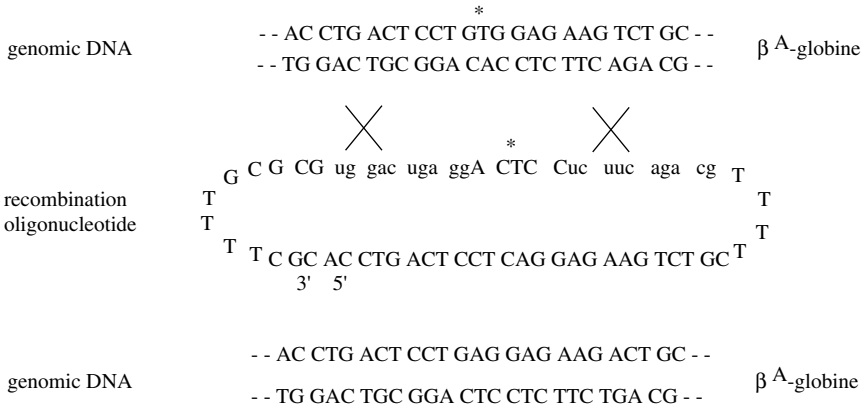


Figure 2.32 Use of RDOs to induce genomic mutation. The capital letters designate RDOs whereas the lower case letters designate ribonucleotides. The homologous recombination, which is facilitated by the presence of ribonucleotides, induces a targeted mutation in the genome

a homologous recombination of the corresponding gene in cell-free systems, in mammalian cells and also in mouse embryos after microinjection (Pati, 1998). These data have not been confirmed so far.

Obviously, many additional studies are required to get a better understanding of the mechanisms responsible for homologous recombination in higher vertebrate cells. They could lead to the optimized use of these mechanisms for targeted recombination in animal genomes. They should also reveal to what extent bacterial recombinases can be used for this purpose in animal cells (Sonoda *et al.*, 2001; Vasquez *et al.*, 2001).

2.3.10 Vectors for gene trapping

Gene trapping is a general method to identify a rare recombination event. It is based on the reconstitution of a functional gene by recombination which may or may not be homologous.

Several types of vector may be used for gene trapping. The basis of the method is described in Figure 2.33. In this case, the vector contains a splicing acceptor site, a reporter coding sequence and a transcription terminator. The reporter sequence can only be expressed when the vector has been integrated into a host gene. The reporter gene reveals in which cell type and when the host gene is expressed. The integrated vector is also a marker and a hook to identify and clone the host gene from a genomic bank or using 5′RACE.

Gene trapping is currently used to select cells in which a specific recombination event has occurred. Typical examples are shown in Figures 2.34–2.36. The vector contains one part of the reporter gene and the genome harbours the other part. The functional gene is reconstituted only when a targeted recombination has occurred. The cells in which the event took place can be selected by an antibiotic, a specific medium or by a direct visualization if the protein coded by the reporter gene can generate a colour. This is the case for the β-galactosidase and the GFP (green fluorescent protein) genes.

Gene trap is used to select ES or somatic cells in which a homologous recombination has occurred. After selection, the cells can be used to generate embryos by the formation of chimaerae or by nuclear transfer (Denning *et al.*, 2001). This method eliminates the cells in which a non-strict homologous recombination has occurred more efficiently and it reduces the number of cell clones to be examined before embryo generation.

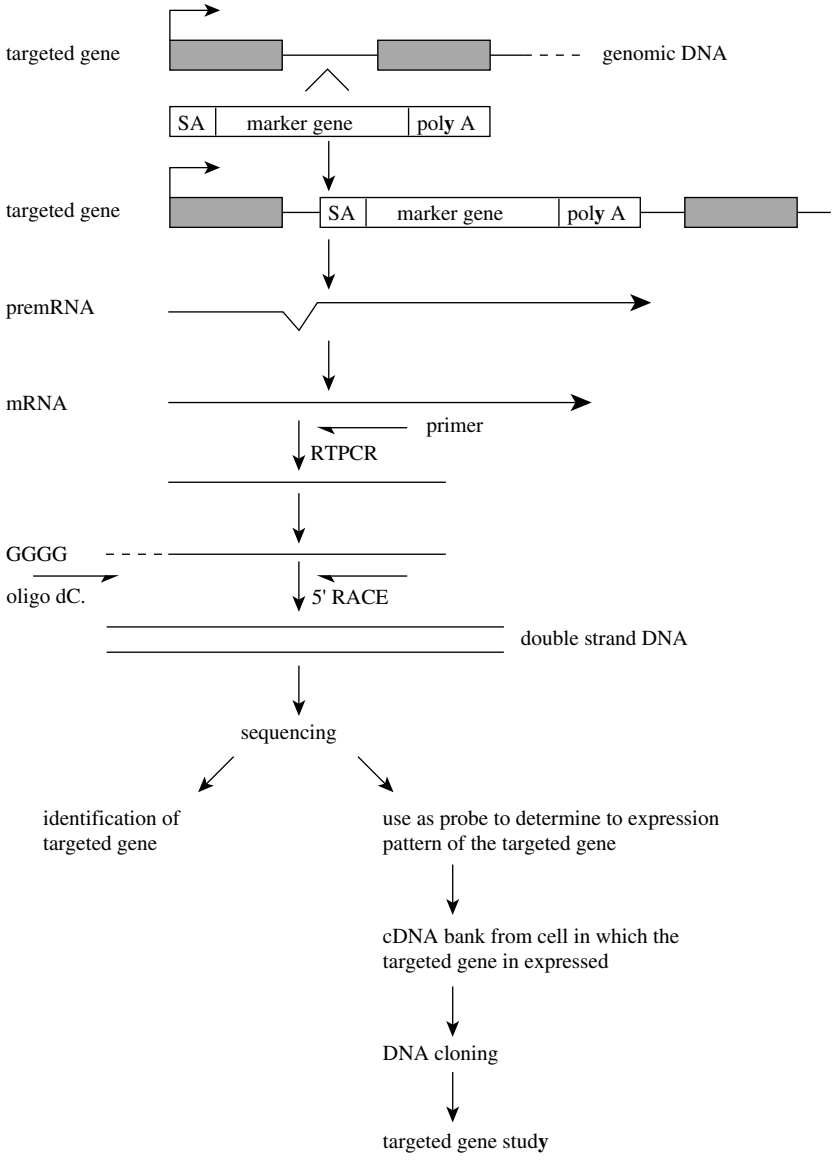


Figure 2.33 Principle of gene identification by gene trapping. A vector containing a marker gene preceded by a splicing acceptor site (SA) and followed by a transcription terminator (poly A) is expressed only when integrated within a functional gene. The cells that express the marker gene are selected and the gene in which the marker gene is inserted is identified and studied

Gene trapping is more specifically used to identify unknown genes by their function. Genome and EST sequencing is providing researchers with essentially all the coding sequences of genomes. Several techniques such as mRNA display, DNA array and others are capable of showing the pattern of gene expression of a given cell type under defined physiological conditions. This is often insufficient to reveal the biological function of a gene. The use of microsatellite markers in individuals of families that either show or do not show a phenotypical property may contribute to identifying the role of a gene (Figure 1.6). This is possible only if specific alleles are responsible for the phenotypical effect and if families of individuals that either harbour or do not harbour the mutation can be generated.

A systematic search of genes involved in a given biological event can be achieved with gene trapping. In practice, the following protocol is implemented (Cecconi and Meyer, 2000; Medico *et al.*, 2001). The vector is transfected into ES cells. A first selection eliminates the cells in which the vector has not been transferred or integrated. The selected cells are those introduced into recipient embryos to generate chimaerae. During the development of the embryo or later, in newborns or adults, the expression of the reporter gene can be revealed by the colouration of the cells or by the enzymatic activity of the protein coded by the reporter gene. The most commonly used reporter gene is a fusion sequence containing the β-galactosidase and the neomycin-resistant genes. This reporter gene is known as β-geo for this reason. It makes a visualization by the β-galactosidase region and a selection by the neomycin-resistant gene possible. Some researchers have constructed another fusion gene, β-phleo, where phleo is the phleomycin-resistant gene, which leads to a better selection of the clones.

In a certain number of cases, the integration of the vector interrupted and inactivated a host gene. This is, to some extent, a non-targeted gene knock-out. The inactivation of the host gene may induce an observable phenotypical effect. A correlation is then established between a given biological mechanism and a gene. The interrupted gene may then be cloned and studied.

This method has specific advantages and drawbacks. The major advantage is that an unknown gene or a gene of unknown function may be identified in this way without any prior hypothesis. The limitations are multiple. The frequency of integration of the vector into a functional gene is low. The interruption of the host gene may generate no observable phenotypical effect. The phenotypical effect or the expression of the

reporter gene may only occur during a short period of the animal's life. This event may then not be observed by the experimenter. The expression of the reporter gene reflects the presence of the mRNA for the interrupted host gene. It may happen that this mRNA is present but not translated for some biological reason. This may prevent the reporter gene from being translated. An mRNA may be present and translated without having a real function in a given cell type and without being involved in a specific physiological event. Indeed, it is by no means certain that evolution has developed mechanisms strictly restricting expression of the genes having a function in a given period of the life of the organism. The trapped gene may thus not have the expected function suggested by the fact that it is expressed.

In practice, gene trapping is a laborious method but still offers certain advantages. It has been essentially used only in mice so far. This limitation is mainly due to the fact that the ES cells are currently available only in this species and that cloning by nuclear transfer has not been implemented yet for this purpose. Gene trapping has proven itself to be quite useful for identifying genes involved in embryo development. Observations made so far have revealed that one-third of the trapped genes are expressed in specific cell types. Another third are expressed in all cell types. In the other cases, the expression was transient and not observed by the experimenters.

The success of the gene trap method relies on the possibility of screening many ES cell lines and embryos. To facilitate the research task and to generate a more potent tool, banks of ES cells in which gene trap vectors have been introduced have been prepared. Chimaeric mice can be prepared in large numbers using these ES cells, improving the chance of identifying an interesting mutation (Zambrovicz and Friedrich, 1998; Jackson, 2001).

Different types of gene trap vectors have been designed to enhance the efficiency of the method or to identify genes coding for proteins having particular properties.

2.3.10.1 Vectors to trap coding sequences

Vectors containing only an acceptor splice site, the reporter cDNA without any initiation codon and a transcription terminator may become functional when they are integrated at any site of the transcribed region of a host gene. The protein coded by the recombined

gene is always a fusion protein if the integration of the vector occurs after the initiation codon of the interrupted gene. When the β-geo and the host genes are not in the same reading frame, the fusion protein does not code for β-geo. When the β-geo vector is integrated before the host gene initiation codon, the β-geo sequence is not translated. This significantly lowers the number of exploitable ES cell clones (Figure 2.34).

An initiation codon may be added to a β-geo sequence. This enhances the chance of the fusion mRNA being translated into a β-geo protein.

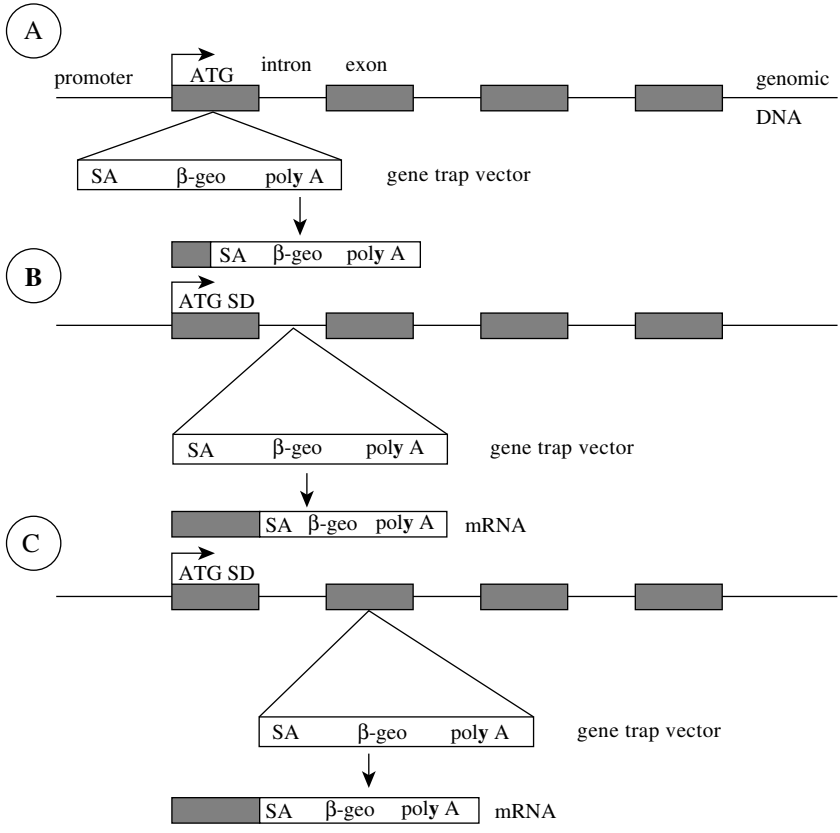


Figure 2.34 Different integration of the gene trap vector in a functional gene. The splicing acceptor site **(SA)** allows the marker gene to be expressed when integrated into an intron or an exon. The β-geo marker gene makes a visualization of its β-galactosidase activity and a selection by its neo^r activity possible. The gene trapping is performed in ES cells that are further used to generate chimaeric animals

2.3.10.2 Vectors to trap promoters and enhancers

In many cases, it may be important to trap a gene promoter or an enhancer rather than the gene itself. Vectors containing the β-geo sequence with their own initiation codon and terminator but devoid of a splicing acceptor site may be used. This vector may be expressed even when it is integrated before the host gene initiation codon (Figure 2.35(a)).

The vector will trap an enhancer if it contains a minimum promoter followed by the complete β-geo cDNA and a transcription terminator (Figure 2.35(b)).

2.3.10.3 Vectors to trap genes not expressed in ES cells

The vectors depicted above have been designed to work in ES cells selected by Geneticin. Although many genes are expressed in ES cells, it may

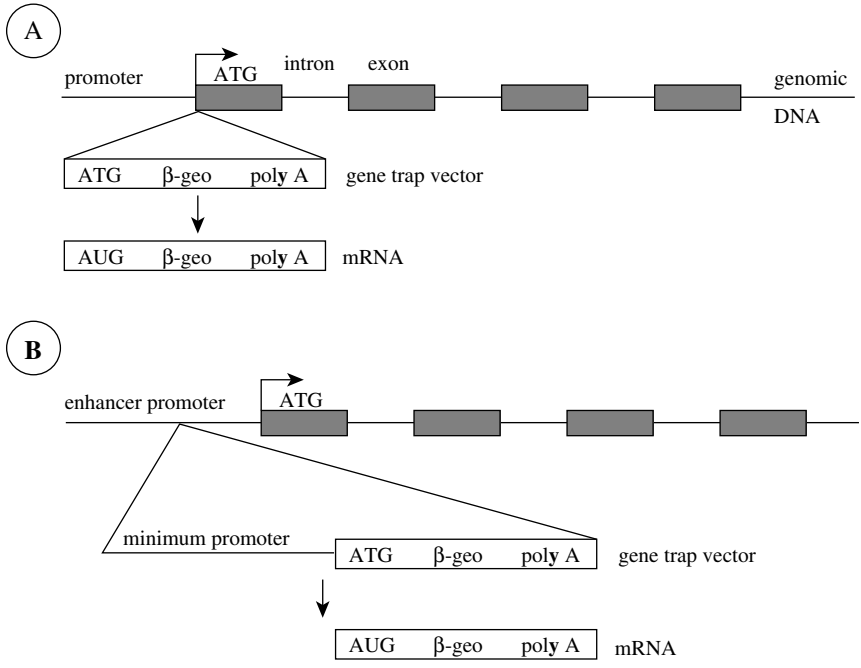


Figure 2.35 Promoter trapping. This vector is similar to the one described in Figure 2.34 but the β-geo gene contains its own ATG initiation codon. The reporter gene is therefore active when integrated between the transcription initiation site and the ATG of the gene

happen that the vector is integrated into a gene which is not active in ES cells. These cells are lost during selection although they might express the reporter gene later in development. To circumvent this problem, the reporter gene may be dependent on a promoter active in all cell types, including ES cells. The PGK (phosphoglycerate kinase) gene is generally used for this purpose. To select only the clones in which the vector was integrated within a gene, the β-geo sequence is followed by a splicing donor site but not by a transcription terminator (Figure 2.36). The presence of the PGK gene promoter makes β-geo gene expression and selection of the ES cells possible, even if the interrupted gene was not active in ES cells.

A more sophisticated version of this kind of vector is described in Figure 2.37. This vector contains two selection genes. The first is β-geo, containing its own initiation codon, a splicing acceptor site, an IRES, which allows the β-geo initiation codon to be used irrespectively of the presence of another initiation codon upstream and a transcription terminator. The second selection gene may be the puromycin-resistant gene preceded by a promoter active in ES cells and followed by a splicing donor site.

The second selection gene allows ES cells to be selected irrespectively of the expression of the interrupted gene. The first selection gene is expressed only in the cells in which the promoter of the selected gene is active. The expression of the first selection gene can be visualized by the β-galactosidase at any time during the life of the animal.

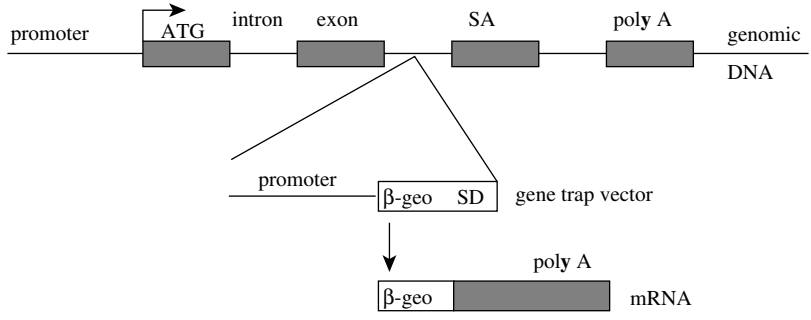


Figure 2.36 Non-expressed gene trapping. This vector contains it own promoter that is active in all cell types. The reporter gene is followed by a splicing donor site (SD), which makes it possible to use the transcription terminator of the expressed non-targeted gene

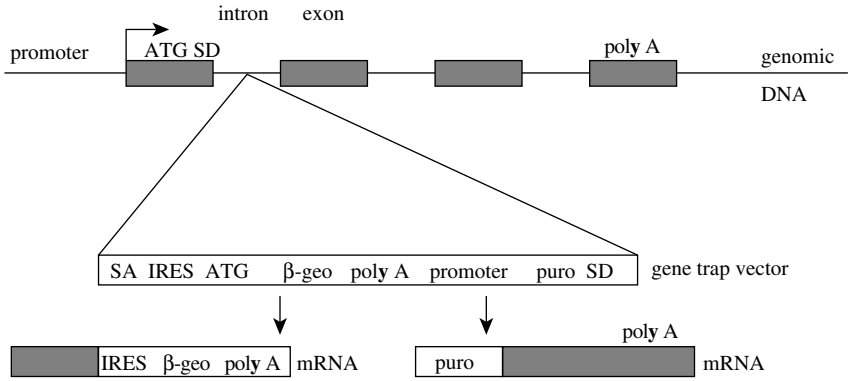


Figure 2.37 Trapping of expressed or non-expressed genes. This vector contains the β-geo reporter gene preceded by a splicing acceptor site (SA), an IRES sequence making it possible to use its own ATG initiation codon and a transcription terminator. The vector also contains a selection gene (puro), making it possible to trap non-expressed genes

2.3.10.4 Vectors to trap genes coding for secretory proteins

A more focused selection of trapped genes may be required in some cases. It may be interesting, for example, to selectively trap genes coding for secreted proteins. This may be achieved by using the reporter β-geo gene preceded by a sequence coding for a hydrophobic transmembrane (TM) domain capable of anchoring the protein in the Golgi apparatus membrane. The vector also contains a splicing acceptor site and a transcription terminator (Figure 2.38).

When the vector is integrated into a gene coding for a non-secretory protein, the β-geo protein is anchored by its TM domain in the membrane of the endoplasmic reticulum. The enzymatic sites which are in the lumen of the endoplasmic reticulum are not accessible to the substrates and the cells remain uncoloured. In contrast, when the interrupted gene codes for secretory protein, the β-geo protein is preceded by the signal peptide of the host gene. This drives the fusion protein in the endoplasmic reticulum in such a way as to anchor it by the TM domain, leaving the enzymatic sites outside the lumen and accessible for the substrate.

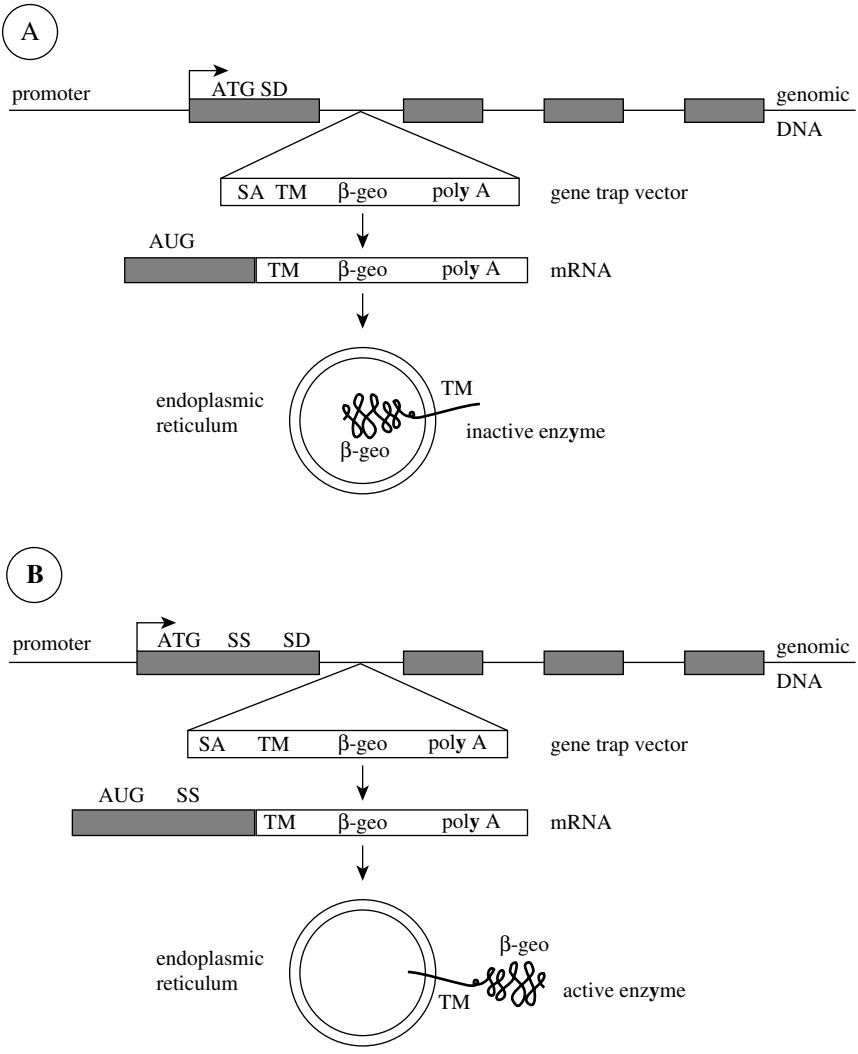


Figure 2.38 Trapping of genes coding for secretory proteins.
(A) This vector contains a TM sequence allowing the β-geo protein to be anchored in the endoplasmic reticulum membrane. The enzymatic region of β-galactosidase is then inside the endoplasmic reticulum and not accessible by its substrate
(B) This vector is integrated into a gene coding for a secretory protein. Its signal peptide **(SS)** makes it possible for the β-geo protein to be outside the lumen of the endoplasmic reticulum. The substrate can reach the enzyme and cells appear to be blue

2.3.10.5 *Vectors to trap transiently expressed genes*

A number of genes are expressed transiently and only in a given cell type. In this case, the experimenter has little chance of trapping this kind of gene using the vectors described above. A more sophisticated system based on the use of two vectors has been designed (Figure 2.39). The vector trap proper contains the Cre-recombinase cDNA, preceded by a splicing acceptor site and followed by a transcription terminator. This integrated vector may express its cDNA even with a transiently active promoter. The transient presence of the Cre-recombinase is sufficient to withdraw an inhibitory sequence of the reporter β-geo gene preceded by a promoter active in all cell types. The reporter β-geo gene is restricted to the cells in which the Cre-recombinase gene was expressed at an unknown stage of the life of the animal. The Cre-recombinase gene may be used as a hook to identify the gene in question.

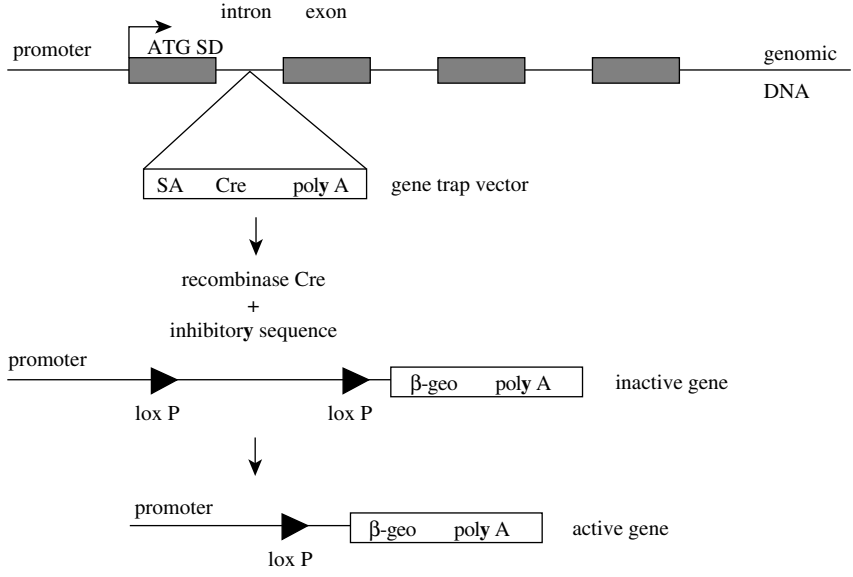


Figure 2.39 Trapping of transiently expressed genes. These vectors contain the Cre recombinase gene. When the targeted gene is active, the expressed Cre recombinase induces the irreversible expression of the reporter gene by the elimination of an inhibitory sequence. The expression of the β-geo gene indicates that the gene it has inserted was active, at least transiently

2.3.11 Vectors for the expression of transgenes

The first experiments, performed more than 20 years, ago revealed that transgenes were often expressed in a poorly predictable manner. One fact stood out in particular. Gene constructs perfectly active in cultured cells were silent in transgenic mice. A limited number of studies have been performed to explain this phenomenon and to tentatively improve expression vectors. This is obviously due to the particularly great complexity of the problem. Indeed, it is obvious that many signals are present in a gene and that their natural order is upset by gene construction. Many combinations of these signals are possible and transgenic animals are required each time to evaluate their effect. The first systematic study clearly showed that the chance of rapidly finding the optimal conditions to prepare a gene construct for transgenesis was low (Palmiter *et al.*, 1991; Peticlerc *et al.*, 1995). On the other hand, the generation of transgenic mice has become a standardized and relatively efficient technique. This allows researchers to obtain these animals with a moderate time investment. It is acknowledged that a gene construct may be too badly expressed in transgenic mice, thus preventing any exploitation of these animals. The general rule is also that several lines of transgenic mice (five to 10) are required to have a reasonable chance of obtaining a few lines expressing the transgene in an appropriate manner.

After almost two decades, a few rules for transgene expression have emerged, as a result of many observations.

The expression of a transgene is frequently not strictly dependent on its promoter. It is recognised that this is due to a chromatin-position effect. The presence of endogenous enhancers in the vicinity of the transgene is responsible for its leaky expression. The random integration of the foreign gene after microinjection into the embryo leads to the generation of lines, each showing a different position effect.

The fact that transgenes often remain silent was similarly attributed to the presence of silencers in their vicinity. The leaky expression and the silencing of the transgenes progressively appeared as not being really symmetrical. Indeed, transgene silencing has a more frequent and more potent effect than leaky expression. It appears, in fact, that a conventional transgene is poorly expressed if (i) cDNA is used rather than its genomic counterpart in the transcribed region (Palmiter *et al.*, 1991), (ii) the foreign gene is integrated in multiple copies (Whitelaw and Martin, 2001) or (iii) if the cDNA is of bacterial origin and, in fact,

rich in CpG motives (Cohen-Tannoudji et al., 2000a, 2000b). The last phenomenon has been directly demonstrated with the study of the bacterial lacZ gene. This gene is extensively used as a reporter in transgenic animals. The addition of a substrate gives off a blue colour in the cells expressing the lacZ gene. The use of this reporter gene is limited by its relatively high propensity to become silent during the course of animal development. This effect has been greatly attenuated by the systematic suppression of the CpG motives in the gene (Henry *et al.*, 1999). These observations strongly suggest that transgenes are often extinguished by an active mechanism, which might also be the one responsible for transposon and retroviral genome silencing.

A certain number of rules have been defined to limit leaky expression and the silencing of transgenes. These rules have been described in detail in a book in press (Houdebine *et al.*, 2002).

For the sake of simplicity, a gene and a transgene can be divided into three domains although the action of each domain is not independent of the other domains.

2.3.11.1 The promoter–enhancer region

Scientific literature contains a huge amount of information describing the regulatory regions of many different genes. A minimum promoter comprised in about 200 bp upstream of the cap site may be added to a gene construct. The presence of enhancers is highly recommended to obtain significant expression of the transgenes. In some cases, the major enhancers of a gene have been described and can be added to the vector. They are usually found in a few kb upstream of the cap site. In the event that the enhancers have not been identified, the use of a 5–10 kb fragment upstream of the cap site is recommended.

Promoters expressed in all cell types are needed for some experiments. The combination of β-actin gene promoter followed by its first intron and preceded by the CMV (cytomegalovirus) early gene enhancer seems to be the best tool for expressing foreign genes in all mouse cells at this time. The viral promoters from SV40 and CMV are weak *in vivo* although they are strong in cultured cells. The promoter from several housekeeping genes can also be used. PGK gene promoter is weak whereas EF1α gene promoter is very strong. Both are CpG rich and quite sensitive to silencing (Taboit-Dameron *et al.*, 1999).

It is well established that the action of promoters and enhancers is mediated by transcription factors which bind to specific DNA sequences. An increasing number of factors are being discovered but much remains to be done before their interactions are completely understood. The present knowledge of these mechanisms, although limited, led to the testing of combinations of regulatory sequences to tentatively create artificial promoters. Ideally, such promoters could be potent, specific and compact. This would facilitate their use. Multiple combinations have been tested and they have often proven to be disappointing. Indeed, the association of potent enhancers may not lead to high expression of the linked gene. The present level of knowledge about action mechanisms of promoters and enhancers is inadequate for the construction of efficient artificial regulatory elements on a rational basis. However, it is interesting to observe that multiple and systematic combinations of several DNA sequences involved in gene expression in muscle showed that some combinations gave very potent, quite specific and compact regulatory elements for expression of foreign genes in muscular cells (Somia *et al.*, 1999; Li *et al.*, 1999).

2.3.11.2 Boundary regions

Genes or gene clusters are bordered by specific DNA sequences which insulate genes from their neighbours. These regions are known as insulators (Bell, Welt and Felsenfeld, 2001). Their actions seem multiple and complementary (West, Gaszner and Felsenfeld, 2002). They prevent the action of a neighbour enhancer in a directional manner. This allows a gene or genes in a cluster to be efficiently and specifically expressed. Insulators also prevent the formation of silent heterochromatin. It is now clear that chromatin is opened by a local acetylation of histones which prevents DNA methylation. On the contrary, in heterochromatin, histones are deacetylated and methylated and DNA is methylated. These mechanisms are known as histone code. Some of the boundaries are long and complex and contain enhancers. Such regions, known as LCRs (locus control regions), contain insulators. One of them, found in the chicken β-globin locus, has been well characterized (Bell, Welt and Felsenfeld, 2001). The enhancers present in LCRs are supposed to stimulate expression of the locus genes. The formation of loops is then required to bring enhancers into the vicinity of promoters. An interesting hypothesis has been formulated recently. Transcription factors as well as

RNA polymerase II have been found in the LCR of the β-globin locus although a promoter is not present. It has been postulated that the LCR might be a reservoir concentrating some of the proteins required for efficient transcription (Johnson *et al.*, 2001a).

High transgene expression can be elicited by long genomic DNA fragments (Li, Harju and Peterson, 1999). Alternatively, predetermined elements can be used to favour transgene expression. The 5′HS4 region from the chicken β-globin locus, which contains an insulator, has been shown to prevent silencing of transgenes driven by several promoters having quite different specificities. The EF1 α gene promoter, which is quite sensitive to silencing, remained active when it was associated with two copies of the 5′HS4 region (Taboit-Dameron *et al.*, 1999). However, this insulator was not able to prevent the completely variegated expression of transgenes. Other elements present in the β-globin LCR must be required to make a perfect expression of transgenes possible.

The number of identified LCRs and insulators is increasing and their functional elements are being identified and described. It is likely that compact insulators will have been delineated and added to expression vectors in a few years. Boundary regions seem to play a major role in cell differentiation at the chromatin level. Indeed, they contribute to open chromatin regions, which contain genes to be expressed in a given cell type. This means that specific boundary regions may be required to efficiently express a transgene in a given cell type.

2.3.11.3 The transcribed region

An increasing number of signals are being found in mRNAs. They control pre-mRNA maturation, mRNA transfer to cytoplasm, mRNA half-life and translation etc. (Jacobs *et al.*, 2002).

Introns It was observed several years ago that the addition of at least one intron to a cDNA is essential for the expression of the transgenes (Palmiter *et al.*, 1991). Many of the first introns in genes contain sites for the binding of transcription factors. This makes it possible for chromatin to keep an open form on both sides of the cap site. The addition of enhancer introns may quite significantly enhance transgene expression (Petitclerc *et al.*, 1995).

Another possibility consists of using minigenes composed of the whole coding sequence but of some of the introns only. The choice of the

introns to keep in minigenes can be determined only on a case by case basis

Exon splicing is mediated by spliceosomes, complexes formed by proteins and small RNAs. After splicing, some of the proteins of the spliceosome remain bound to the mRNA. This complex is essential for the transfer of the mRNA to cytoplasm (Keys and Green, 2001). This may explain why at least one intron is required for a cDNA to be expressed in transgenic mice.

Various combinations of introns have been tested. One conclusion is that it is preferable to add introns before the cDNA rather than after. A mechanism that controls the quality of mRNA and is known as NMD (nonsense mediated decay) may explain why introns added after a cDNA prevent its expression. A terminator codon must not be farther than 50 nucleotides from the following intron. If this distance is greater, the mRNA is considered to be abnormal and destroyed. Gene construction may follow this rule but the task is a bit more complicated in this case. This NMD mechanism may explain why some constructs that have an intron after the cDNA are not efficiently expressed in transgenic animals.

5′UTR The ribosome scanning mechanism is slowed or prevented by GC rich 5′UTR that forms stable secondary structures. The generation of such structures should be avoided during transgene construction.

An initiation codon, essentially AUG, is used in animal cells primarily when it is located within the consensus sequence, GCC A/GCC AUGG. The essential nucleotides are a purine after the AUG codon and another purine (most frequently, A at position 3) before the AUG codon. The consensus sequence may be optimized by mutation, if necessary.

5′UTR often contains signals for the control of translation. These signals may be added to or withdrawn from the gene construct according to the expected expression level of the transgene.

Secretory proteins contain a signal peptide formed by 15–30 hydrophobic amino acids. The proteins not naturally secreted may be secreted when the corresponding DNA sequence is added before their coding sequences.

3′UTR 3′UTR may contain signals targeting the mRNA in a given cell compartment. These signals may or may not be added to the gene constructs, as required.

Signals controlling mRNA half-life have been found in the 3′UTR of numerous mRNAs. AU rich 3′UTR containing AUUUA motives desta-

bilizes mRNA. In contrast, CU rich regions are mRNA stabilizers. The presence of these sequences may greatly modify mRNA half-life and gene constructs must be examined to see whether or not they contain these signals.

A certain number of proteins are anchored in the plasma membrane by a GPI structure (glycophosphatidylinositol). This structure is covalently bound to the terminal region of the protein. This mechanism is triggered by a specific aminoacid sequence which may be added to a protein not naturally anchored to plasma membrane.

Codon usage It is well known that some codons are preferentially used in given species or organs. Optimization of the codons by mutation may greatly enhance (up to 100-fold) the expression of a foreign gene in animals. The modification of the codons may also eliminate or generate sequences that modify mRNA transport to the cytoplasm, mRNA half-life or translation.

The simultaneous expression of several cistrons In a number of cases, several coding sequences have to be expressed simultaneously after gene transfer. This is the case when cellular clones stably expressing a foreign gene have to be established. The cotransfection of a selection gene that may cointegrate with the gene of interest is generally needed. Some proteins are formed by several subunits which have to be present in the same cell to be assembled in an appropriate manner. Antibodies formed of a minimum of a light and a heavy chain are a case in point.

The simultaneous expression of two or more cistrons can be achieved by different approaches. The simplest way is to prepare separate gene constructs and cotransfect them into cells or coinject them into embryos. The cointegration that occurs in about 70 per cent of the cases frequently leads to the simultaneous expression of the two genes. However, cointegration is often accompanied by rearrangements, which somewhat inactivate the genes. The coinjection of two gene constructs into embryos makes it necessary to eliminate the transgenic animals that harbour only one of the constructs or to cross them to cumulate both transgenes. This is time consuming and not very compatible with the use of large animals, which reproduce at a slow rate.

Another possibility consists of introducing both gene constructs into a single plasmid. The two constructs are then stoechiometrically integrated and expressed accordingly. This approach implies that complex gene

constructs are being prepared. This may be particularly laborious if long regulatory genomic DNA fragments have to be used.

The cointegration of two genes at the same site may lead to poor expression of the transgenes. A systematic study carried out in cells revealed that two similar gene constructs integrated into a given site of a genome are expressed, to some degree, as a function of their relative position. When the transgenes are in tandem, the first is expressed, but at a reduced rate, and the second is much less active. This partial silencing is suppressed when the promoter of one of the two genes is inactivated. When the two genes are linked in a divergent orientation, each of them works about one-fifth as well as each gene alone. When the two genes are in a convergent orientation, their expression is severely reduced and does not exceed 10 per cent of their maximum potency. This phenomenon, called transcription interference, must have a considerable impact in transgene expression when multiple copies are integrated. This phenomenon must be taken into account when several cointegrated genes have to be expressed. It is not known whether transcription interference may be attenuated by the addition of some DNA sequences in the vectors (Eszterhas *et al*., 2002).

Research carried out for more than a decade revealed that two cistrons added to a single vector and expressed as a single mRNA are poorly expressed. The first cistron is translated but generally not the second.

A certain number of sequences present in the 5′ UTR of different mRNAs made it possible to express the second cistron in bicistronic mRNA when added between the first and the second cistron. This kind of stimulatory sequence was originally found in various viral genomes and especially in those in which the mRNAs are translated despite the absence of a cap. These observations suggested that ribosomes might be recruited directly by 5′ UTR without a scanning process from the cap to the initiation codon. This hypothesis was supported by some experimental data. The 5′ UTR of the mRNAs devoid of a cap added between both cistrons of bicistronic mRNA enabled the second cistron to be translated. This type of 5′ UTR is known as an IRES (internal ribosome entry site).

IRESs sharing these properties have been found in a number of viral and cellular mRNAs. Their action mechanisms have been questioned (Houdebine and Attal, 1999; Kozak, 2001b). Indeed, several experimental data are not compatible with the notion of an IRES. The capacity of many IRESs to stimulate translation is dependent on the physiological state of cells. IRESs often operate when translation is slowed down for some reason in a cell. It was therefore suggested that IRESs are specific

translation stimulators rather than elements capable of directly recruiting ribosomes (Houdebine and Attal, 1999).

Regardless of the mechanism that is involved in IRES action, these elements make possible the expression of the second cistron in biscistronic mRNAs and even of three cistrons if two IRESs are added to the vector.

Gene constructions containing two cistrons can be easily prepared. Many researchers have been disappointed after using such constructs. A certain number of rules must be followed to obtain an optimized expression of bicistronic mRNAs. The IRESs have quite different properties. The well known polyoma IRES is only fully active in infected cells. The IRES from the encephalomyocarditis virus (EMCV) is potent and active in all cell types, regardless of the physiological state of the cells. This property is shared by other IRESs. The second cistron added after the EMCV IRES must be in a precise position. The initiation codon of the cistron must be in the position of the 13th AUG of the IRES. This complicates gene construction. Most of the IRESs do not share this property. The terminator codon of the first cistron must be at distance corresponding to about 80 nucleotides from the following IRES to allow for an optimum translation of the second cistron (Houdebine and Attal, 1999).

One of the major problems of IRESs is that not just the second cistron of bicistronic mRNAs, but sometimes the first as well, are often translated at a low rate in comparison with similar monocistronic mRNAs. Hence, IRESs are quite useful elements for expressing several cistrons from single mRNAs but only if moderate expression levels are needed.

2.3.11.4 Vectors coding for RNA inhibiting the expression of specific genes

It is as important to inhibit the expression of a gene as to express a foreign gene in a cell or a living organism. Gene knock-out depicted above is one of the methods used to inhibit the expression of a specific gene. This approach has several drawbacks. It is laborious, irreversible and not easily possible in a single-cell type. Ideally, a cellular gene should be inhibited in a controlled manner in a given cell type but only when the experimenter decides to do so. Several techniques that can theoretically be applied to transgenic animals are available or presently being studied. They specifically allow for the inhibition of a gene at DNA, mRNA or protein levels. Natural mechanisms involve RNAs to control gene expression at the transcription or translation level (Altuvia and Wagner, 2000).

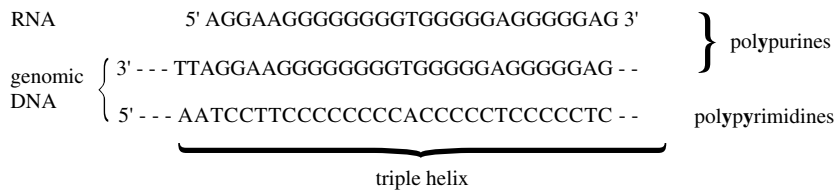


Figure 2.40 Targeted inhibition of a gene by the formation of a triple helix. A pyrimidine rich region of a gene can form a stable triple helix with a purine rich RNA oriented as described in the figure. The triple helix blocks transcription by RNA polymerase II. The RNA synthesis may be directed by a transgene

RNA forming a triple helix with DNA Single-strand DNA or RNA can form a triple helix with DNA in pyrimidine-rich regions. This property is being exploited to inhibit specific gene expression by DNA oligonucleotides (see Section 2.2).

It is conceivable to prepare gene constructs coding for RNAs capable of forming a triple helix with DNA in transgenic animals (Figure 2.40). This technique was successful for inhibiting IGF1 and IGF1 receptor genes in cultured tumoural glial cells. This inhibition attenuated the tumorigenic properties of the cells (Upegui-Gonzales *et al.*, 2000). The gene construct that inhibited IGF1 gene expression in cultured cells was used to generate transgenic mice. Unexpectedly, no transgenic animals were obtained, suggesting that the transgene was lethal. This is somewhat surprising since IGF1 gene knock-out is not lethal in mice. One possible explanation is that the RNA expected to form a triple helix with the IGF1 gene also acted on another gene that plays a key role in mouse embryo development (Houdebine *et al.*, unpublished data).

Although it is an interesting option, the blocking of gene expression by the formation of a triple helix between an RNA and DNA has not yet been successful in transgenic animals.

Single-strand antisense RNA The formation of an RNA double helix between an RNA and a complementary antisense RNA is a natural mechanism to control the activity of targeted cellular mRNA. Antisense RNAs are thus believed to inhibit targeted mRNA translation after formation of a double helix with two complementary RNAs.

Although this method can be implemented with no particular difficulty, it is marginally used at most. This is clearly due to its low and unpredictable level of efficiency. It is generally acknowledged that

the poor reliability of this method is essentially due to the fact that the targeted mRNA and the antisense RNA do not associate to form a double helix. The targeted RNA and the antisense RNA both have a good chance of being structured in a double helix in multiple regions and of being covered by cellular proteins.

To enhance the probability of forming a sense–antisense hybrid, it has been suggested that the regions of the targeted and antisense RNAs that are naturally in a single-strand structure be identified. These regions have the best chance of recognizing and forming a hybrid. Similarly, vectors expressing relatively short antisense RNA having no complementary sequences have been designed. Such a vector is depicted in Figure 2.41. The sequence coding for the short antisense RNA is then inserted into a loop of an RNA transcribed by RNA polymerase III (tRNA, 5S RNA, VA1 or VA2 RNA from adenovirus, etc.). These genes are transcribed at a high rate. Their products are highly structured and stable RNAs, which leave the best chance for the antisense to remain in a single-strand structure.

Despite this potential improvement, the strategy based on the use of single-strand antisense RNA has not been retained to inhibit specific gene expression in transgenic animals.

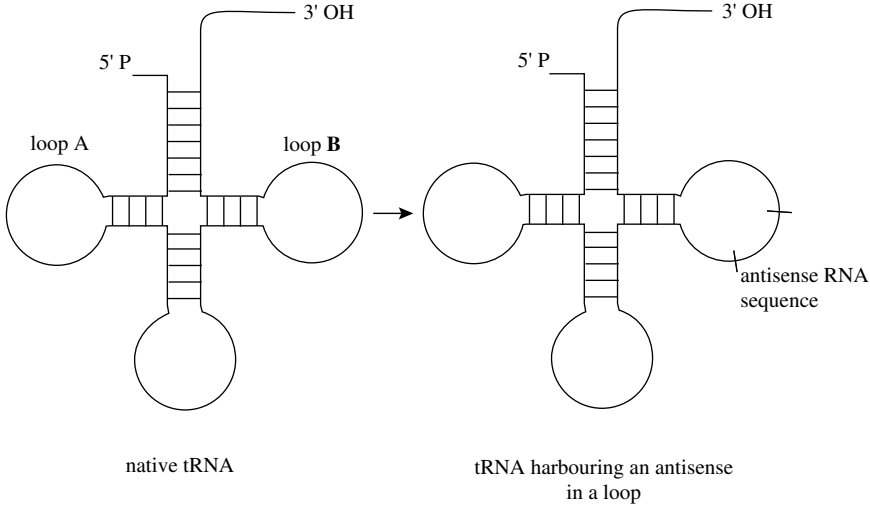


Figure 2.41 Use of vectors transcribed by RNA polymerase III to express antisense monostrand RNA complementary to mRNA. The sequence coding for the antisense RNA is introduced into a loop of the RNA. This makes a high level of transcription of the transgene possible as well as good accessibility of the targeted gene by the antisense RNA

Ribozymes Some natural RNAs have an RNAse activity, which degrades targeted RNAs. These RNAs, known as ribozymes, have different basic structures. Some ribozymes form a hairpin-like structure whereas others are reminiscent of hammerheads (Bramlage, Luzi and Eckstein, 1998).

Both types of ribozyme are formed by two domains. One contains the catalytic activity and the other has sequences complementary to the targeted RNA. A double-strand structure is thus formed between the targeted RNA and the ribozyme. This ensures the specific degradation of an RNA by a given ribozyme. The cleavage of the targeted RNA by a ribozyme occurs only at specific consensus sites. The hammerhead ribozymes cut a targeted RNA, preferably after the C of a GUC motif (Figure 2.42). Other motives are the target of hammerhead ribozymes. Hammerhead ribozymes can cleave RNA after the sequence UX, where X can be A, C or U but not G.

Multiple sites are present in a given RNA although the specific degradation of a cellular RNA by a ribozyme is not easily achieved. It is acknowledged that the reasons for this poor efficiency are the same as those that reduce the antisense RNA action. The two partners have little chance of meeting and forming a hybrid. A systematic and modelized study of RNA structure may indicate the regions that are in a single strand. This may be confirmed by the action of some chemical compounds that are known to form covalent links with single-strand

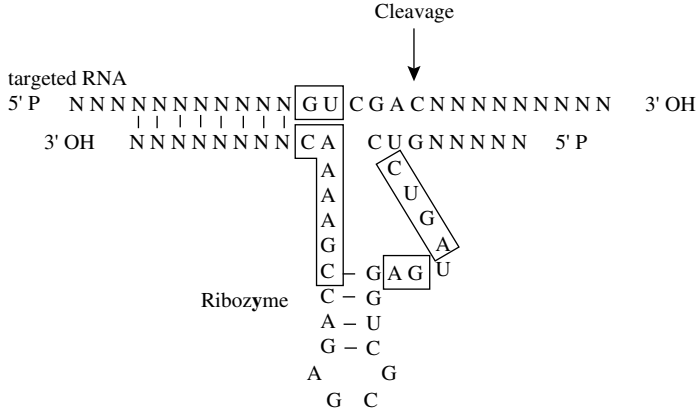


Figure 2.42 Structure of hammerhead ribozymes. The ribozyme active site is in the boxed regions. The targeted RNA is cleaved at the site indicated by the arrow. The GUC codon as well as other codons are ribozyme targets. The targets of the ribozyme are NUX where N is any base and X is any base but G

RNA. In this way, RNA regions that are probably the most accessible to ribozymes are determined and ribozymes can be constructed accordingly.

A recent study revealed that the efficiency of a ribozyme may be greatly enhanced if the sequence coding for the ribozyme is linked to the coding sequence of a helicase. The enzyme opens the double-strand RNA and thus increases the chance of the ribozyme reaching its target (Warashina et al., 2001). It is interesting to note that a poly-A stretch added to the hammerhead ribozyme was able to bind PABP, PAIP and eIF4A. The last factor has a helicase activity and it considerably enhances ribozyme action (Kawasaki and Taira, 2002).

Double-strand RNA Experiments carried out in plants revealed that the antisense RNA strategy is relatively efficient. This could not be explained by the fact that the targeted RNAs were in a single strand or not covered by proteins. The reasons for the high efficiency remained unknown for years.

Another unexpected observation was made on transgenic plants. A transgene may frequently become extinguished over a period of time. More surprisingly, some transgenes inhibited the expression of endogenous genes and of other transgenes as soon as they shared common sequences.

Another surprising fact was that double-strand antisense RNA was much more potent for inhibiting the specific expression of a gene than the corresponding single-strand antisense RNA.

Integrated transgenes are sometimes transcribed in both directions. This is due to the fact that transcription terminators in plants are weak. A transgene may be transcribed in the sense direction by its promoter and in the antisense direction by the promoter of an endogenous gene located in its vicinity. These two RNAs may form a double-strand RNA, which leads to the inhibition of the corresponding gene.

The mechanisms that explain these phenomena have been deciphered in *Caenorhabditis elegans*. A double-strand RNA longer than 300 bp does not exist in natural RNAs. Such abnormal structures induce an RNAse, which cleaves the double strand in multiple and overlapping double-strand RNA fragments containing 21–23 base pairs with two- to three-nucleotide 3 overhangs. These RNAs may be amplified by an RNA polymerase. They recognize the RNA having the same primary structure and this induces a specific and potent destruction of the targeted RNA (Figure 2.43).

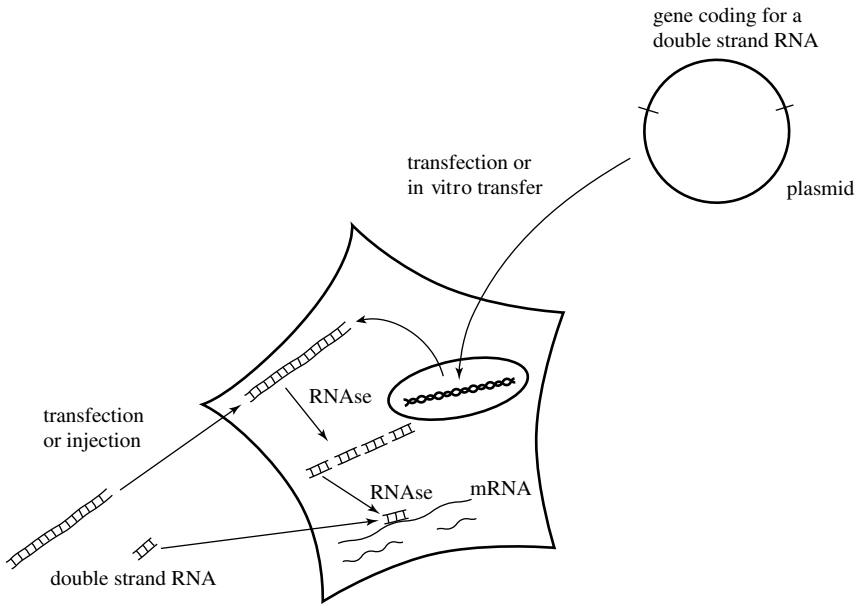


Figure 2.43 Mechanism of RNA interference leading to a specific degradation of the complementary RNA. A double-strand RNA is introduced into the cell or synthesized from a foreign gene. This RNA is cleaved into 21–23 bp fragments that recognize homologous sequences in a cellular RNA. This recognition induces an RNAse, which cleaves the targeted RNA

This mechanism has been given the name of RNA interference (RNAi) (Voinnet, 2001). RNAi related mechanisms also inhibit gene expression at the transcription level. These mechanisms, known as TGS (transcriptional gene silencing), as opposed to those degrading RNA (post-transcriptional gene silencing: PTGS) are mediated by an adenosine deaminase, which mutates the targeted gene (Voinnet, 2001).

The RNAi mechanism has been found more recently in other animals. In *Drosophila*, the expression of a double-strand RNA inhibits the corresponding gene by degrading its mRNA. This phenomenon is reliable and the systematic inhibition of endogenous gene expression by transgenes coding for double-strand RNAs has been undertaken. Gene replacement by homologous recombination proved possible in *Drosophila* (Bernards and Hariharan, 2001). The RNAi approach appears simpler and more flexible, since the expression of the transgene coding for double-strand RNAs can be easily targeted in a given cell type and regulated by a specific promoter.

Still more impressive is the implementation of the RNAi mechanism to inhibit specific gene expression in *C. elegans*. Up to 5000 genes have been silenced in this species by plasmids coding for double-strand RNAs and given orally to the animal (Timmons and Fire, 1998). Alternatively, dsRNAs can be injected into embryos or the animals can be soaked in a dsRNA solution. A large proportion of these transgenic animals show phenotypic properties due to RNAi action. Hence, although it is unable to degrade all the targeted mRNAs, the RNAi mechanism in invertebrates may be considered as the equivalent of gene knock-out.

The RNAi mechanism has been observed in lower vertebrates and in mammals (Wianny and Zernicka-Goetz, 2000). However, specific silencing by double-strand RNA was clear in embryonic cells but not in differentiated cells. It was observed in different cell lines after gene transcription (Ui-Tei *et al.*, 2000). However, higher concentrations of double-strand RNA are required to induce a degradation of the corresponding mRNA in CHO cells as opposed to *Drosophila* cells.

Repeated experiments carried out with various mammalian cells cast doubts on the idea that RNAi is an essential mechanism in these species. In fact, the results were not easily interpretable in most species. Indeed, long double-strand RNAs are known to induce interferon synthesis, an inhibition of protein synthesis and mRNA degradation in mammalian cells. This defence mechanism hides a possible RNAi effect.

Experiments performed by several groups have shown that synthetic 21–22 bp double-strand RNAs transfected into mammalian cells induce a specific degradation of the corresponding mRNA without interferon side-effects. (Elbashir *et al.*, 2001; Elbashir, Lendeckel and Tuschl, 2001; Bass, 2001; Billy *et al.*, 2001). Hence, the RNAi may be implemented in cultured cells to inhibit gene expression, provided that short double-strand RNAs are transfected. It is not clear whether the expression of transgenes coding for double-strand RNA can inhibit specific gene expression in mammals. This possibility remains to be explored. It is conceivable that medium-size double-strand RNA not exceeding 300 bp induces RNAi without triggering the interferon mechanism (Paddison *et al.*, 2002). Quite interestingly, several independent studies demonstrated that short double-strand RNAs synthesized in cells by vectors using promoters recognized by RNA polymerase III have a potent and specific RNA interference effect (Sui et al., 2002; Brummelkamp, Bernarde and Agami, 2002). It remains to determine in which conditions these vectors are active in transgenic animals. Another interesting fact is that one of the promoters used in these studies, the U6RNA gene promoter, can be

controlled by the tetracycline system (Section 2.3.11.6) (Ohkawa and Taira, 2000).

RNAi was discovered quite recently, although it now appears to be a universal mechanism possibly involved in the defence against viruses, at least in plants and invertebrates, where it is very potent. In addition and independently, RNAi might be a basic mechanism controlling gene expression at the mRNA level.

The study of the RNAi mechanism is making rapid progress. Several genes involved in this mechanism have been identified. It is interesting to observe that several genes play an essential role both in RNAi and in transposon inactivation. A better understanding of these mechanisms might contribute to the definition of the optimal conditions for implementing RNAi in mammals.

Transdominant negative proteins Natural proteins inhibit the action of homologous proteins. This is sometimes the case for proteins resulting from alternative use of initiation codons or alternative splicing. These mutants share some but not all the properties of the normal protein. Competition between the two forms of the proteins results in an inhibition of the active form. The mutant is then known as a transdominant negative.

A certain number of these mutants have been found in cells. Others can be generated by a systematic search for the mutants showing a transdominant negative effect. Transgenes coding for transdominant negative proteins inhibit the natural protein with good specificity provided that they are present in sufficient quantities. This concept is exemplified in Figure 2.44. An overexpression of an insulin receptor mutant in transgenic mice induced a type 1 diabetes.

This sophisticated approach is flexible and quite diverse. Its success relies essentially on the availability of the gene coding for appropriate transdominant negative proteins.

2.3.11.5 Genetic ablation

To evaluate the function of specific cell types in a given tissue or to mimic some human diseases, it is essential to induce the specific destruction of these cells. This can be achieved by specifically expressing genes coding for cellular toxins in these cells. Diphtheria toxin is currently being used for this purpose. This toxin ADP ribosylates an elongation factor, leading to a complete inhibition of protein synthesis and to cell death.

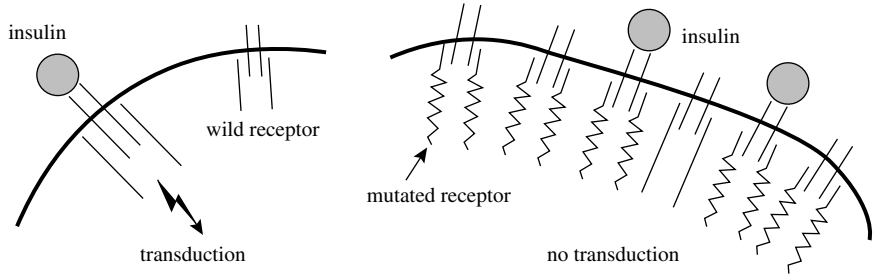


Figure 2.44 Specific inhibition of a gene by the overexpression of a transdominant negative protein. The insulin binds to the overexpressed mutated receptor, acting as a decoy. The insulin signal is no longer transduced in the cell and the transgenic mice are diabetic

The toxin is extremely potent and only a few molecules are sufficient to induce cell death.

In its classical version, this method implies that the toxin gene is driven by a cell-specific promoter. The toxin potency is so high that the cells are destroyed as soon as the promoter starts to be active during development or when a leaky expression of the transgenes occurs.

A more sophisticated version of the method offers higher flexibility. Rodents are devoid of diphtheria toxin receptors and are not sensitive to this molecule. The receptor can be expressed in specific cell types of transgenic mice using appropriate promoters. Injection of the toxin into the animals induces a rapid and specific destruction of the cells expressing the receptor gene (Saito *et al.*, 2001).

2.3.11.6 Vectors for the control of transgene expression by foreign inducers

The conventional vectors for transgene expression contain regulatory elements from animal cells. This implies that the transgenes are controlled by natural or synthetic inducers, which also regulate endogenous genes. An inducer may thus activate not only the transgene but many of the host genes as well. This may complicate the interpretation of the data, since it may be difficult to discriminate between the effects of the transgene and those of the endogenous genes.

It is therefore highly desirable to use promoters sensitive only to foreign inducers not active on the host genes. Several systems have been

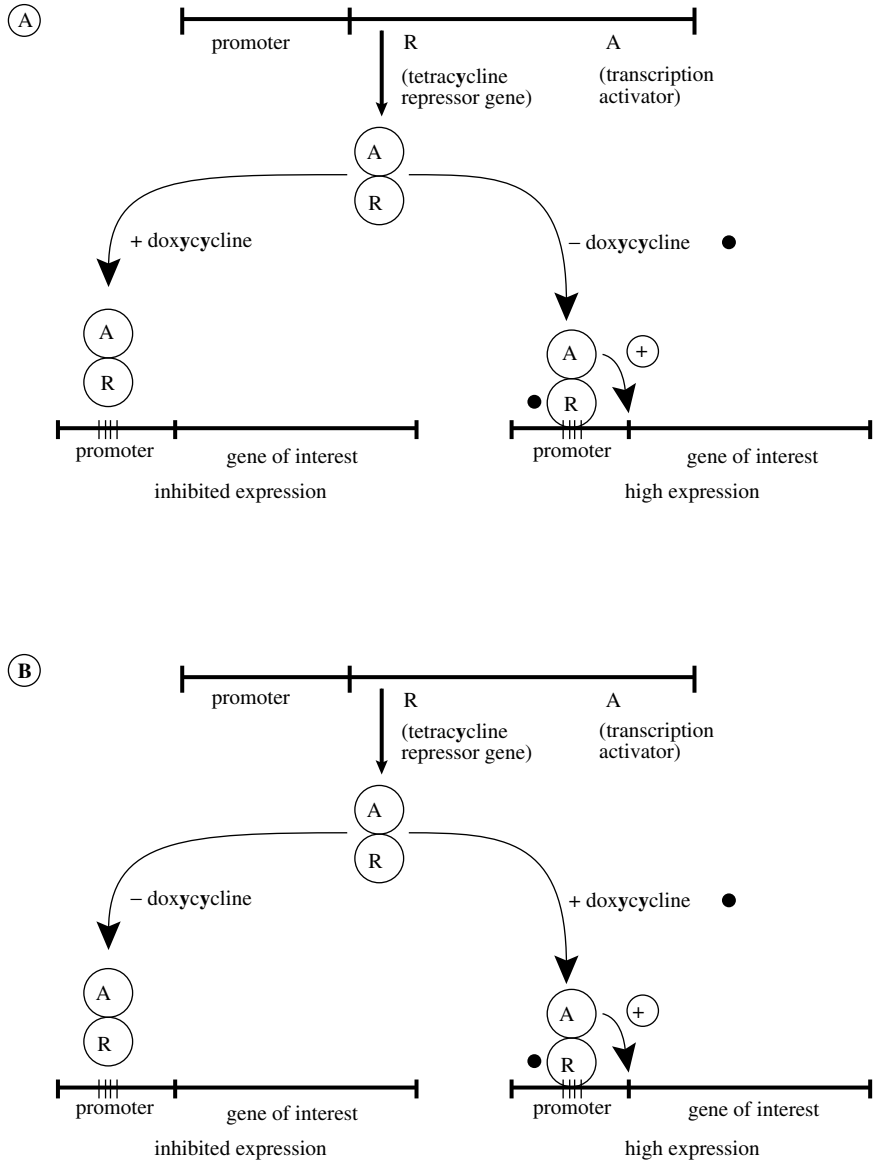


Figure 2.45 Controlled expression of transgenes by tetracycline.
(A) The transcription factor, AR, binds to the promoter, directing the expression of the gene of interest in the presence of doxycycline or a tetracycline analogue
(B) A mutant of R has the opposite action. In both cases, the gene enhancer is in A

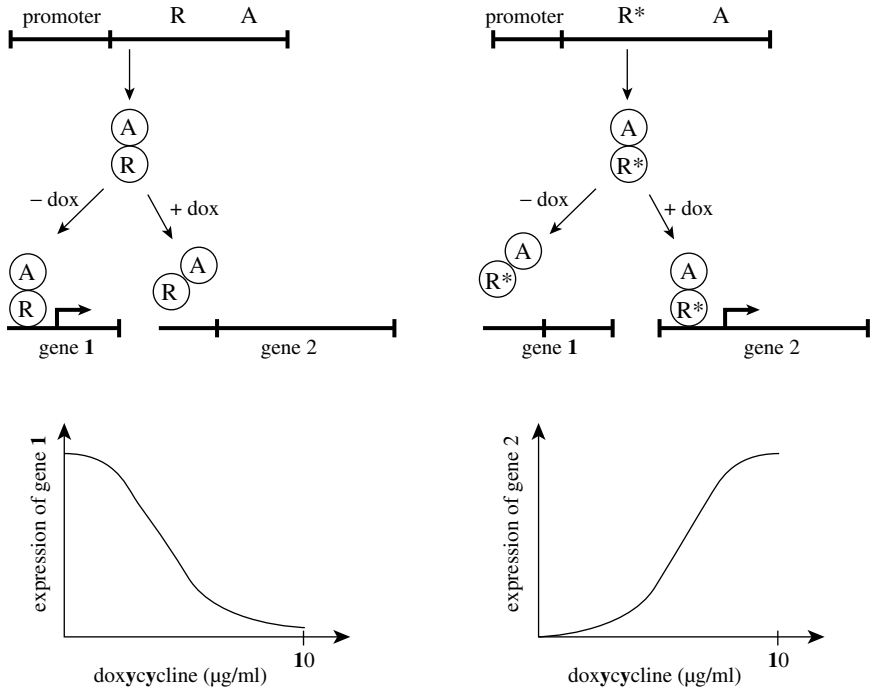


Figure 2.46 Simultaneous induction and inhibition of two genes by doxycycline. Two R mutants having opposite actions on the tetracycline-resistant gene are used simultaneously. Doxycycline induces one gene and inhibits the other

proposed and are currently used. In all cases, these systems are hybrids using both animal and quite distant mechanisms. The most popular system relies on the use of tetracycline (or derivatives such as doxycycline) as an inducer. This system is described in Figures 2.45–2.47.

Other systems have proved to be satisfactory although they are less frequently used. These systems use different inducers: rapamycin (Rivera *et al.*, 1996), ecdysone (No, Yao and Evans, 1996), RU 486 (Wang *et al.*, 1997) or streptogramin (Fusseneger *et al.*, 2000).

All these systems use hybrid transcription factors containing at least two domains, one capable of activating animal genes and the other activated by a foreign inducer, which may or may not bind to the transgene promoter.

The system depicted in Figure 2.45 uses doxycycline as an inducer. A first transgene codes a transcription factor containing an activator

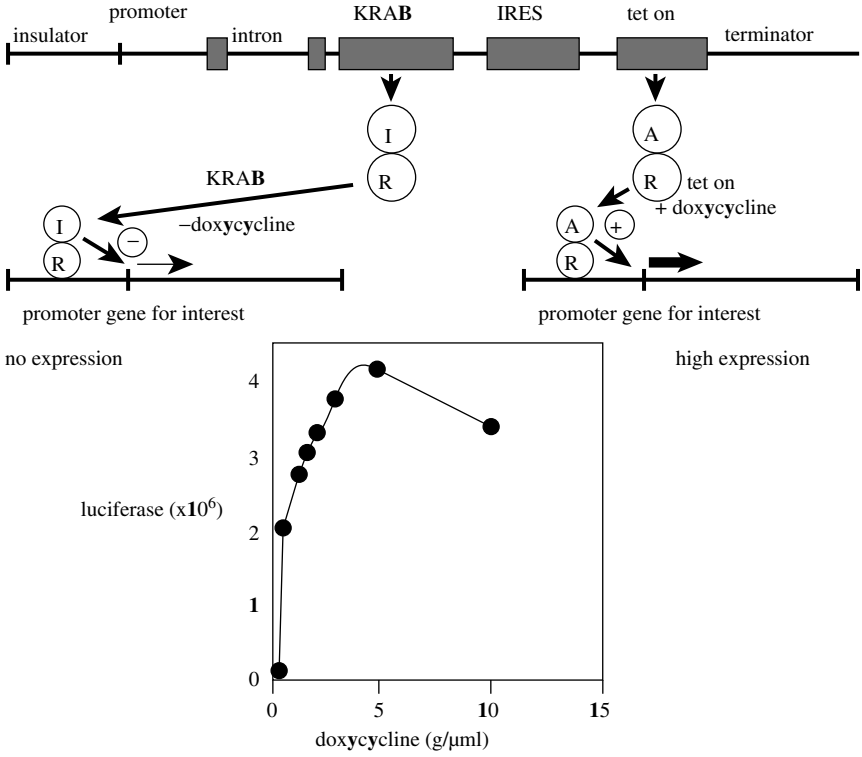


Figure 2.47 Use of an activator and a repressor to enhance the control of a foreign gene by doxycycline. The KRAB region is a silencer (I) and A is an enhancer. In the absence of doxycycline, the gene of interest is not stimulated and is therefore repressed. In the presence of doxycycline, the gene is stimulated and no longer repressed. This system reduces background expression

domain (the VP16 protein from *Herpes simplex* virus) and a region from the tetracycline repressor gene binding doxycycline. A second transgene harbours the gene of interest and its promoter contains the minimum CMV promoter and repeated DNA sequences binding the tetracycline repressor gene. Several mutants of the repressor have been generated. This allows doxycycline to be a transgene repressor (Figure 2.45(a)) or a transgene inducer (Figure 2.45(b)).

The two genes are generally transferred separately into mice, which are later crossed to harbour both transgenes. The effect of doxycycline can be obtained by injecting the drug into the animals or by adding it to their drinking water.

The simultaneous expression in the same mice of the two versions of the repressor–activator hybrid may make it possible for two target genes to be regulated in a reverse manner. Doxycycline then induces one of the target genes and inhibits the expression of the second in a dose-dependent manner.

The versions of the system depicted above and all the others suffer from a fundamental limitation. The regulatory region of a gene of interest contains a minimum promoter and the binding site for the hybrid transcription factor. In the absence of the inducer, the minimum promoter is active, leading to a background expression of the transgene of interest. Although this expression is low, it may invalidate the model.

To reduce the basal expression of the transgene of interest, systems based on the simultaneous use of an activator and a repressor have been designed. One of these systems is shown in Figure 2.47.

A first transgene contains two cistrons separated by an IRES. This transgene expresses both a transcription stimulator and a transcription silencer. In the absence of doxycycline, the activation is absent and the repressor is present in the regulatory region of the transgene, reducing its expression to a very low background. In the presence of the inducer, the opposite situation is obtained.

These systems have been described in many publications (Rossi and Blau, 1998; Forster *et al.*, 1999; Blau and Rossi, 1999; Mills, 2001).

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

3 Applications of Cloning and Transgenesis

3.1 Applications of Cloning

Cloning offers the unique opportunity to obtain totipotent cells from fully differentiated cells. Cloning is therefore a new and attractive experimental model for studying the mechanisms responsible for cell dedifferentiation and differentiation. Cloning is also a new method of animal and potentially human reproduction, producing individuals genetically identical to an adult. Potential applications have been developed with a view to better control animal and human reproduction. It is now acknowledged that transgenesis is in some cases facilitated by cloning. A number of diseases could probably be cured by cell therapy. This implies that stem cells are available to regenerate damaged tissues. Embryonic stem cells can be obtained by cloning and organ stem cells can be derived from embryonic stem cells. Hence, theoretically, cloning could contribute to cell therapy by autografting. In practice, these applications are still in their infancy and some of them may not become a reality due to technical or ethical problems.

3.1.1 Basic research

The birth of Dolly demonstrated that the genetic composition of differentiated cells is essentially the same as that of an embryo. This was generally considered to be the case but evidence has now confirmed that the genome of an adult cell can generate a living embryo.

Repeated although not yet systematic observations have shown that the genotype of the animals from which oocytes and nuclear donor cells

are obtained has a quite significant impact on the success of cloning. Ovarian follicles can be recovered repeatedly from a given animal by an ultrasound-guided follicle aspiration technique named OPU (ovum pick-up). Oocyte maturation can be achieved *in vitro* providing a number of recipients sharing the same mitochondrial genome. This approach reduces the cost of oocyte production and it allows the study of mitochondrial genome impact (Brüggerhoff *et al.*, 2002).

One of the major questions raised by cloning is that of knowing how the numerous genes that are silenced in differentiated cells can be reactivated to participate in embryo development. This process very rarely occurs properly as judged by the constant low efficiency of cloning. The reasons for the numerous failures are unknown. It is conceivable that a certain number of genes in differentiated cells are mutated and inactive. Embryo development may be impaired if essential genes not required for adult cell life are no longer functional. It is impossible to examine the genome structure of an individual cell to be used for cloning. The state of the genome could be studied in cells of embryos and foetuses which have failed to develop. This kind of experiment does not seem to have been performed so far. It could reveal the role of genes playing an essential role in embryo development. This task is particularly difficult since a maximum number of genome domains should be compared in cells from normally developing embryos and from those whose development has stopped.

One factor is striking. Most of the cloned animals unable to survive after birth show repeated syndromes, some of which, interestingly, are similar to certain human development diseases. Cloned animals could therefore become models for the study of these diseases. This suggests that the same group of genes is not functional in the clones. This fact is inconsistent with random gene mutations in adult cells used as nuclear donors. The idea that the major causes of cloning failure are due to epigenetic phenomena appears more likely.

Repeated observations have led to the conclusion that some cell types can be a source of nuclei for cloning whereas others always result in failure. Cumulus cells are considered as one of the most suitable nuclear donors. Nerve cells or ES cells have a much lower cloning efficiency. No convincing hypothesis has been proposed to explain these facts. Correlations between the number of divisions that cells were subjected to before being used as nuclear donors and the success rate of cloning have not been clearly established. The accumulation of mutations in older cells cannot currently be considered as a major cause of cloning failure.

The state of chromatin is thought to be essential for the genome reprogramming required for cloning. It is difficult to imagine how chromatin structure may differ so much between the cell types that produce successful cloning and those that regularly fail.

The stage of the cell division cycle has been invoked as a possible parameter favouring cloning or not. Quiescent cells in G0 of the cell cycle were used to generate Dolly. This experiment was repeated successfully in several other species but with donor cells in G0 or G1. A recent study even indicated that G1 is a more appropriate stage for nuclear transfer (Kasinathan *et al.*, 2001b). Hence, the major rule that seems to be respected is the compatibility between the cell stages of the recipient enucleated oocyte and of the donor cell.

It is recognized that cloning efficiency is highest when performed with non-cultured embryonic cells and lowest with adult cells. The pluripotent cells show moderate efficiency, which decreases when the cells are cultured for long periods of time. This observation is consistent with a hypothesis suggesting that cloning is successful mainly or perhaps only with stem cells still present in pluripotent or adult cell populations. Experimenters are currently unable to preferentially use cells on the basis of whether or not they have the characteristics of stem cells.

A recent work demonstrated without ambiguity that fully differentiated cells can generate live animals after nuclear transfer. Indeed, B- and T-cells in which immunoglobulin and T-cell receptor genes were specifically rearranged produced cloned mice harbouring the rearranged genes in all their cells (Hochedlinger and Jaenisch, 2002). However, clone development was possible only when the pluripotent cells from the cloned embryos were introduced into tetraploid blastocysts (Eggan *et al.*, 2002). This procedure is known to give rise to animals containing only the genome of the pluripotent cells. The completely differentiated B- and T-cells were therefore inefficient nuclear donors. This argues in favour of the idea that the success of cloning could be due to the presence of a few stem cells in the cultured primary cells used as nuclear donors (Colman, 2002).

Recent data strongly support the idea that cloned animals are epigenetically modified. Indeed, DNA is abnormally methylated in embryos generated by cloning instead of normal fertilization. Subsequently, many genes are not reactivated by DNA demethylation and unavailable for embryo development (Reik, Dean and Watter, 2001). A systematic identification of the methylated genes in clones that develop normally

or not could explain why so many clones are abnormal. This study could reveal the role of some genes in organogenesis. This does not necessarily imply that gene methylation in clones could be controlled and that this would enhance cloning results. DNA methylation seems to be a subtle process, which cannot be easily modified in a few specific genes. The defect in genome reprogramming and abnormal DNA methylation might account for the late effects of cloning on the life of animals.

It is striking that a large proportion of foetuses obtained by cloning do not survive during pregnancy and mainly after birth. The numerous genes required for life after birth are probably not all functional. A recent observation indicated that cloned mice have a shorter lifespan than normal animals. This does not necessarily imply that they were prematurely aged at birth. Rather, some of their biological functions do not work properly, thus leading to accelerated ageing (Perry and Wakayama, 2002; Ogonuki *et al*., 2002). The same is probably true for Dolly, who is suffering from early arthritis. It is, however, hazardous to draw conclusions from this observation, made so far in only one cloned farm animal. The other cloned animals are still too young to allow any conclusion to be drawn. It remains that early ageing could be observed in cloned mice because the life cycle and normal biological events in this species are short. The same study could be conducted in cloned farm animals (Enright *et al*., 2002). It would be quite interesting to determine the methylation status of DNA mice born after normal reproduction of cloned mice showing early ageing.

A particular point was the topic of discussion after the birth of Dolly. The telomeres of this animal were abnormally short, suggesting that Dolly was prematurely aged at birth. Dolly gave birth to two normal lambs after two pregnancies. These lambs have normal telomeres. All the cloned animals born after Dolly that were examined have essentially normal telomeres. Several clones obtained from a 17-year-old bull have normal telomeres. This suggests that cloning *per se* does not induce a shortening of telomeres. In fact, Dolly's telomeres were shorter than normal because the telomeres of the cells used as nuclear donors were already short before cloning. These cells were obtained from a six-year-old sheep, but more importantly they were cultured for long periods of time. Interestingly, the telomeres of Dolly's genome reached normal length in offspring. Sexual reproduction is thus able to restore damaged telomeres. This event probably occurred during gametogenesis.

It is currently impossible to correlate early arthritis in Dolly with the length of her telomeres. This phenomenon may not be related to cloning or may result from an ageing process indirectly linked to telomere length.

Clones must therefore be considered as epigenetically modified animals. It is striking that the capacity of the cells to give birth to cloned animals after nuclear transfer decreases as the number of passages increases. Little is known of the events that induce this phenomenon. They seem to be related to culture conditions. Many studies are obviously required to understand these problems. An empirical approach is currently expected to find the best culture conditions for improving cloning efficiency and for defining the experimental situations in which cells keep or lose their capacity to give rise to cloned animals.

The reprogramming of the differentiated cell genome after nuclear transfer implies that many chromatin domains adopt a new conformation. Most likely, there are a number of heterochromatin regions in which silent genes of the differentiated cells turn into euchromatin, leaving the possibility for the enclosed genes to be transcribed. This significant phenomenon is believed to result from the replacement of proteins from relatively methylated DNA by other proteins favouring euchromatin formation and gene expression. These proteins are present in oocyte cytoplasm. The same proteins are likely to transform highly condensed inactive chromatin in spermatozoa into decondensed active euchromatin. Little is known about the nature of these proteins. It is conceivable that the introduction of oocyte cytoplasmic proteins into differentiated cells could induce a reprogramming of their genome. It cannot be excluded that cloning may become possible in this way. A study using cell-free systems has shown that gene activation can be obtained by importing proteins into isolated nuclei (Landsverk *et al.*, 2002; Hakelien *et al.*, 2002). This approach may in time lead to the identification of proteins playing a key role in chromatin transconformation, gene expression and silencing.

Most cloned animals have been obtained by transferring intact cells between zona pellucida and oocyte plasma membrane. Many trials have shown that the transfer of isolated nuclei into enucleated oocytes is generally not followed by embryo development. Nevertheless, this method is currently the best way to generate cloned mice. These observations suggest that nuclear structure must be preserved during transfer to make embryo development possible. An interesting question is whether naked DNA introduced into an enucleated oocyte might be able to drive the development of a normal embryo. One technical difficulty is that

a simple genome should be introduced into the enucleated oocyte. This means that the amount of DNA corresponding to one genome has no chance of inserting a stoichiometric number of DNA sequences corresponding to each chromosome into an oocyte. Hence, DNA extracted from a single diploid cell should be injected into each enucleated oocyte. One possibility might be to complex naked DNA with oocyte proteins before its microinjection. It is by no means certain that the experiment will eventually be successful. It is conceivable that the high degree of complexity required for chromatin and the nucleus to be functional is incompatible with the use of naked DNA to integrate a diploid genome into an enucleated oocyte.

3.1.2 Transgenesis

The only real application of animal cloning is currently transgenesis. This point has been discussed above in Section 2.1.2.

It is now clear that gene addition is easier by cloning than by DNA microinjection into pronuclei in ruminants. In pigs, rats, rabbits and mice, which are prolific species, microinjection and perhaps infection with lentiviral vectors and DNA transfer by sperm (see Section 2.3 above) is still easier than cloning.

Gene replacement by homologous recombination has been achieved in sheep (Mc Creath *et al.*, 2000) and pigs (Dai *et al.*, 2002; Platt, 2002; Lai *et al.*, 2002) by the cloning technique. This remains the only way to do it. Gene replacement by cloning could also be obtained in mice (Rideout *et al.*, 2000). Clearly, gene replacement in mice through the classical use of ES cells and the generation of chimaeric embryos remains easier than by cloning. Moreover, optimal conditions for the culture of the cells used as nuclear donors have to be found before gene replacement by cloning can become a routine experiment (Denning *et al.*, 2001).

The implementation of cloning for transgenesis is advantageous, partly because it is a more efficient and comfortable technique for some species. Cloning theoretically allows us to choose the site of the genome at which a gene is going to be modified. A genome of an animal in which genetic modification has been conducted may be used preferentially because it comes from a male or a female. Furthermore, different cell clones obtained after transfection may generate transgenic animals that differ in terms of the integration of foreign genes into their genome. Another advantage of cloning is that the transgenic animals are not mosaic, as

opposed to DNA microinjection, which gives rise to about 30 per cent of mosaic animals.

Independently, simple reproductive cloning of transgenic founders may greatly accelerate the introgression of the genetic modification in herds.

3.1.3 Animal reproduction

Reproduction by cloning can theoretically accelerate genetic selection in different ways. Clones of animals with a valuable genome can be generated and crossed with other animals of the same breed or not, in a limited number. The phenotypic properties of the animals may suggest that the traits present in cloned animals are worth being introgressed by classical reproduction in a large number of breeding animals.

Despite its low efficiency and the uncertainty surrounding the long-term survival of cloned farm animals, cloning may be implemented to generate genetically identical animals used as sires in order to accelerate genetic selection. In agreement with this, sires prematurely forced to stop their activity for certain reasons and particularly after a lethal accident may be cloned using their frozen cells and pursue their reproduction via the cloned sires. This protocol is currently being followed in several laboratories. The impact of cloning is not expected to be significant as long as its efficiency remains low. Nevertheless, its use appears to be justified in its present state for several purposes (Westhusin *et al.*, 2001).

Racing horses are sometimes high-value animals and their cloning could produce champions more rapidly than classical reproduction.

For understandable reasons, pets will be cloned in the future. Recently, a cat was cloned by classical nuclear transfer. The owners of these animals believe that their favourite companions will continue to live with them although they are dead. A significant market has been identified in this field and companies have been created to clone pets. The clients of these companies might be bitterly disappointed if the health of their cloned pets happens to deteriorate prematurely or if they die or are simply not strictly similar to their genitors as is the case for the cloned cat.

Cloning is considered as a unique way to save species threatened with extinction. Attempts to save the gaur – a rare wild ruminant – by cloning, using cow oocytes as recipients gave rise to one animal which did not survive. This failure may be due to the cloning technique *per se* and

not necessarily to the fact that two different species contributed to the generation of the clone. The same approach was successful for another wild ruminant. Cells were collected from mouflons found dead in the pasture. Their nuclei were transferred into enucleated sheep oocytes. One apparently normal mouflon was born (Loi *et al*., 2001). Other trials are being performed with other threatened species. The efficiency of the cloning technique as well as limited financial support currently reduce the chance of this approach achieving significant success.

3.1.4 Human reproduction

The major visible impact of the birth of Dolly on public opinion was that adult humans could be reproduced by cloning.

Although cloning has only been achieved yet in a rather limited number of species, it is hard to believe that it is impossible in humans. Attempts to generate human embryos by nuclear transfer have been made in Korea and the USA (Cibelli *et al*., 2001). The results were too preliminary to be convincing. In both cases, only a few embryos started to develop after nuclear transfer and rapidly stopped. This suggests that cloning in humans by nuclear transfer will most likely not be easier than in other species. The poor success rate of monkey cloning achieved so far supports this hypothesis. Clearly, huge efforts are required to improve the cloning technique before its extension to humans can be reasonably envisaged.

Apart from technical difficulties, human cloning raises major ethical problems, which are discussed in Chapter 4.

3.1.5 Therapeutic cloning

Soon after the birth of Dolly, the issue of therapeutic cloning was raised. The basis of this approach is summarized in Figure 3.1.

Cell therapy is an extremely promising approach for humans. The idea is to graft differentiated cells into patients to restore a damaged organ (Griffith and Naughton, 2002). This is a very common practice in a limited number of cases. Blood cell transfusion is the most popular example. *In vitro* skin regeneration starting from the patient's skin cells is a classical way to treat burned victims. Attempts to graft pancreatic human cells into diabetics have not been successful to date for various

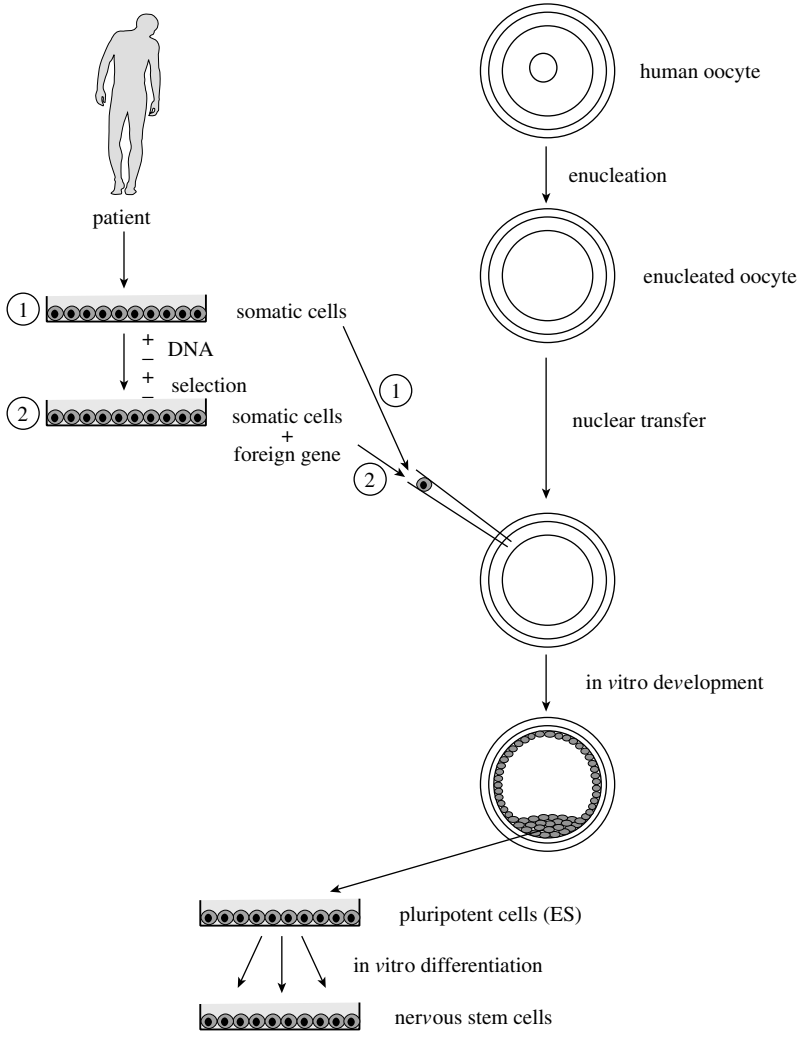


Figure 3.1 Use of human therapeutic cloning to obtain nervous stem cells to regenerate brain cells in patients. Somatic cells from the patient are used as nuclear donors to generate embryos. Embryonic cells are differentiated *in vitro* and the nervous cells are reimplanted in the patient's brain

technical reasons. Implantation of human hepatic cells to cure fulminant hepatitis has also met with quite limited success. A partial regeneration of cardiac muscle after grafting homologous skeleton muscle cells has been achieved recently and this approach appears promising.

The regeneration would be greatly amplified if organ stem cells could be used rather than differentiated cells. By definition, organ stem cells have the capacity to divide at a low rate without becoming differentiated. This process is sufficient to maintain a stock of stem cells in the organs. If the organ has to be regenerated, the stem cells divide rapidly and differentiate to produce the functional organ cells. A few stem cells can differentiate in only one cell type. This is the case for spermatozoon precursors. Most stem cells give rise to several cell types. Haematopoietic stem cells located in bone marrow produce red blood cells and the entire family of white cells (lymphocytes, macrophages, mastocytes, platelets etc.) (Figure 2.1).

It is known that the transplantation of haematopoietic stem cells into the bone marrow of a patient is followed by the appearance of numerous red and white blood cells. To reach this goal, the patient must be previously irradiated locally to reduce the number of his or her own bone marrow stem cells.

Stem cells have so far been found in a limited number of organs. They are rare and currently difficult to culture in most cases. For these reasons their use in cell therapy is at present quite limited.

Stem cells potentially capable of regenerating an organ can have different sources. Organs from adults are of course one possible source, albeit quite limited. Embryonic stem cells in mice are known to differentiate *in vitro* under the influence of various inducers. Different combinations of inducers added to ES cells are capable of producing red cells synthesizing haemoglobin, lymphocytes synthesizing antibodies, nervous cells synthesizing dopamine, muscle cells, cardiac cells, etc. This protocol has been extended successfully to human pluripotent cells (Solter and Gearhart, 1999; Thomson and Odorico, 2000; Lee *et al.*, 2000; Studer, 2001). It should be noted that the human cells differentiated from ES cells into neuron precursors colonized the mouse brain as normal neuron cells (Studer, 2001).

Pluripotent cell lines (ES cells) were established about 15 years ago from mouse blastocysts. Similar cell lines have been obtained independently from spare human embryos resulting from *in vitro* fertilization (Thomson *et al.*, 1998). Many spare embryos are available and can potentially be used for this purpose. Banks of pluripotent cell lines might be set up. The different lines might be classified according to their histocompatibility in order to minimize rejection in patients.

An alternative is to generate embryos by nuclear transfer using donor cells from the patients (Figure 3.1). In these conditions, the pluripotent

and differentiated cells derived from the cloned embryos are, inherently, histocompatible with the patient.

This protocol has been named therapeutic cloning before being implemented. To date, this approach is still only theoretical. Indeed, experimental conditions for cloning in humans are still essentially unknown. It is not known whether the pluripotent and differentiated cells would have the biological properties expected to repair damaged tissues. It should be noted that the defects observed in cloned embryos, which severely alter their development, may not be real problems for cell therapy. Indeed, a more limited number of functional genes are required to ensure the normal activity of a differentiated cell than the full development of a normal embryo.

A recent experiment carried out in mice provides quite significant experimental support for therapeutic cloning (Rideout *et al.*, 2000). Cells from adult mice suffering from a genetic disease were used to generate embryos by nuclear transfer. The genetic defect was previously corrected in the cultured cells by the transfer of the non-mutated form of the gene responsible for the disease. The mice born after cloning were healthy. This elegant experiment involves two therapies: cell and gene therapy.

Parthenogenesis may be an alternative approach to establish lines of pluripotent embryonic stem cells. In a certain number of species, an oocyte can develop without fertilization, giving rise to normal animals. In mammals, the development of an embryo from an oocyte is also possible, but only in experimental conditions and to a strictly limited degree. In all cases, development is slower and soon stops. It is therefore recognized that both paternal and maternal genomes are required for the complete development of mammals.

Parthenogenesis can be induced in mice, rabbits and other species, including monkeys (Mitalipov, Nusser and Wolf, 2001). The inducers are chemical agents. Ionomycin, which induces calcium flux in oocytes, and roscovitine, an inhibitor of MPF (maturation promoting factor), are sufficient to induce parthenogenetic development of rhesus monkey oocytes. Embryos artificially reconstructed in this way are diploid. The same method has been extended successfully to another primate, *Macaca fascicularis*, and could probably be used in humans as well.

Embryonic stem cell lines were established from the monkey parthenotes. Interestingly, the stem cells differentiated into cells that belong to the three germ layers mesoderm, ectoderm and endoderm after having been injected into the peritoneal cavity of immunocompromised mice

(Cibelli *et al*., 2002a). Hence, the injected stem cells migrated to different organs, where local factors induced their differentiation. This suggests that pluripotent stem cells, even derived from parthenotes, might contribute to restore damaged organs. The advantage of this approach is that a sufficient number of human oocytes could be recovered to establish multiple embryonic stem cells histocompatible with many potential recipients. Moreover, this approach raises only limited ethical problems. Numerous additional studies are required to validate the method (Thomson, 2002).

Another potentially interesting source of pluripotent stem cells or organ stem cells is the blood from the umbilical cord. Studies are under way to identify the nature of the cells present in blood from that area and to define conditions for their *in vitro* isolation and expansion.

It is generally acknowledged that the differentiation process is essentially irreversible in natural conditions (Figure 2.1). Experiments carried out recently indicate without ambiguity that this dogma is no longer acceptable. Indeed, mouse nerve stem cells were injected into the bone marrow of mice previously irradiated to eliminate their haematopoietic stem cells. A significant proportion of the transplanted nerve cells was transformed into haematopoietic stem cells capable of differentiating into functional red and white blood cells. Similarly, the transplantation of organ stem cells into another recipient organ was followed by their transformation into various organ stem cells (Figure 3.2). This phenomenon, named transdifferentiation, seems to occur widely and it might be a natural tissue regeneration mechanism. Indeed, it is conceivable, but not proven, that stem cells can migrate from a healthy organ to a damaged organ and participate in its restoration.

Lines of stem cells from a limited number of organs might thus be injected into damaged organs to trigger their reconstruction. Although this approach appears exciting and promising it remains to be demonstrated that transdifferentiation can be understood and controlled to become a therapy (Snyder and Vescovi, 2000; Lemischka, 2002).

The local factors that induce stem cell transdifferentiation have not generally been identified. It is conceivable that these inducers will be progressively discovered. They could be used to induce transdifferentiation *in vitro*. Transdifferentiation is most likely a complex and subtle phenomenon that requires particular cocktails of inducers and possible specific cell–cell contacts. It remains that these mechanisms must be understood to a sufficient degree to allow transdifferentiation for cell therapy.

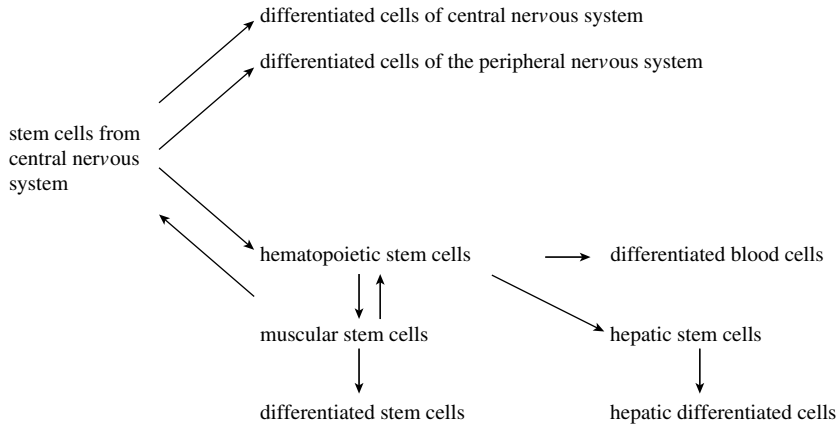


Figure 3.2 Transdifferentiation of organ stem cells. Stem cells of different organs can become stem cells of other organs after implantation into these organs. This method might be used to obtain missing organ stem cells for cell therapy

The number of known transcription factors is rapidly growing. It appears that a few of them are specifically expressed and active in a given cell type. PDX-1 is thus present in pancreatic islets of Langerhans. This factor is believed to play a major role in the differentiation of pancreatic cells into islets secreting insulin. A gene construct expressing PDX-1 cDNA was transferred into hepatic cells. The cells essentially retained their morphological aspect but they became capable of secreting insulin in a controlled manner, as a function of glucose concentration. Hence, a relatively minor transformation was sufficient to change the function of a differentiated cell. It should be kept in mind that the PDX-1 expressing cells retained most of their hepatic characteristics (Kahn, 2000).

It is tempting to imagine that a new approach has been found for cell therapy. Indeed, it is conceivable to recover hepatic cells from a patient suffering from diabetes to transfer a PDX-1 expressing gene construct and to reintroduce these cells into the patient. The transplanted cells might provide the patient with a sufficient amount of insulin when required.

This elegant approach might be possible in a number of cases in which the essential transcription factors controlling differentiation are found. Another example supports this hypothesis. Indeed, dexamethasone (a synthetic glucocorticoid) added to the culture medium of pancreatic cells transdifferentiated them into hepatic-like cells. This phenomenon was consistent with the induction of the C/EBP β transcription factor (Shen, Slack and Tosh, 2000).

3.1.6 Xenografting

The number of patients who need cells or organs is growing and it is acknowledged that the sources of human cells and organs are and will remain insufficient. The idea of using organs from animals emerged one century ago. It was followed by experimental transplantations of organs from different species to patients. In several cases, the experiments resulted in successful transplantation followed by a rapid and severe rejection response. Two organs, the testis and the internal chamber of the eye, were rejected much more slowly. It is now known that these tissues express a molecule named Fas ligand on their surface, which induces the death of activated immune cells.

Different animal species have been tested as a source of organs for humans. Primates including the chimpanzee were originally considered the most appropriate. It rapidly appeared that this choice was probably not the best. Organs from primates are rejected. Primates are protected species and their breeding is extremely expensive. Moreover, primates have the highest risk of transferring pathogens to humans. The idea of using primates as organ donors has currently been abandoned. The pig has been retained as the best compromise. This species is relatively close to humans. It is omnivorous and similar in size to humans. It is considered too different from humans to easily transfer its pathogens. Moreover, pig breeding can be performed in pathogen-free conditions at a moderate cost. In addition, the pig is currently and abundantly used as a source of food for humans. The additional sacrifice resulting from xenotransplantation would be rather small.

The use of immunosuppressors decades later did not improve the situation, indicating that rejection was not due to the known induced immunoreaction that occurs after allotransplantation of organs or cells.

Rejection of xeno-organs is known to be due to three different mechanisms, which operate sequentially. The first mechanism is hyperacute rejection. Certain natural antibodies permanently present in human blood rapidly recognize antigens on the surface of pig cells. The major antigen has been identified. It is a carbohydrate motif composed of galactose and is bound to proteins. This motif is synthesized by an enzyme named α-1,3-galactosyltransferase. This enzyme is present in most vertebrates but not in humans or in old-world higher primates. The gene coding for the enzyme was recently inactivated during evolution by a mutation, which has been identified. Humans, who are

meat-eaters, develop antibodies against the carbohydrate motif. These antibodies are the major natural antibodies responsible for hyperacute rejection of pig organs. The antibodies bind to the antigen of the endothelial cells in the grafted pig organs. This activates complement, which induces very rapid endothelial cell death and necrosis of the organs.

Transgenic pigs expressing anti-complement human genes have been generated in different laboratories. The heart and the kidney of these animals are less or no longer subjected to hyperacute rejection (Houdebine and Weill, 1999). A few days after transplantation, the pig organs are subjected to a delayed acute rejection mechanism and even later to a classical-induced rejection mechanism. These mechanisms are under study with the hope of defining essential genes controlling rejection. These genes might be counteracted in different ways by using genetic modification in pigs.

Recent publications from two independent laboratories reported that the gene coding for α-1,3-galactosyltransferase was activated by homologous recombination. This was possible only with the implementation of the cloning technique to generate living embryos from mutated somatic cells (Lai *et al*., 2002; Butler, 2002). The organs from these animals will be tested to evaluate their capacity not to be rejected after grafting in experimental primates.

It is by no means certain and is even unlikely that this will be sufficient to prevent rejection mechanisms. Indeed, even if the α-galactose motif is the most potent pig antigen, hundreds and even thousands of other antigens are present in pig cells and not in human cells. Moreover, the α-galactose motif is a potent inducer of the hyperacute rejection mechanism but not necessarily of the other mechanisms, which have to be counteracted by different approaches.

Xenografting remains a challenge with no certainty of success (Couzin, 2002). The inactivation of the rejection mechanisms is not the only problem facing xenografting. It remains to be proven that the function of pig organs is compatible with human physiology. Another concern is the possible transfer of pathogens from pig organs to recipients.

Retroviruses named PERV have been shown to infect some cell lines *in vitro*. This suggested that these viruses might infect patients. Gene inactivation by homologous recombination might make it possible to knock out the active pig endogenous retroviral sequences. This approach involves cloning. In practice, it appears simpler to select pig breeds expressing no retroviral particles. This has been achieved by different groups.

In the future, cloning might be quite useful to generate genetically identical animals. The individuals with the best genotype for xenografting in humans might be obtained in this way.

The different possible methods of obtaining cells or organs from humans or pigs for grafting in patients are summarized in Figure 3.3. It is clearly impossible to predict which of these will be the most appropriate. It is highly likely that one method will appear better to solve a particular problem, whereas another method will be more efficient for another problem. The simultaneous evaluation and optimization of these different methods are pursued.

A totally different means of preventing organ or cell rejection is theoretically to induce tolerance in recipients. For this purpose several approaches have been devised. Cells from the thymus, T-lymphocytes, are educated early in the life of animals to definitively recognize cellular components as self-antigens, which results in tolerance of the organism towards its own cells but not towards foreign cells.

Partial tolerance can be experimentally induced by injecting pig T-lymphocytes into the future recipient. This cannot be easily extrapolated

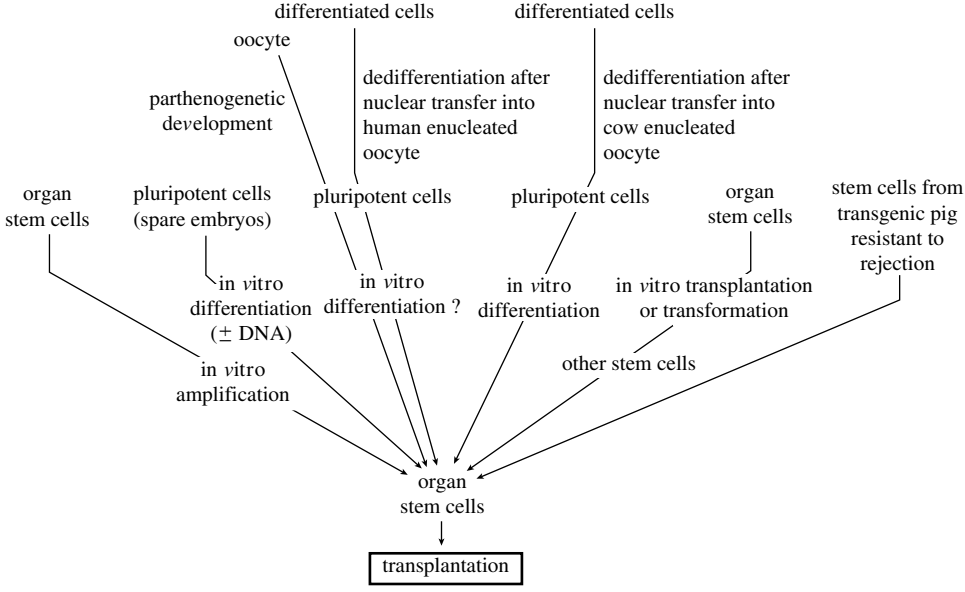


Figure 3.3 Different theoretical possibilities leading to the preparation of organ stem cells for transplantation to patients. None of these methods is validaded

to patients. Indeed, the pig lymphocytes would have to be introduced into the bone marrow of patients previously treated to transiently reduce their own number of T-lymphocytes (Butler, 2002).

A rather exciting advance seems to have been made recently. Pluripotent rat cells were introduced into rats previously irradiated to destroy their T-lymphocytes. The pluripotent cells became implanted in to the bone marrow and differentiated into T-lymphocytes. The rats harbouring these T-lymphocytes became tolerant to a heart transplant from another rat line known to be rejected in control animals (Adler, Bensinger and Turka, 2002).

These results emphasize once more the high potentiality of pluripotent cells. This protocol could be technically extended to humans. It remains to be demonstrated that this elegant approach is valid for animals other than rodents, namely of humans.

3.2 Applications of Animal Transgenesis

Transgenesis is more and more extensively used by researchers but also by industry for quite different purposes. For the sake of convenience, it is possible to divide the applications of animal transgenesis into three categories. The first and the major reason for using transgenic animals is to obtain fundamental information on living organisms and human diseases. The second reason is to use animals as a source of pharmaceutical proteins, cells and organs. The third is to improve animal production.

3.2.1 Basic research

The vast majority of transgenic animals are generated to be used as models for the study of gene function and mechanisms of action. This point has been largely discussed in Chapters 1 and 2. The trend in this field is to use animals in an increasingly systematic and integrated way. Special laboratories belonging to universities or to private companies generate experimental transgenic animals (mice, *C. elegans*, *Drosophila*, ...) by gene addition or replacement. In theses cases, the genes used are known and researchers need transgenic animals to evaluate a hypothesis on the role or the control mechanisms of their genes.

A more recent approach consists of systematically knocking out laboratory animal genes without any particular hypothesis on their

function. Banks of animals with a gene knocked out by homologous recombination or gene trapping are available for researchers and companies (Goodwin *et al.*, 2001). This systematic approach and the complete sequencing of the laboratory animal genomes is expected to provide the best chance of rapidly obtaining maximum information on gene function (Petters and Sommer, 2000; Hamilton and Frankel, 2001).

3.2.2 Study of human diseases

The methods described above to study gene function are also largely used to study human diseases. Indeed, a gene involved in a human disease can be added or specifically inactivated in mice. The models obtained in this way may provide researchers with invaluable experimental animals to describe human diseases and to evaluate the therapeutic effects of chemical compounds, recombinant proteins or genes transferred to patient tissues.

The mouse is the most commonly used species for this purpose. Although satisfactory in many cases, the mouse does not mimic some of the human diseases. This is essentially due to the fact that the mouse is a rodent and not a primate and some biological functions are different in mice and humans. A striking example is this of cystic fibrosis. This human disease is essentially caused by mutations of the CFTR gene. Mice whose CFTR gene has been knocked out do not develop the disease. It is acknowledged that the action of the no longer functional CFTR chloride channel is compensated by the effect of another gene that does not have the same pattern of expression in humans.

Mice are known to have a lipid metabolism different from that of humans. The study of atherosclerosis with in transgenic mice is thus not satisfactory. Rabbits, which are closer to humans in terms of this biological and other functions (Novacek, 1996), are a more appropriate model (Fan *et al.*, 1999b).

Transgenic rats appear to be the best model (Charreau et al., 1996) for the study of specific functions

Mice have another fundamental limitation. They are quite small animals, making surgical operations potentially difficult. Rabbits and rats are sometimes preferred for this reason.

Many transgenic mice are candidate models for the study of human diseases. In practice, only a limited number of mice appear relevant. This is due to the fact the transgene or the knocked-out gene does not always

generate a disease resembling a human pathology. The genetic background of the mouse is also quite important. This sometimes implies that the genetic modification obtained in a mouse line has to be transferred into another line by crossing (Carvallo, Canand and Tucker, 1997).

A few examples taken from different fields may illustrate the importance of transgenic mice for the study of human diseases.

3.2.2.1 Genetic diseases

About 6000 of all known human diseases are of genetic origin. In the simplest case, a gene and its mutations playing a major role in the disease have been identified. This is the case for cystic fibrosis. As mentioned above, the mouse models are not relevant due to the existence of a chloride channel other than that coded by the CFTR gene. Rabbits and sheep in which the CFTR gene has been knocked out are expected to be better models for the study of this disease.

Mice in which the LAM 2 gene coding for a laminin isoform has been knocked out show congenital muscular dystrophy syndrome. The transfer of the agrin gene, which codes for a protein known for its role in the formation of the neuromuscular junction, restores muscle function. It should be noted that another gene coding for a quite different protein can also suppress the pathology (Moll *et al.*, 2001).

In other cases, several genes are involved in the pathology and this was revealed by the simultaneous action of two independent transgenes. Amyotrophic lateral sclerosis is much better mimicked in mice expressing the superoxide dismutase 1 gene and the gene coding for a neurofilament subunit than each gene separately (Kong and Xu, 2000).

3.2.2.2 Infectious diseases

Many pathogens are species specific and animals are often not relevant models for the study of a human infectious disease. This limitation is often due to the lack of receptors for the pathogen in the animal. In a certain number of cases, animals expressing a human pathogen receptor have become sensitive to infection.

A model for listeriosis E-cadherin in enterocytes is known to recognize internalin, in a protein of the *Listeria monocytogenes* surface. Mice expressing the human E-cadherin gene in their enterocytes are good models for the study of infection by *L. monocytogenes* (Lecuit et al., 2001).

Models for viral infections Mice expressing the human CD 46 gene are sensitive to the measles virus (Oldstone *et al.*, 1999). Similarly, mice have become sensitive to the polyomyelitis virus after transfer of the viral receptor gene (Ren *et al.*, 1990).

Mice sensitive to human hepatitis C virus have also been obtained. In this case, the presence of the U-plasminogen activator gene greatly improved the model by prolonging the life of human lymphocytes grafted into immuno-suppressed SCID mice (Fausto, 2001).

Transgenic rabbits expressing the human CD4 gene in T-lymphocytes were shown to be transiently sensitive to HIV infection (Dunn *et al.*, 1995). The discovery of the second HIV receptor, CCR5, suggested that the transfer of both genes could generate better models (Cohen, 2001). A quite different approach gave rise to a rather attractive model. Transgenic rats harbouring the whole HIV genome minus the gag and the pol genes showed syndromes substantially similar to AIDS (Reid *et al.*, 2001).

Models for prion diseases It is no exaggeration to consider that prion diseases would have remained largely misunderstood without the contribution of transgenic mice (Moore and Melton, 1997). These models revealed the essential role of the PrP gene in prion diseases. Animals whose PrP gene has been knocked out are no longer sensitive to prion diseases. These animals develop a new sensitivity towards prion diseases according to the origin and the allele of the PrP gene. It is also clear that the PrP gene and the PrP protein are necessary but not sufficient to induce a prion disease in animals. The mice models in which the PrP gene was replaced by the bovine PrP gene revealed that they became sensitive to bovine spongiform encephalopathy and that the development of the disease was influenced by environmental factors (Manloakou *et al.*, 2001).

3.2.2.3 Alzheimer's disease

Alzheimer's disease is a complex pathology, which occurs after long periods of time. It is characterized by the accumulation of insoluble

β-amyloids and polymerized tau protein. The genes coding for prenisilin 1 and 2 also play an important role in the disease. Mice whose apolipoprotein E and interleukin-1 genes have been knocked out are significantly less sensitive to the disease. Insoluble β-amyloids result from a cleavage of amyloid precursors by secretases α, β and χ. Various transgenic models have greatly contributed to establishing these facts. Some of these models are currently used to evaluate the therapeutical effects of chemical drugs. Ibuprofen, an anti-inflammatory molecule, and several secretase inhibitors, have thus been shown to delay the formation of β-amyloid plaques in the brain of transgenic mice (Chapman *et al.*, 2001).

3.2.2.4 Atherosclerosis

Atherosclerosis is a complex disease, which develops at a slow rate. Numerous genes involved in lipid metabolism (apolipoproteins, lipase and others) have been added or knocked out in mice (Miller and Rubin, 1997). Although the mouse is not a good model for the study of human lipid metabolism, the transgenic animals provide invaluable data. Many of the above-mentioned genes have also been transferred in to the rabbit, which is definitely a better model for the study of atherosclerosis (Brousseau and Hoeg, 1999). Some of these animals are used to evaluate the anti-atherosclerosis effect of chemical drugs and also to define conditions for a putative gene therapy.

3.2.2.5 Cell death

Apoptosis plays a major role in organogenesis and tissue involution. Defaults in apoptosis may lead to autoimmune diseases, tumour formation and neurodegeneracy. At least 25 genes involved in apoptosis have been identified and knocked out (Ranger, Malynn and Korsmeyer, 2001). This enables a better understanding of the mechanisms involved in apoptosis and the potential design of drugs controlling these mechanisms.

3.2.2.6 Cancer

It is acknowledged that cancers result from several gene mutations leading in a step-wise manner to cell immortalization, formation of

primary tumours and finally metastasis. This implies the activation of several oncogenes and the inactivation of anti-oncogenes. Transgenic mouse models greatly contribute to revealing the complex roles of these genes (Bartek and Lukas, 2001).

C-myc, erbB2 and cyclin D1 have a cooperative effect in inducing mammary tumours (Siegel, Hardy and Muller, 2000; Yu, Geng and Sicinski, 2001). Other oncogenes such as neu, ras or Wnt 1 also contribute to the generation of mammary tumours. It has been observed that mice in which the cyclin D1 gene has been knocked out are more resistant to mammary tumour formation. Drugs inhibiting cyclin D1 might have an anti-tumour action. It is well established that mammary cells are stabilized by extracellular matrix. It is also known that metastatic cells degrade extracellular matrix, thus enabling them to migrate in the body and colonize different tissues, where metastasis develops. A gene coding for an anti-metalloproteinase factor, Akt, delays extracellular matrix degradation in the mammary gland and potentially mammary metastasis (Schwertfeger, Richert and Anderson, 2001). All this information supplies researchers with clues to design new anticancer drugs.

Tumours are known to derive from a single cell in which mutations have occurred. This explains why tumour formation is a slow process in its early stages. These mutations occur progressively and classical transgenic models do not take this factor into account. Specific gene constructs capable of sporadically activating oncogenes have been designed (Berns, 2001). This method has allowed activation of the K-ras oncogene according to a random process and at a low frequency. This approach leads to the definition of better models for the study of lung cancer (Johnson *et al.*, 2001b).

3.2.2.7 Xenografting

The problems of xenografting have been largely discussed in a previous section (3.1.6). Transgenesis has two different roles in this field of research. Transgenic laboratory animals (mice, rats and rabbits) are generated to decipher the rejection mechanisms and to define genes that could inhibit them. The cells and organs to be grafted in patients are expected to be recovered from pigs in which the genetic modifications performed have been previously studied in laboratory animals (Houdebine and Weill, 1999).

3.2.3 Pharmaceutical production

All human communities are searching for substances to combat diseases. Originally, essentially crude plant extracts played this role. This method is still quite popular. The emergence of chemistry made it possible to determine the structure of the active molecules in plant extracts and in some cases to obtain these molecules or analogues by chemical synthesis. The case of aspirin is illustrative. Willow leaf extract was used extensively until it was discovered that its active compound was salicylic acid. The chemical synthesis of an analogue, acetyl-salicylic acid, named aspirin, was achieved in the early 20th century. This protocol is still followed successfully to prepare quite efficient drugs.

Despite their major role in life, proteins were totally absent from the traditional pharmacopoeia. This is clearly due to the fact that orally absorbed proteins rapidly loose their biological activity.

Some of the proteins discovered subsequently came to be used as pharmaceuticals. Insulin is a good example. For decades, this hormone was extracted from pig pancreas and injected into diabetics to regulate their glucose metabolism. Other extracted proteins such as growth hormone, coagulation factors, and antibodies have been and still are used.

This approach has some advantages and severe drawbacks. The major advantage is that proteins are in their native state and quite active. Some proteins can be prepared in sufficient amounts for patient treatments. This was the case for pig insulin and human growth hormone. In contrast, human erythropoietin is present in blood at a much too low concentration to be extracted for patient treatment. Pig insulin is not strictly similar to its human counterpart. Furthermore, proteins extracted from human blood or tissues can be contaminated with pathogens such as viruses or prions.

Genetic engineering has profoundly changed this situation. A gene coding for a protein of pharmaceutical interest can be isolated, inserted into an expression vector and transferred into cells or organisms, which become producers of the protein on an industrial scale.

This was achieved for the first time with human insulin, which was prepared from genetically modified bacteria about 15 years ago. The majority of the diabetics are now treated with recombinant insulin and no longer with extracted pig insulin. The recombinant hormone has a higher degree of purity and is structurally identical to native human insulin.

This first success was rapidly followed by another. Human growth hormone of bacterial origin is the only form that has been used for more than a decade. This eliminated the risk of contamination by the human prion responsible for Creutzfeldt–Jakob disease.

Although potent, this method rapidly showed its limits. Some proteins are poorly synthesized by bacteria. Others become insoluble in bacteria and their purification is difficult. Moreover, many proteins of pharmaceutical interest, namely human proteins, are glycosylated or have to be post-transcriptionally modified to be biologically active. Bacteria are unable to proceed to most of these maturation stages. The same is relatively true for recombinant yeast. Proteins such as α1-antitrypsin or erythropoietin prepared from bacteria show biological activity *in vitro* but not *in vivo*. This is clearly due to the fact that these proteins are not glycosylated and are thus highly unstable in blood circulation, since they are rapidly taken up by the kidney and liver.

These observations have led to the use of genetically modified animal cells as the source of recombinant proteins. This approach is used successfully to prepare proteins on an industrial scale. Monoclonal antibodies can be obtained from cultured immortalized B-lymphocytes named hybridoma. Their production remains limited and many hybridoma are genetically unstable and lose the genes coding for the antibodies.

Antibodies are prepared in relatively large amounts from animal cells, such as CHO (chinese hamster ovary) in which gene constructs containing the genes for the two chains of antibodies have been introduced.

Other complex proteins are also prepared from genetically transformed cells. The only source of human erythropoietin currently consists of animal cells cultured in large fermentors. This protein is properly glycosylated and shows excellent biological activity.

The production of recombinant proteins from cells is thus a reality but this process has limited potency. Amounts of proteins not exceeding a few kilogrammes per year can be prepared from cells at a reasonable cost. Another possible approach is to use animal cells in living organisms. The advantages are that cells are as numerous as required and maintained in ideal metabolic conditions. This possibility was raised for the first time in 1982, when the first mice expressing their transgene at a high level were obtained. These animals harbouring the rat or human growth hormone gene exhibited accelerated growth and some of them had a high growth hormone concentration in their blood (up to 50 μg/mL). This gave rise to the idea that blood from transgenic animals could be a source of recombinant proteins.

After the generation of transgenic animals larger than mice (rabbits, pigs and sheep) (Hammer *et al.*, 1985), it appeared conceivable to prepare transgenic farm animals secreting foreign proteins in blood, milk etc.

Experiments published in 1987 demonstrated that active foreign proteins could be secreted in the milk of transgenic animals (Simmons, McClenaghan and Clark, 1987; Gordon *et al.*, 1987). To reach this goal, milk protein gene promoters were fused to the coding region of the sheep β-lactoglobulin gene and human tissue plasminogen activator.

This method allowed the experimental secretion of 100 foreign proteins in milk. A few of these are secreted in large amounts in rabbit, sheep, goat and cow milk and they are subjected to clinical tests. It is expected that the first recombinant protein extracted from milk will be on the market in the near future. One of these proteins, human α-glucosidase prepared from rabbit milk, improved the clinical condition of babies suffering from Pompe disease (Van den Hout *et al.*, 2001).

Although efficient, this method for preparing pharmaceutical proteins has limitations. Some proteins are secreted at a very low concentration, hardly compatible with industrial use. Some proteins alter the health of the animals. The leaky expression of the transgenes and the transfer of small proteins from milk to blood are sufficient to produce a significant biological effect on the animals.

The mammary gland shows an excellent capacity to synthesize complex foreign proteins such as collagen, fibrinogen, spider silk and EC-superoxide dismutase. However, a certain number of post-translational modifications of the recombinant proteins secreted in milk are not properly achieved. For instance, human protein C was not properly cleaved to eliminate its propeptide and to generate subunits. Several proteins such as human antithrombin III are not fully glycosylated. A few experiments have shown that the expression of furin from a transgene can greatly improve protein C cleavage. It seems likely that genes coding for various glycosidases will improve glycosylation of the recombinant proteins secreted in milk. This is generally not a particularly difficult task (Lubon, 1998).

The proteins can be purified from lactoserum in an undegraded state since milk is essentially devoid of proteases. The recombinant proteins extracted from milk must not be contaminated by pathogens. Guidelines have been defined to solve this problem. This point is discussed in Chapter 4.

The available data accumulated over one and a half decades indicate that the mammary gland of transgenic animals is an appropriate

fermentor for preparing a number of functional recombinant proteins at a low cost. The animals used to prepare recombinant proteins in their milk are and will probably remain in the future rabbits, sheep, goats and cows. Pigs might also be used. Each species has specific advantages and drawbacks. Transgenesis is efficient in rabbits, which are also very prolific animals. Prion diseases and other severe diseases transmissible to humans are not observed in rabbits. Up to 10 kg per year could be prepared from the milk of about 1000 rabbits. Ruminants are more appropriate but less flexible than rabbits for preparing large amounts of protein.

The recombinant proteins prepared from milk are quite different. They may be hormones, growth factors, enzymes, blood factors, vaccines, antibodies, structural proteins etc. Antibodies, which are potentially extremely diverse, should become the proteins most frequently prepared from milk.

Milk is the biological fluid that is currently the most studied and the closest to industrial use. Other fluids could be an attractive alternative, one of which is blood from mammals. Urine and pig seminal plasma are other possibilities. Secretions from the silk glands of silkworm could also be a source of recombinant proteins. The egg white of transgenic chickens appears to be an attractive option, which remains ineffective until transgenesis has become routine in this species (Houdebine, 2000).

Although milk offers an attractive alternative to cultured cells, recombinant proteins can also be prepared from transgenic plants. These different systems will most likely be implemented according to the problems to be solved. The number of proteins to be prepared is currently high and rapidly increasing. It is now recognized that all the available methods for preparing recombinant proteins will be insufficient to meet the demand for a decade.

3.2.4 Xenografting

The different possible applications of xenografting have already been discussed in Section 3.1.6. Xenografting is not yet a reality and it is impossible to predict when and even whether this will eventually occur.

The first problems to be solved are in the scientific field. Rejection mechanisms are not fully understood and all the protocols for inhibiting these mechanisms, involving transgenesis or not, have not yet been defined. It is expected that the severity of rejection will be different

when isolated cells or organs are grafted. Indeed, it is known that vascularized tissues are strongly rejected. This is due to the fact that the pig endothelial cells of the grafted organs are rapidly destroyed by human complement. This induces thrombosis and a rapid destruction of the organ. Isolated cells or aggregates are not sensitive to these phenomena. Yet, pig pancreatic islets are rapidly rejected when grafted into primates.

The second category of problems arises from the fact that pig organs may not function efficiently in humans. It is acknowledged that a pig heart would work properly in a human. A kidney would not be so well adapted to human hosts. It currently seems unrealistic to envisage grafting of a pig liver into a human. The functions of this organ are too complex to be easily compatible between different mammals. Isolated cells or aggregates may raise fewer problems. Porcine pancreatic cells secreting insulin might work correctly in humans. The same could be true for neurons secreting dopamine or for skin.

It should be kept in mind that xenografting may offer a transient solution, enabling a patient to survive until a human organ is available. In particular cases, pig cells may be present no longer than one day after grafting in patients or may be used in extracorporeal circulation. Pig liver cells are used to help patients suffering from fulminant hepatitis. The cells are maintained in an extracorporeal reactor and they detoxify the human blood circulating in the reactor. The pig cells could remain functional longer if they were prepared from transgenic pigs expressing genes that provided resistance to human complement.

The third problem is the possible infection of patients with pig pathogens, namely viruses. This point is discussed in the next chapter.

Although xenografting is not about to become routine practice, it remains an attractive method for organ replacement and cell therapy. The use of human stem cells may appear to be more appropriate for cell therapy in the future. Generating human organs *in vitro* remains a challenge. Only skin and blood cells grown *in vitro* are regularly prepared to be transplanted into patients. Pigs thus remain a potential source of organs for humans.

3.2.5 Breeding

Genetic selection is one of the major methods used to improve animal production. This approach has become more powerful since the discovery

of the heredity rules and it is still being improved by the use of genetic markers (see Section 1.6). Transgenesis is logically expected to generate mutants of breeding interest that could not have been obtained by classical selection, at least within a reasonable period of time (Wheeler and Watters, 2001). Transgenesis in plants, which is based only on gene addition for technical reasons and not at all on gene replacement, has generated some new varieties highly appreciated by farmers. The same is theoretically possible for animals. In practice, this is only partly true. Indeed, transgenesis is still more complex and costly in farm animals than in plants. Furthermore, the introgression of the genetic traits introduced by the transgenes cannot be a rapid process, even if the cloning of founders is implemented. The number of possible trials in animals is and probably will remain lower in farm animals than in plants. The candidate genes for improving animal production by transgenesis must therefore be carefully selected. The number of validated genes to generate relevant transgenic farm animals is presently quite limited. The study of their positive and potentially negative effects must be performed in many cases in transgenic mice.

A large part of the investment in transgenic farm animal research has been restricted for years to applications in the medical and pharmaceutical fields depicted above. The success achieved in preparing pharmaceuticals from the milk of transgenic animals and in generating transgenic pigs for xenografting has recently given some experimenters an incentive to start developing projects aimed at improving breeding. The important advances in the techniques used to add and replace genes in farm animals by cloning are progressing in the same direction.

3.2.5.1 The genes that may improve animal production

Although the number of projects under way to improve animal production is limited, the theoretical possibilities, trends and result are worth considering. The available data were reported in a recent review (Houdebine, 2002).

Resistance to diseases In agriculture, up to 40 per cent of the harvest may be lost as a result of different diseases. In breeding, the loss may reach 20 per cent. Reducing disease frequency in farm animals seems justified for different reasons (Müller, 2000). Animal resistance to dis-

eases will in some cases reduce the use of antibiotics. This should slow down the emergence of resistant pathogens. Animals free of disease enjoy a better quality of life. Breeding healthy animals is simpler for farmers. A reduction of disease frequency is expected to increase production or to reduce production costs. Another reason for generating farm animals resistant to diseases is to attempt to reduce the transfer of such diseases to humans. Indeed, a number of diseases are transferred from farm animals or pets to humans. Reducing disease transfer from animals to humans is a priority in most developed countries. Transgenesis can contribute quite significantly to this endeavour.

The genes capable of preventing diseases in animals are theoretically quite diverse. These genes may inhibit virus infection or replication. The tools described in Section 2.3.11 are good candidates for this goal, although evidence supporting this is still lacking in most cases. Some natural alleles are known to protect individuals against diseases. These genes have rarely been identified. The systematic study of farm animal genomes will progressively contribute to identifying these genes, which can be transferred to animals of the same species and possibly of another species to protect them against a given disease.

A general approach to disease prevention may be to transfer the genes coding for the different chains of antibodies known to neutralise the pathogen.

Transgenesis may indirectly contribute to the fight against animal diseases through the synthesis of recombinant vaccines in milk. The antigens in milk can be administered orally to animals in some cases and injected in the other cases.

Digestion and metabolism Optimizing digestion in farm animals may have several advantages. It may diminish the risk of rejection and pollution, reduce feed consumption and enhance production yield. Metabolic modifications may generate animals better adapted to available or low-cost feed.

Milk composition Milk is the source of about 30 per cent of the proteins consumed by humans in developed countries. Milk production has been greatly improved by selection and optimized feeding. Milk composition has only been slightly changed. Essentially lipid concentration has been enhanced.

It is well known that several cow milk proteins are allergens. Moreover, most humans cannot digest lactose after weaning. This is due to the inactivation of the lactase gene. Undigested lactose is used by bacteria in the intestine. This generates severe disorders, which dissuade many people from eating milk products.

The protein composition of farm animals could be optimized to enhance the quality of curd and its ability to prepare cheese, yoghurt and other milk derived products.

Milk could contain human milk proteins rather than cow proteins. This milk would be less allergenic and more digestible.

Milk could be supplemented by proteins, providing an oral vaccine. Antibodies capable of neutralizing pathogens in the digestive tract could be added to milk.

Milk used to feed offspring could be more nourishing and favour newborn growth.

All these modifications have a greater chance of being achieved by transgenesis than by selection. In many cases, transgenesis even appears to be the only way (Houdebine, 1998; Pintano and Gutierez-Adan, 1999).

Carcass and wool growth Although the growth of farm animals has been enhanced by selection for centuries, further improvements are conceivable. The accelerated growth of animals may be required in some cases. Specific muscle growth and a reduction of fat storage is preferable in other cases. This is expected to diminish production cost and pollution or to enhance meat quality.

The accelerated growth of wool and modification of its composition are expected to have a significant economic impact.

A higher prolificacy of farm animals is also expected to play a significant role in animal production.

3.2.5.2 The transgenes under study to improve animal production

Various projects aimed at improving animal production in entirely different ways are under way. Some of the transgenes are being studied in mice, whereas others have been transferred already into farm animals (Houdebine, 2002).

Resistance to diseases Mice expressing the lysostaphin gene in their mammary gland have been obtained. This bacterial protein has a bactericide action on Gram-negative bacteria. Lysostaphin is expected to prevent mastitis.

Lactoferrin and lysozyme genes have been transferred to mice and cows for the same purpose.

Preventing prion diseases in ruminants is essential. One possible approach consists of knocking out the PrP gene out by homologous recombination. Experiments in progress have not yet been successful (Denning *et al.*, 2001). Another attractive option might be to transfer genes coding for antibodies capable of neutralizing PrP protein activity. Transgenic mice secreting such antibodies in their blood show reduced sensitivity towards scrapie (Heppner *et al.*, 2001). This quite encouraging result suggests that infectious diseases, including prion diseases, can be prevented by passive immunization in transgenic animals.

Digestion and metabolism Pigs are known to be a major source of pollution. They reject phosphorus and nitrogen, which contaminate soil and water. This induces eutrophication in ponds, rivers and even in the sea. Phosphorus rejection results from the fact that pigs ingest phytic acid present in plants but are unable to digest it. Phytic acid is a highly phosphorylated carbohydrate, which is also a chelator, trapping ions rejected by the animals, which may consequently suffer from ion deficiency. Phytic acid is digested by bacteria, after which phosphorus is released, leading to pollution.

To circumvent this mechanism, several complementary solutions have been envisaged and are under study. One possibility consists of selecting plants depleted in phytic acid. A quite different approach may be to add phytase to pig feed. This enzyme can be extracted from different bacteria and fungi. This method has shown some efficiency. Lower amounts of phosphate are rejected by pigs and the phosphate released into the digestive tract of the animals contributes to their growth. Transgenic plants containing phytase have been obtained. They are currently being tested to evaluate their capacity to diminish phosphate rejection.

A recent work offers a rather attractive solution. Transgenic pigs secreting bacterial phytase in their saliva have been generated. These animals are healthy and they reject only 25 per cent of the phosphate they absorbed with phytic acid. The released phosphate enhances animal growth and supplementation of feed by mineral phosphate is no longer

needed. These pigs, named ecopigs, are under study to evaluate their harmlessness for human consumers, their impact on breeding and the real advantage over other methods (Golovan *et al.*, 2001a, 2001b).

The adaptation of animal metabolism to available feed has been envisaged. Two ambitious projects are worth mentioning. Most farmed fish eat proteins, which are their major source of energy. These animals have a very low capacity to digest carbohydrates, which are available in large amounts and lower in cost than proteins. A study under way has revealed that several genes involved in carbohydrate digestion are missing or poorly expressed in fish. These genes taken from other species could be transferred into fish. These animals could use carbohydrates from cereals as an important source of energy.

Ruminants essentially use acetate generated by the rumen as the source of circulating energy for cells. Genes coding for enzymes capable of transforming acetate into glucose could be transferred into ruminants. These animals might utilize their feed more efficiently (Ward, 2000).

Milk composition Overexpression of k-casein in mouse milk is expected to reduce micelle size and enhance milk stability (Bösze *et al.*, 2001). Results obtained by different groups are conflicting, and it is not clear whether this experiment, currently extended to farm ruminants, will really improve milk composition.

Experiments to knock out the β-lactoglobulin gene are under way. This protein is one of the major milk allergens.

Overexpression of αs1-casein is expected to enhance overall protein secretion in milk (Chanat, Martin and Olliver-Bousquet, 1999).

Reducing lactose concentration in cow milk would have several theoretical advantages. The intolerance of consumers to lactose would be less severe. The quantity of water in milk would be lower, leading to reduced mammary gland engorgement and infections (Whitelaw, 1999). Several approaches have been evaluated to reduce the lactose content of milk. The most promising method appears to be to express a lactase gene in the mammary gland. The enzyme digests lactose within the mammary cells but also in secreted milk. This significantly reduces lactose concentration in milk, without altering mammary gland physiology (Jost *et al.*, 1999).

Several monoclonal antibodies exhibiting a neutralizing effect on gastric pathogens have been secreted in transgenic mouse milk (Saif and

Wheeler, 1998; Castilla *et al.*, 1998; Kolb *et al.*, 2001). Some of these antibodies protect mice against coronavirus infection.

A mutated human α-lactalbumin secreted in the milk of transgenic cows is a potential source of protein devoid of phenylalanine for patients suffering from phenylketoneurea.

Cow α-lactalbumin secreted in the milk of transgenic pigs improves the nutritional value of the milk. This allows a higher number of piglets to survive after weaning (Bleck *et al.*, 1998).

Carcass and wool growth Transgenic pigs overexpressing the pig growth hormone gene have enhanced muscle development with a very moderate increase in body size. About 60 independent lines of transgenic pigs have been generated to select those exhibiting significant muscle growth without any alteration of animal health (Nottle *et al.*, 1997).

Pigs overexpressing the IFG1 gene specifically in their muscle have a moderate but significant and quite specific enhancement of muscle development (Pursel *et al.*, 2001).

Transgenic mice overexpressing a gene coding for a transdominant negative myostatin develop muscle hypertrophy. This mimics the double-muscling trait of several cow breeds. This is known to result from a mutation in the myostatin gene. Interestingly, the promoter used to drive expression of the transdominant negative myostatin gene in the transgenic mice was active only after birth. This only allows a post-natal muscle hypertrophy, avoiding calving problems (Yang *et al.*, 2001). This protocol is being extended to cows.

Fish growth hormone genes have been transferred to various fish species, including trout, salmon, carp, catfish, tilapia and loach. These animals show considerably accelerated growth, ranging from two fold in most species to seven fold in salmon and up to 35-fold in loach (Devlin *et al.*, 1997). Transgenic salmon have been studied in depth. Their feed consumption is lowered to reach the same degree of development. This reduces pollution, cage occupancy and production cost. Their flesh seems to be of good quality. Experiments are under way to determine whether these animals can be eaten by human consumers without any risk and whether their dissemination in the environment can be controlled.

Accelerated wool growth in transgenic sheep could be obtained by enhancing cysteine synthesis. This project, although promising in

mice, proved inapplicable in sheep due to transgene side-effects (Ward, 2000).

Modifying wool composition to improve its mechanical property is an attractive project. Several genes coding for wool proteins have been expressed in transgenic mice and sheep. Modifications in wool composition have been obtained but with no apparent economic interest so far (Bawden *et al.*, 1999).

Other unreported and unpublished projects are under development. These projects are either at a too preliminary stage or kept confidential. An interesting project, which is worth mentioning, is being developed by a Japanese group. The desaturase gene from spinach has been used to generate transgenic pigs. Thee animals have a lower content of saturated lipids, known to induce atherosclerosis in humans (Iritani *et al.*, unpublished data).

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

4 Limits and Risks of Cloning, Gene Therapy and Transgenesis

The emergence of cloning and transgenesis has been abrupt and most people are not able to evaluate the advantages and drawbacks of these techniques at this time.

Cloning and transgenesis, both of which may or may not have a strong direct impact on humans, have every reason to be the subject of debates. The exchange of ideas in these domains is difficult at this time and may be confusing. This can be attributed to independent and complementary factors. The vast majority of people still have a quite limited knowledge of biology and biotechnology, as well as modern agriculture and breeding. A certain number of events have cast doubts on and even generated fear of science and its applications. Stories of human blood contaminated by viruses, mad cow and foot-and-mouth disease, global warming etc. are contributing to this state of mind. A certain number of people are always wary of novelty *per se*, whereas others feel threatened by the emergence of new techniques. Some people just hate science, and intellectual work in particular. Scientists have been unable to explain what they do, their expectations and what their real role is in research and research applications. The media often do not cover scientific issues in sufficient depth, particularly in the biotechnology field. This leads to publications that are a poor reflection of reality. An increasing number of people have doubts about the way that modern societies are governed. Science is accused of catering to multinationals and not respecting workers, consumers, people seeking medical care or the environment. Many researchers want to preserve their freedom at all costs and are reluctant to collaborate with private industry. It may thus seem

paradoxical that the scientific community is not considered to be a natural ally of some of biotechnology's adversaries. Indeed, the myth that scientific or technical progress is always beneficial to humans has almost completely disappeared from researchers' minds. This myth is still exploited as a form of publicity by some companies. Science is supposed to have increasing power over human beings. Instead, it seems as if economic considerations, advertising and the control of the media are playing an increasingly important role in modern societies, including science and its applications.

The idea that biotechnologies must be tightly controlled by society has met with a very broad consensus. The problem is how to manage the rules capable of preventing excesses while respecting innovation. Indeed, the principles on which modern ethics concerning biotechnology are based are not all new. The Greek doctor, Hippocrates, recommended using a medication only if it was beneficial and not harmful to the patient. It is now increasingly acknowledged that a treatment must respect patient autonomy to the greatest degree possible and must be offered to all those suffering from the same disease. E. Kant said that human actions should always give priority to the well-being of humans. More recently, H. Jonas has defined the 'precautionary principle', which is a modern and formalized form of traditional wisdom. This principle may be applied even if there is no scientific proof to indicate that the risk is real. The application of the precautionary principle must be accompanied by intensive research into the risks. It must therefore be reversible. Moreover, this principle can only be applied if its cost is reasonable for society.

The application of the precautionary principle implies that a certain level of fear towards potential risks is maintained in the public opinion. This level seems to be excessively high in several domains of biotechnology at this time.

Clearly, bioethics and risk management are in permanent evolution and this is cause for confusion for people waiting for simple answers to complex problems. A common source of confusion consists of making no distinction between risks and ethics. Risk management is no more that the implementation of techniques to guarantee that ethical principles are respected.

Strangely enough, ethical principles are almost exclusively evoked at this time as a means of preventing the negative impact of science and technology. It is rarely brought to mind that one of the ethical obligations of scientists is to alleviate human suffering through research.

Animal cloning and transgenesis raise specific problems. Some of them are scientific and technical and limit the way they can be used. Biorisks and ethical problems are also concerns for animal experimentation and breeding. These problems cannot be separated from those that could result from the application of cloning and transgenesis to humans.

4.1 Limits and Risks of Cloning

The first limit in the use of cloning is its low efficiency and the fact that the health status of the animals born after cloning is often poor and unpredictable (see Section 3.1).

4.1.1 Reproductive cloning in humans

The technical problems encountered when cloning animals preclude all reasonable projects of reproductive cloning in humans. Indeed, cloning techniques in their present state are far below the standard level of reliability required for use in human reproduction (Mollard, Denham and Trounson, 2002). The data reported recently indicating that cloned mice have a shorter life span than control animals reinforce this point (Ogonuki *et al.*, 2002).

Quite unexpected observations are made in cloned animals. A recent work revealed that cloned mice have an obese phenotype. It is interesting to note this phenotype was not transmitted to offspring, suggesting that the syndrome resulted from epigenetic mechanisms operating in clones but blunted by sexual reproduction (Tamashiro *et al.*, 2002). This supports the assumption that cloned mammals might be essentially abnormal (Wilmut, 2002). The precautionary principle need not be invoked for reproductive human cloning since the risks are real and high.

Cloning techniques are presently quite empirical and may stay that way for a long time (Western and Surani, 2002). Unidentified details in cloning protocol seem to be important for obtaining a high yield of cloning. It is striking to observe that two experimenters working simultaneously in the same laboratory, using the same biological material, may have quite different success in cloning (Perry and Wakayama, 2002). It is also surprising that cloned cells used as nuclear donors may generate a different number of clones, depending on culture conditions before nuclear transfer.

Clearly, a huge amount of work is needed before any reproductive cloning can be reasonably envisaged in humans. This opinion is shared by the vast majority of scientists. It was expressed with no ambiguity by two of the leading researchers in this field (Jaenisch and Wilmut, 2001). The very small number of people who pretend to be able to clone humans in the very near future are not really aware of the scientific and technical state of the art in this field or they are just impostors who are trying to make money or who are in search of glory.

One historical consideration should not be overlooked. The first animal cloning performed on a laboratory toad more than 40 years ago had no other goal than basic research. The cloning of sheep undertaken about 15 years ago aimed at improving animal selection. The purpose of the project that resulted in Dolly's birth was to tentatively simplify transgenesis techniques. Hence, in no case did the researchers aim at cloning humans. Many people have the opposite view of reality. They believe that a number of laboratories are running a race to be the first in cloning humans. Many people are also convinced that human cloning is inevitable. They obviously underestimate the technical difficulty and base their opinion on the idea that a technique that is imperfect at the beginning can be improved and will inevitably meet with success. This is clearly not the case for cloning. It is not possible to know whether cloning techniques may be optimized to become an acceptable medical treatment, independent of ethical considerations. Thus, public opinion overestimates the real threat of human reproductive cloning at this time.

The major ethical problem of reproductive cloning in humans is that the newborn will be genetically identical to a living adult. This unique new situation could generate identity problems in the cloned persons. The confrontation with an adult having the same genotype may obviously be uncomfortable for a child. Strangely enough, the reverse situation is rarely evoked. An adult may feel humiliated to discover that a young person with his or her genotype is more successful in life than the model.

It is conceivable that this confrontation would not generate problems but, on the contrary, would instead create an exceptionally harmonious relationship between the adult and the cloned child. The probability of problems seems higher than the reverse. This is one of the reasons why reproductive cloning in humans is generally not considered to be ethically acceptable.

The number of situations in which reproduction by cloning is medically justified is small. In many cases, sterility can be overcome by *in vitro*

fertilization including ICSI, as well as by embryo, oocyte and sperm donation. Child adoption is another possibility. The space left for reproduction by cloning is therefore narrow.

A particular situation might perhaps be theoretically acceptable. The genome of individuals still living or not, and not subject to genetic disease, might be used by sterile couples to generate children. This is comparable to embryo, oocyte and sperm donation. It could be called genome donation. This would be preferable, of course, if it were done anonymously, as it is for sperm donation, and without any intention of eugenism.

The question of the identity of the clones is also important. Obviously, cloned individuals are essentially genetically similar to their genitors. However, the real genotype of a somatic cell used as a nuclear donor is not known. On the other hand, the mitochondrial genome of clones is essentially that of the oocyte. The cloned individuals are thus chimaeric organisms. Moreover, epigenetic phenomena may explain why cloned animals are not totally similar. Furthermore, the fate of a human being is not entirely determined by his or her genome. Cloning could be theoretically used as a treatment for sterility but not to satisfy the whim of a person who hopes to become immortal. Indeed, the first situation is a medical act, whereas the second is not. In any case, the genitor would probably be disappointed by the replicate, once born.

4.1.2 Reproductive cloning in animals

Reproductive cloning has become a reality to facilitate transgenesis in several species (see Section 3.1). The number of founder transgenic animals to be generated is generally low in this case and imperfections in cloning techniques are acceptable.

Systematic reproduction of farm animals by cloning cannot be presently envisaged because of the low yields inherent in the technique. The cloning of precious animals is more easily justified. This situation is being evaluated for high-value genitors that have become unable to reproduce. This practice may lead to a reduction of biodiversity and potentially endanger breeds. This problem can be easily avoided since the genetic management of farm animals is closely controlled.

Reproductive cloning may be implemented to save endangered species that have become unable to multiply by themselves at a sufficient rate. In this case, as in the case of farm animal breeds, reproduction by cloning of

adults may generate a relatively small number of animals that are genetically identical. The fate of a herd derived from such animals may be quite uncertain because of their considerably reduced biodiversity.

The cloning of pets has become a reality since the birth of a cloned cat. This may give the owner the impression that a dead loved animal is still alive. The low yield of cloning means that numerous oocytes must be collected from adult females and that multiple recipients will be used. This may inflict significant suffering on the animals or even the sacrifice of unwanted animals. A distinction must be made between cloning to generate individual pets for one person and cloning to generate reproducers used to improve a breed of pets. The ethical acceptability appears lower in the first case than in the second.

4.1.3 Therapeutic cloning

This approach, described in Section 3.1, is essentially aimed at obtaining stem cells for cell therapy in humans.

Other possibilities are being studied to reach similar goals. It is not yet known which of these techniques will be the most appropriate to become a real therapy. Therefore, the success of human cloning does not appear currently to be a critical step. This might become the case in several years if it becomes clear that cloning is the only way to generate stem cells for humans. In the meantime, researchers have much work to do with animal models, including non-human higher primates, to improve cloning techniques.

They also have enough spare human embryos resulting from *in vitro* fertilization and kept frozen to define optimal conditions to establish stem cell lines. The ethical guidelines for the generation of stem cells from spare embryos and for therapeutic cloning are particularly confusing. Quite different situations are found in countries having distinct legislation on research on human embryos. The establishment of embryo stem cell lines is authorized in the USA in laboratories supported by private financial interests but not in those receiving public funding. In some countries, experimentation with embryo stem cell lines is authorized but not the establishment of the lines.

A present trend in thinking tends to establish a clear distinction between reproductive and therapeutic cloning. The first is considered as unacceptable while the second is supposed to raise no ethical problems. To avoid confusion, the expression 'nuclear transplantation' has been

proposed to replace therapeutic cloning (Vogelstein, Alberts and Shine, 2002).

In its present version, human cloning implies the use of numerous woman oocytes. The collection of this biological material is an ethical problem *per se*.

On the other hand, viable embryos are generated only for the purpose of creating stem cells. Their development is therefore interrupted at the blastocyst stage to establish stem cell lines. This protocol is considered to be ethically acceptable when is can be used to cure human diseases.

This appears to be a rather strange way of looking at the situation. An embryo obtained by cloning is an embryo proper since it can generate a living organism. Some people pretend that embryos obtained by cloning are not really embryos but just cellular artifacts. If this is true, Dolly is not a sheep. An embryo is defined by its capacity to develop and give rise to a normal living organism and not by the method used for generating it. The fact that the yield of cloning is lower than that of fertilization cannot justify the idea that therapeutic cloning does not raise ethical problems. The argument that embryos were generated for cell therapy but not for reproduction does not in any way change the fact that viable embryos were created to be destroyed.

In a few years, when the evaluation of the different approaches to obtain human stem cells is completed, therapeutic cloning may appear necessary and acceptable. The ethical price to pay will remain the generation and the deliberate destruction of viable human embryos.

The main subject of dispute in this field is the status of human embryos. For some people, a human embryo is a human being and thus sacred. For other people, a human embryo is nothing more than a cell aggregate until a project has been defined by its parents. An intermediate position consists of considering that a human embryo is not a human being but it should not be treated as a mouse embryo. Our ethical values are often dictated by practical considerations. Life and embryos have been traditionally considered as sacred since they were relatively rare and precious. Modern societies now favour contraception as much as fertility. This may result in a deep change in our view on human embryo status.

4.2 Limits and Risks of Gene Therapy

The technical problems of gene therapy and also the hope raised by the development of new techniques have been discussed in Section 2.2.

The specific ethical problems of gene therapy arise from its complexity and its novelty. Indeed, the implemented tools are sophisticated and, in many cases, based on the use of viral vectors. Gene therapy uses biological material, which can be assimilated to classical pharmaceuticals but only up to a certain point. The effects and side-effects of gene therapy are obviously more complex than those generated by most chemical drugs. The guidelines for gene therapy have not yet been standardized. They have to be adapted to technical advances. This leaves room for risks which cannot be easily evaluated.

Gene therapy has received considerable financial support for a decade without giving satisfactory results. The pressure to accelerate trials in patients is strong. This has instigated some experimenters to proceed with trials under poorly controlled conditions, resulting in the death of patients under unacceptable conditions (Zallen, 2000). This experimental field is now under more intense scrutiny from scientists, governmental commissions and the public.

One particular concern about gene therapy is the possible transfer of genes to the germ cells of patients. Indeed, most of the vectors used at this time are poorly targeted, if at all. In one case, it was reported that adenoviral vectors transmitted the foreign gene to a patient's sperm precursors at low frequency. It was not mentioned that this event was reproducible and that the vector was integrated into the germ cell genome. This observation reinforced the idea that this technology must be carefully controlled.

In a recent publication, it was shown that lentiviral vectors were able to transfer a gene in testis Sertoli cells and correct a genetic disease in mice. The sterile animals became fertile, while the foreign gene was not transferred to germ cells (Ikawa *et al.*, 2002). This approach could be extended to humans. The children born after such gene therapy would have every chance of being sterile. This raises an ethical problem, which should be taken into consideration.

4.3 Limits and Risks of Transgenesis

4.3.1 Technical and theoretical limits

The data reported in Chapter 2 summarize the difficulties encountered by researchers in generating transgenic animals as well as in obtaining a satisfactory expression of the transgenes. These difficulties are progres-

sively reduced and improvement can be reasonably expected in this domain.

Theoretical limits in the use of transgenesis will remain. Numerous experimental data show that the addition of a foreign gene or a gene inactivation may have deleterious effects and even induce death in animals. This may greatly complicate interpretation of the results.

More surprisingly, gene knock-out is often followed by no apparent phenotypic effects. Up to 30 per cent of the mice having an inactivated gene pursue a normal life. This was unexpected, especially for genes that have been highly conserved during evolution. In some cases, the absence of a gene is compensated for by the activation of alternative mechanisms. Living organisms have developed redundant conserved mechanisms acting as a safeguard in case of mutation. The role of the experimentally inactivated gene cannot be easily identified under these conditions. This is the inevitable price to be paid for the study of a gene in a complex situation.

In other cases, gene inactivation has effects that are not observed in mice kept in cages. Indeed, these animals are assisted. They do not have to find their food, to defend themselves against cold, heat stress, to search for a sexual partner, to save themselves from destruction by predators etc. Hence, genes that are essential for life in the wild and thus expected to be highly conserved show no visible effect in breeding facilities. More and more sophisticated tests aimed at revealing abnormalities in mutants are being used.

A researcher is inclined to study the role of a given gene after having observed that its expression is modified under specific physiological conditions. The idea that expression of a gene occurs because it is useful for the organism prevails in the mind of researchers. In higher organisms, gene regulation is extremely complex and not necessarily fully optimized during evolution. It is tempting to think that the expression of a gene should progressively disappear during evolution if it is not mandatory for the organism. Its expression has every chance of being suppressed if it has a negative effect, but its expression may be neutral and thus not eliminated. This may complicate the work of researchers, who may follow the wrong clues without imagining that this is the case.

4.3.2 Biosafety problems in confined areas

It is generally admitted that genetic engineering and gene transfer have no fundamental reasons to generate living organisms more dangerous

than those found in nature. The conditions necessary to manipulate different pathogens in confined areas were defined decades ago. They are considered as appropriate for genetically modified organisms. The problems with the safe manipulation of genetically modified organisms is not to define fundamentally new rules but to determine how and when to follow them. This means that the commissions controlling experiments involving genetically modified organisms have no reason to forbid any experiments. Their role is to evaluate the potential risk on a case by case basis and to determine the confinement under which experiments must be performed, accordingly.

Pathogens and infected organisms have been divided into four classes. In the first class, the living organisms show no risk for experimenters and the environment. In the second class, the living organisms are mild pathogens considered to be moderately dangerous. The diseases they can induce are not severe, and can be cured and prevented. The living organisms of the third class are highly dangerous and those of class four extremely dangerous.

The living organisms of class 1 can be manipulated without any particular precautions. The living organisms of class 2 must be manipulated under an appropriate safety hood, using gloves. All of the biological material must be inactivated by autoclaving or another method at the end of the experiment. The laboratory in which class 2 organisms are manipulated must just be isolated from other rooms.

The organisms of class 3 must be manipulated in special areas under negative pressure. Air must be evacuated from the laboratory through HEPA filters capable of trapping all microorganisms and viruses. An airlock should separate the laboratory from the rest of the building. Experimenters must take uncontaminated clothes into the airlock and take a shower before leaving the laboratory. All of the biological material must be inactivated in an autoclave with one door in the laboratory and the other door on the other side of the wall.

In class 4, the living organisms must be manipulated in a glove box maintained under negative pressure. The box must be in a laboratory essentially similar to those used in class 3. The experimenters must use special clothes providing complete isolation and protection.

This classification, originally defined for natural pathogens, has been extended to genetically modified organisms. As far as transgenic animals are concerned, the following rules have been retained.

– The animals of class 1 show no risk. They can be bred in conventional facilities. They must be specifically labelled as GMOs. Their breeding conditions must prevent any uncontrolled reproduction with other animals. At the end of the experiments, the animals must be sacrificed to make sure that they will not be reproduced.

The confinement of transgenic animals must be defined according to the living habits of each species. Conventional facilities are appropriate for laboratory rodents. Specific devices must be imagined and used for some species to prevent any dissemination of transgenes. In the case of aquatic animals, filters and grids must be used to prevent transgenics from escaping. Water that may contain gametes or young animals must be sterilized by adding appropriate chemical compounds before being rejected. Specific grids or filters must be used to prevent dissemination of air-borne transgenic animals.

– The transgenic animals of class 2 (harbouring mild pathogens) must be kept in cages having a lid with a filter preventing dissemination of micro-organisms. The cages must be opened only under a safety hood. All the biological material and waste, including animals, must be inactivated by autoclaving or other methods.

– The transgenic animals of class 3 must be kept in facilities under negative pressure and manipulated like class 3 pathogens.

– The rules defined decades ago to manipulate class 4 non-transgenic animals are the same as those to be followed for class 4 transgenic animals.

These guidelines, which have been adopted by all experimenters, may be considered as satisfactory, since no accident has been reported after more than 15 years of intense manipulation of transgenic animals.

4.3.3 The intentional dissemination of transgenic animals into the environment

No transgenic animals have been intentionally disseminated into the environment so far, but this situation will logically occur some day. The problems raised by such a practice are fundamentally similar to those encountered with genetically modified plants.

Farm animals (poultry, rabbits, pigs, and ruminants) are usually bred in confined areas. The breeding of transgenic farm animals is not expected to pose a threat for the environment. Domesticated farm animals and wild animals do not meet and have no chance of interbreeding.

The situation is far more complex for aquatic animals. Most of these animals can escape, live and reproduce in wild water. Some of them can even interbreed with wild animals, since they have not yet been intensively selected. The uncontrolled dissemination of their transgenes is therefore possible for these animals.

Studies with model animals have been performed to tentatively evaluate the risks of transgene dissemination. American catfish harbouring an additional growth hormone gene have accelerated growth on farms. Some of these animals were released in a pond under well-controlled conditions. The development of the transgenic animal was compared to that of the wild catfish. Unexpectedly, the transgenic catfish showed delayed growth under wild conditions. One possible explanation for this surprising result is that the transgenic catfish have a metabolism that is overstimulated by growth hormone. This is an advantage for bred animals who do not to have catch their own food. The transgenic catfish may need more food than the wild animals. They are unable to catch more food than the wild fish and they suffer from underfeeding. In this particular case, any escape of transgenic catfish into wild waters cannot be considered as a threat for the environment.

Another experiment carried out with transgenic medaka has led to quite different conclusions. The additional growth hormone gene in these animals accelerates their growth and thus their sexual maturity. The transgenic medaka reproduce earlier than the control animals. This might lead to a progressive invasion of the wild water by the transgenic animals. On the other hand, it has been observed that the additional growth hormone gene shortens the life of medaka. The initial invasion by transgenic medaka might be followed by the progressive local extinction of this species (Muir and Howard, 2002).

The scenario is probably more complicated than this model (Hackett, 2002). Yet, the model argues strongly in favour of the idea that the dissemination of transgenic fish with accelerated growth may have a detrimental and uncontrolled effect on wild animals. These data, although imprecise, justify the fact that the release of transgenic fish with accelerated growth is not allowed at this time. It should be noted that up to 30 per cent of salmon escape from their cages (McDowell, 2002). Other breeding conditions must therefore be found for transgenic fish.

Fish farming is becoming increasingly important. Consumer demand is growing. One of the reasons may be that fish is recommended by nutritionists since it contains anti-atherosclerosis lipids. Fishing activities are more and more limited by the decreasing number of fish in the sea. Fish farming techniques are making rapid progress and an increasing number of marine species are being bred. The productivity of fish farms is also increasing due to better knowledge of fish nutrition, improved control of fish diseases and genetic selection, as well, which started only recently. The fish bred in farms are thus still genetically close to those in the wild. This is one of the reasons why wild fish can cross-breed with those living on fish farms. This is also the reason why an increase in yield is expected through genetic selection and why transgenesis may contribute to improving fish production.

To prevent the dissemination of transgenic fish in the wild, several approaches are theoretically possible. One consists of sterilizing the animals by rendering them triploid. This can be achieved by subjecting one-cell embryos to hyperbar or heat shock, which alters chromosome dispatching in daughter cells. The triploid animals experience normal growth but they are sterile. However, the process is not totally efficient and a few animals remain fertile. This precludes this method from being used to prevent transgene dissemination in wild water.

Another possibility might be to sterilize the animals by transgenesis. The expression of antisense RNA directed against the mRNA of gonadotropin hormone has been shown to inhibit the synthesis of this hormone in trout. Fish cannot reproduce unless they are treated by exogenous gonadotropin hormone. This trait is inheritable and the method might be used to control transgene dissemination. Its feasibility has not yet been completely determined

The best solution may be quite different. The breeding of transgenic fish in isolated lakes or bays has been envisaged. However, biosafety did not appear to be guaranteed in any of the cases and no project along these lines has actually been undertaken. The breeding of most domesticated fish can be achieved in closed tanks. This implies somewhat expensive facilities and a renewal of water, which is costly. There are advantages to this approach, independently of transgenesis. Fish farming is an important source of pollution. The systematic breeding of fish in closed areas would solve these problems and the problem of transgene dissemination as well. This approach would probably also protect fish from diseases transmitted by water in which wild fish live.

4.3.4 The risks for human consumers

It appears inevitable that transgenic animals will be proposed to human consumers as is already the case for transgenic plants and microorganisms. The introduction of a foreign gene into a genome is a kind of mutation that may alter nutritional properties or the safety of the food stuffs.

The introduction of a foreign gene is generally non-targeted. The foreign DNA may inactivate or activate an endogenous gene. On the other hand, the product of the foreign gene may be toxic or allergenic. These points must be taken into consideration, as is already the case for transgenic plants.

Until the structure of the whole genome of domestic animals is known, it will be difficult to determine whether the foreign gene has interrupted an endogenous gene. The product of a transgene may be tested to evaluate its possible toxic and allergenic properties. The same evaluation may be performed with food stuffs containing the product of the transgene.

The evaluation of toxicity is expected to be quite reliable. Indeed, these kinds of test are routinely used to measure the toxicity of various products used by humans, namely pharmaceuticals. It should be noted that a toxic substance resulting from the presence of a transgene has a good chance of affecting animals, a fact that can be easily observed. This is usually not the case for transgenic plants.

The same is not true for allergens. It is well known that animals do not develop immune reactions against the products of their transgenes since they are an integral part of the organism from birth. Allergenicity of a compound cannot be identified with certainty in all cases, especially if it is weak and affects a small proportion of people.

The substantial equivalence has been retained as a good criterion by risk-assessment committees. This method consists of comparing the overall composition of a transgenic organism with that of a non-transgenic control. This approach is rather imprecise and superficial. DNA and protein chips might reveal differences between transgenic and non-transgenic organisms in the future. It may take a long time before these tests are reliable. On the other hand, it may be quite difficult to predict the effects of some of the differences induced the presence of transgenes.

All in all, global evaluations based on measurements of toxicity and allergenicity may be the most relevant approach. Indeed, these tests integrate the different modifications brought about by the transgene.

A biosafety problem that is specific to animals and not encountered with plants is possible increased sensitivity towards pathogens, which might affect breeding and humans. Such events in plants might have an impact only on agriculture.

The emergence of an animal virus favoured by the presence of a transgene may not be observable in its initial stages, but, rather, only after it has reached epidemic proportions.

All these events are unlikely but could possibly occur. Labelling and traceability, in particular, appear to be the best ways to identify a problem resulting from the consumption of genetically modified animals.

An increasing number of cloned farm animals are alive and their consumption by human beings has been questioned. Indeed, in the majority of the cases, the cloned animals are not supposed to be genetically modified but they are often epigenetically modified instead. This may alter animal products in ways that are more or less similar to transgenesis. The consumption of products (meat and milk) from cloned farm animals is not authorized in some countries whereas no restrictions exist in others. A possible compromise might be to allow consumption of clone offspring but not of clone founders. This proposition is based on the fact that a certain number of cloned animals have health problems that are not transmitted to progeny.

4.3.5 Transgenesis and animal welfare

Some people are convinced that transgenesis, by its very nature, generates monsters that are born to suffer. Any observer who has visited transgenic animal facilities will be rapidly convinced that the reality is more subtle.

The first ethical problem concerning transgenesis is that it implies that animals are genetically transformed. Some people are strictly against such a practice, considering that it is a violation of the laws of nature. These people want to know if a drug used as a pharmaceutical has been prepared from transgenic animals or on the basis of knowledge obtained through transgenics. The first situation may be easily identified, whereas this is not true for the second. Knowledge is a complex activity, which cannot be separated into independent domains. In practice, the vast majority of people do not reject the idea of experimental genetic modification in animals *per se*. They realize that major physiological modifications resulting in suffering in some cases have been made to establish lines of

animals by selection, namely of household pets. These genetic modifications, which are numerous and generally unknown, are accepted. On the other hand, modifications having a similar negative impact animal welfare would not be accepted if they resulted from transgenesis.

Animals' rights are often evoked in an attempt to reduce their suffering. This principle is far from being universally accepted. Indeed, animals have a quite limited capacity to determine their own behaviour. This lack of freedom is not considered to be compatible with animals' rights. Rather, socially conscious human beings have a responsibility towards animals. This point of view is probably more easily acceptable, and may ensure respect for animals just as well.

Transgenesis induces suffering in animals for different reasons. The first is the process of gene transfer itself. Indeed, the collection of embryos for microinjection, and embryo transfer to recipient females, imply either surgery or the sacrifice of the animals themselves. A number of embryos do not survive gene transfer. On the other hand, ill constructed genes may be poorly expressed in transgenic animals which cannot then be used for experimentation.

To reduce these problems, the European Centre for the Validation of Alternative Methods (ECVAM) recommended that the '3R rule', defined years ago for conventional laboratory animals, be applied. This rule implies that researchers reduce the number of experimental animals, refine the experimental protocols and replace the use of animals, whenever possible, by *in vitro* tests, and cell cultures, in particular. The list of alternatives to the use of experimental animals including transgenics is growing (Moore, 2001).

It was also recommended that the generation of transgenic animals unable to produce relevant information be avoided. Reliable gene constructions are required to reduce the number of animal lines that do not properly express their transgenes. Another point should also be emphasized. A significant proportion of transgenic animals produce no interesting results, at least for the researchers who generated them. These kinds of result are generally not published and other researchers independently generate similar animals, leading to the same outcome. On the other hand, a transgenic animal may not meet the expectations of the researchers who generated them but they might be a precious tool for others not working in the same field. Better knowledge of the data, even of the negative results, would reduce the number of transgenic animals needlessly produced and generally serving no purpose (Ben Mepham *et al.*, 1998).

To tentatively reduce the use of transgenic animals, some people have proposed the application of the 'no unless' principle whereas others recommend the 'yes but' principle. In the first case, transgenesis is fundamentally banished but exceptionally acceptable. In the second case, transgenesis is basically allowed but must be used with parsimony.

These principles are quite general and do not really help experimenters in their choice. To clarify this situation, transgenic animals might be divided into three categories.

The first category includes experimental animals. These animals are essentially used to acquire knowledge and not to make a profit, at least not directly. For this reason, the effect of a transgene on animal welfare cannot be predicted in all cases. Moreover, the number of animals used for each experiment is limited. Tolerance towards animal suffering might be relatively high in these cases.

The animals in the second category are those that produce highly valued products for human health. The typical animals in this category are those that produce pharmaceuticals in their milk and those used as a source of organs and cells for human therapy. In these situations, the side-effects of transgenes are known and reproducible. These animals are used to alleviate human suffering but they may also be a source of high profit for industry. The suffering of these animals may be acceptable but only on a case-by-case basis, taking all of the advantages and drawbacks for humans and animals into account.

The third category of animals might include farm animals. The side-effects of the transgenes are then perfectly well known and reproducible. These animals may be useful for feeding humans but not mandatory. Their breeding is justified, to a large extent, by the profit they generate. In this case, no animal suffering would be acceptable.

This classification should not just be applied to transgenics but also to conventional, experimental and farm animals or even pets.

Ethical committees, either local, national or at the European Union level, are increasingly solicited to examine the ethical aspects of protocols involving the use of transgenic animals. Industry and investors are also progressively adopting this practice.

4.3.6 Patenting of transgenic animals

Patenting living organisms is a traditional practice but has been restricted for decades to microorganisms. Animal patenting has raised new

problems. Some people consider that patenting is not compatible with the respect due to animals. In practice, the fact that an animal is patented will not increase its suffering. The animals will be used for experimentation if they can be the source of interesting information, independently of their patenting. Patenting is now permitted in the European Union as well as in the USA. The European Union has added a restriction. An experimental transgenic animal cannot be patented if the potential benefit derived from it is not expected to improve the fate of humankind. Transgenic mice generated to serve as models for the study of human baldness were not accepted by the European Patent Office, which considered that baldness is not a human disease.

The number of cases in which the patenting of experimental animals is justified is not great. Indeed, the effect of a transgene is generally repeatable and a gene construct capable of generating a relevant animal model should be patented rather than the animals themselves.

Yet, it may happen that a transgenic animal is unique due to the particular expression of the transgene integrated in a non-anticipated site of the genome. In this case, the patenting of the animals may be more easily justified.

The patenting of farm animals raises quite different problems. The addition of a transgene must, in no case, be a pretext to patent an animal breed resulting from centuries of selection by farmers. Selected seeds are protected by specific licences that are not the same as patents. The same system does not exist for animals at this time and could possibly be extended to them. At most, only a new trait resulting from the action of a transgene should be the object of a patent. As in the case of experimental animals, the patenting of a gene construct transferred to a given animal breed should protect the inventors in an appropriate manner (Figure 4.1).

4.3.7 Transgenesis in humans

It may appear strange that transgenesis has never been attempted in humans and hardly discussed. Indeed, even if the classical techniques for gene transfer are poorly efficient, they might be extended to humans with good chance of success.

This may be due to the fact that transgenesis is not authorized for humans. The lack of relevant genes to be transferred is another strong reason.

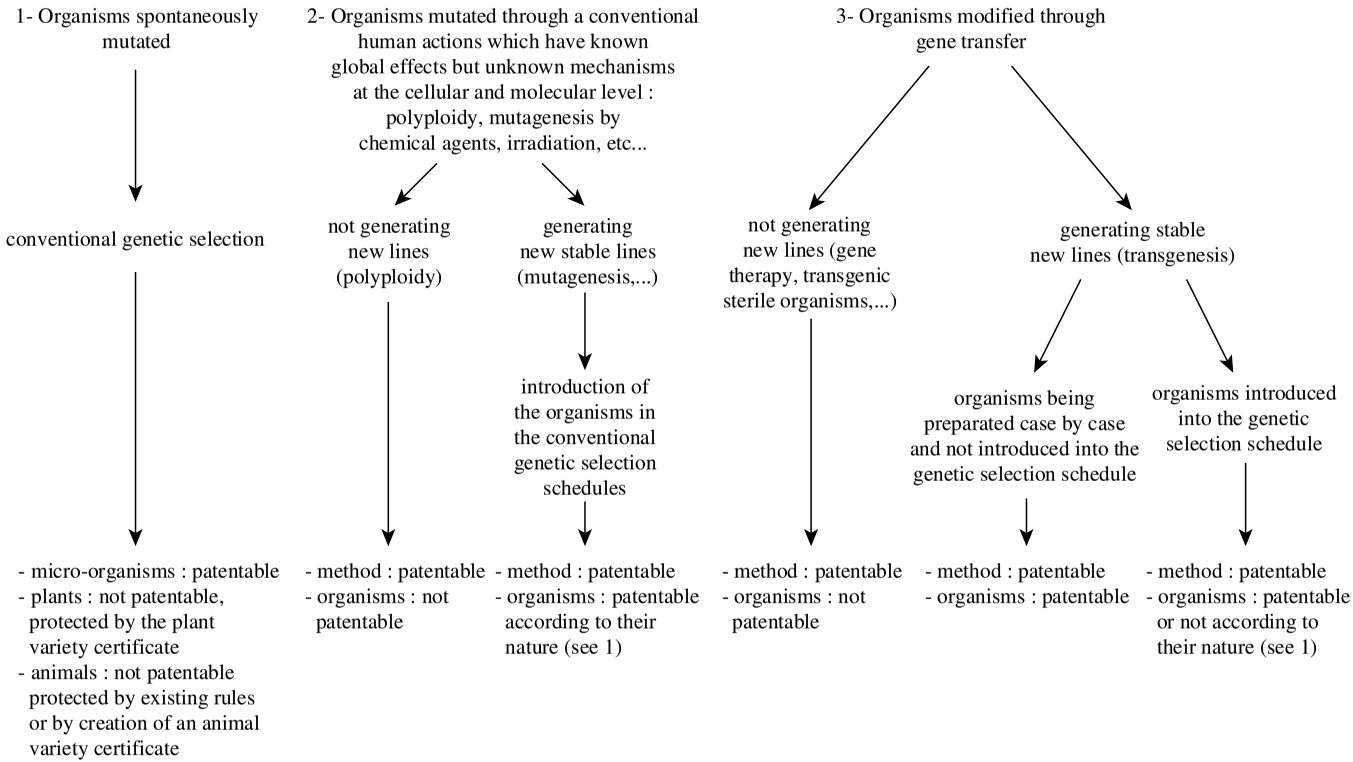


Figure 4.1 Patentability of genetically modified organisms (L.M. Houdebine Cahiers/Agricultures (1993) 2: 343–345 © John Libbey Eurotext)

In order to make this possible, transgenesis techniques should be considerably improved to leave no chance for unpredictable events. This is far from the case at this time but it is possible to imagine that this will become a reality in the future.

A fundamental question remains. The effect of a transgene, even perfectly controlled, cannot be fully predicted. Genes have numerous subtle inferences with multiple mechanisms in the organism. Unwanted side-effects may result from gene transfer late in the life of the individual or in his progeny. On the other hand, some effects of a transgene are known to be dependent on the genetic background of each individual organism. Hence, the long-term effects of transgenes should be evaluated in humans before being used as a therapy. This is clearly impossible.

The Cre–LoxP system may be used to eliminate a transgene. A gene transfer in humans cannot be conceived as reversible if there is a problem, with a system like Cre–LoxP. Gene transfer in humans must be absolutely safe or not permitted at all.

A particular case is the mutation of alleles responsible for sensitivity to a specific disease. It is conceivable that targeted mutation be implemented to modify the faulty alleles. No side-effects would then be expected, since the genetic correction would aim at restoring an allele naturally present in many human beings. However, an unknown genome modification occurring with the targeted allele correction cannot be completely ruled out. Even if all the technical problems are solved, the correction of alleles will probably always appear as exceedingly sophisticated and risky in comparison to embryo selection.

A day may come when the human species decides to modify its genome to improve its life. These modifications might even deliberately alter some aspects of human nature. The natural complexity of human beings would not make it possible to fully predict the effects of transgenes. This would be a leap into the unknown and a real gamble. The present state of the art leaves no reasonable chance for human transgenesis to become a reality in the near future.

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

Conclusion and perspectives

Present progress makes it possible to make some predictions for the use of animal transgenesis and cloning. Others are obviously impossible in a field where things are moving so rapidly.

Cloning techniques are expected to improve significantly. This is exemplified by the recent success with rabbit cloning (Chesné *et al*., 2002). Subtle modifications in the protocol respecting rabbit specificity led to this success. Rabbits are characterized by the rapid kinetics of their embryo cell cycle and by a narrow window of time for embryo implantation. The modifications taking these specificities into account are not expected to be extrapolated, as such, to other species but they do invite researchers to address cloning differently in order to increase their chances of success. The possibility of cloning rabbits opens new avenues to study differentiation and dedifferentiation but also to create relevant animal models to study human diseases.

Cloning techniques should reach a level of reliability and efficiency compatible with the reproduction of farm animals and pets. The possible extension to humans, independently of ethical considerations, is unpredictable.

To avoid the use of human oocytes for cloning, the possibility of using cow oocytes has been envisaged. It is by no means certain that the compatibility between the two species would be sufficient to allow the normal development of embryos. One of the problems would be the presence of cow mitochondria transmitted by the oocytes. The replacement of cow mitochondria by human mitochondria might be a possible approach (Mollard, Denham and Thomson, 2002). The use of cow oocytes might be compatible with therapeutic cloning, which implies only stem cell differentiation and not embryo development.

The future success of cloning might come from the implementation of new techniques. Indeed, better knowledge of the factors inducing somatic cell dedifferentiation in oocytes might lead to the introduction of these factors in somatic cells rather than the transfer of somatic nuclei into enucleated oocytes (Landsverk *et al*., 2002; Hakelien *et al*., 2002).

One aspect may remain quite difficult to control. Genome imprinting is essential for normal embryo development. It is a known fact that after fertilization oocyte factors induce an imprinting of the paternal genome while modifying the maternal genome only moderately. This subtle balance is absolutely required and seems to be controlled by the action of just a few genes. Paradoxically, it seems that one of the major reasons for cloning failure is the inappropriate genome reprogramming due to abnormal DNA demethylation (Xue et al., 2002). Yet, the success of cloning using nuclei from somatic cells might be possible only because the parental imprinting balance for essential genes is maintained in embryos generated by nuclear transfer (Surani, 2002).

One possibility might be that genome reprogramming did not occur similarly in all cell types, giving rise to mosaic individuals. This kind of phenomenon has been shown to be the source of numerous diseases (Pearson, 2002).

Human reproductive cloning is massively rejected by the public and researchers. Yet, some scientists claim that they have undertaken reproductive cloning. Their chances of success appear quite low and these attempts could give rise to non-viable or unhealthy babies. If this did occur, there is little doubt that science would be accused of indulging in high-risk and ill controlled activities. The fact that the scientific community is against human reproductive cloning (Jaenisch and Wilmut, 2001) must be kept in mind, and that the groups attempting to clone humans are obliged to do so in countries where legislation in this field is non-existent. It would therefore be unfortunate if a practice that has met with such widespread disapproval turned into an accusation against science and scientists alike.

These rare but uncontrolled activities have led some politicians to prohibit both human reproductive and therapeutic cloning. They consider that accepting therapeutic cloning is just one step close to reproductive cloning. Clearly, even if therapeutic cloning raises ethical problems, they are quite different from those generated by reproductive cloning (Vogelstein *et al*., 2002; Powledge, 2002).

There is no doubt that therapeutic cloning implies the voluntary destruction of viable embryos created for this purpose, even if some

people claim the contrary. Such an approach may be considered as unacceptable. It should be kept in mind that the widely accepted practice of *in vitro* fertilization implies the generation of spare embryos, whose normal fate is destruction. This practice tends to be lessened with the improved effectiveness of *in vitro* fertilization, but it has a long way to go before it disappears for ever.

Therapeutic cloning is the focus of great hope for humans, despite the numerous uncertainties involved. Cloning problems might have more limited consequences for therapeutic than for reproductive cloning. This is especially the case for genome imprinting, since fewer functional genes are required for the normal working of a differentiated cell than for embryo development.

Therapeutic cloning has been attempted in mice and has met with success. Interestingly, the cells used as nuclear donors were genetically corrected by gene transfection. The recipient mice, which were mutants, were thus simultaneously subjected to both cell and gene therapy, since they recovered their normal phenotype (Rideout *et al*., 2002).

A quite demonstrative study has recently shown that differentiated B and T lymphocytes used as nuclear donors gave rise to cloned mice harbouring mature immunoglobulin and T cell receptor genes used as markers (Hochedlinger and Jaenish, 2002). This clearly shows that fully differentiated cells can become pluripotent after nuclear transfer in enucleated oocytes. Howover, clones were obtained in these cases only when the cells of the cloned embryos were inserted into recipient blastocysts. Hence, the B and T lymphocytes did not become totipotent cells after nuclear transfer but only pluripotent cells. This raises the question of whether cloning is successful using cells from organs only with the stem cells present in the organs (Colman, 2002; Rossant, 2002).

A recent study showed that cells from a cloned cow were not rejected when transplanted into the animal that was the nuclear donor. This means that the proteins coded by the mitochondrial genome of the recipient oocyte were not taken as foreign antigens (Lanza *et al*., 2002a).

A systematic comparison of the genetic stability of ES cells and differentiated fibroblasts revealed that the former mutate less frequently than the latter, whereas chromosome stability is lower in ES cells (Cervantes *et al*., 2002) This means that differentiated cells obtained from cloned embryos after transfer of somatic nuclei might have a higher propensity to give rise to tumours than those derived from ES cells (Freed, 2002).

On the other hand, it is not yet known, for example, to what extent neurons capable of stably secreting dopamine can be derived from embryonic stem cells (Freed, 2002).

Transdifferentiation of organ stem cells in other organ stem cells is extremely promising and could replace the therapeutic cloning approach in some cases. However, organ stem cells are rare, not well known and their culture is difficult. Transdifferentiation was observed repeatedly *in vivo* but *in vitro* transdifferentiation has been recently questioned (Wurmser and Gage, 2002). Indeed, it was claimed that somatic cells cocultured with ES cells acquired pluripotency. A reappraisal of this problem revealed that a fusion between the ES and the somatic cells occurred spontaneously in the culture dishes, giving rise to polyploid cells not exploitable for cell therapy. Hence, it appears that the conditions required to dedifferentiate somatic cells in culture have not yet been found.

In a recent study, it was shown that multipotent cells from mouse bone marrow were dedifferentiated to become pluripotent cells when inserted into a blastocyst (Jiang *et al*., 2002). These cells participated in generating all cell types of the animals. These bone marrow cells can easily be cultured for up to 100 passages, remaining stable and not differentiated. This impressive observation confirms that cloning is probably not the only possibility for generating pluripotent cells from the somatic cells of an individual for cell therapy.

It can be predicted that transgenesis in a few animal models will be more systematically used in the future. The majority of the mouse genes might be knocked out in three years. In some cases, conditional knock-out will be achieved. In *Drosphila* and *C. elegans*, the systematic use of gene activation by RNA interference might lead to a situation similar to what is being done in mice with gene knock-out today.

Interestingly, recent reports indicate that short double-strand RNAs synthesized by RNA polymerase III vectors are quite effective to specifically inhibit gene expression in mammalian cells (Tuschl, 2002). This suggests that gene inhibition, equivalent to knock-out, may be obtained with double-strand RNAs inducing the RNA interference mechanism. Several promoters have been used for these studies. The U6 RNA gene promoter has been the most frequently implemented up until now. This promoter is used by RNA polymerase III in all cell types. This implies that RNA interference will occur simultaneously in all cell types. This may limit the interest of the experimental model. A study published several years ago showed that the U6 RNA gene promoter can be

engineered to become sensitive to tetracycline (Ohkawa and Taira, 2000). Hence, the double-strand RNA may be synthesized under the U6 RNA gene promoter only in the cells expressing the tetracycline dependent transcription factor.

This systematic approach implies the study of numerous mice under different conditions in special facilities considered as mouse clinics. Multiple and complementary tests will be performed in these animals without any *a priori* hypothesis to determine the best way to study them in depth. This includes measurements of various hormonal and metabolic parameters as well as behaviour. Imaging will become increasingly frequent, using different techniques adapted to mice: computed tomography, magnetic resonance, single-photon emission computed tomography, positron emission tomography, confocal observation and episcopy fluorescence (Weninger and Mohun, 2002).

Tools aimed at measuring *in situ* protein phosphorylation are becoming available. Different GFPs with different colours and harbouring phosphorylation sites can be observed in single individual cells, allowing the dynamic observation of the mechanisms implying protein phosphorylation (Sato *et al.*, 2002b).

Transgenic animals, and mainly mice, are expected to be an essential tool for genomics-based drug design and discovery (Dean *et al.*, 2001).

Progress has been made but is still needed to obtain a faithful transgene expression. The use of gene transfer by cloning may become more frequent if cloning yield is improved. Gene transfer in cells further used to generate transgenic clones may be better controlled than after DNA microinjection into embryos. However, this approach does not enhance the prediction of transgene expression. Indeed, for unknown reasons, the GFP gene is correctly expressed after having been transferred into cells used for cloning whereas the same gene is subjected to a variegated expression when it has been previously integrated into cells taken from transgenic animals (Park *et al.*, 2002).

More and more sophisticated tools based on the use of the Cre–LoxP system have led to efficient mitotic chromosome recombination in mouse embryonic stem cells (Liu, Jenkins and Copeland, 2002).

The production of recombinant pharmaceutical proteins in milk or other biological fluids does not seem to face major obstacles (Gavin, 2001). It is interesting to note that a human protein, α-glucosidase, extracted from transgenic rabbit milk was capable of treating Pompe disease in babies (Van den Hout *et al.*, 2001). Numerous proteins are

expected to be prepared from transgenic animals. Among these proteins, monoclonal antibodies capable of blocking infectious diseases and replacing antibiotics appear particularly promising (Casadevall, 2002).

Transgenesis applications in breeding are still in the early stages (Houdebine, 2002). Despite significant progress in the techniques to generate transgenic animals and to control transgene expression, additional improvement would be welcome. This approach is not expected to be well received by consumers at this time. It seems reasonable to consider that the biosafety problems will be well controlled. Hostility towards the method *per se* will remain for some time. Strangely, transgenesis is often regarded as contradictory to the improvement of preexisting breeding techniques. It is fully acknowledged by researchers that transgenesis has no reason to be used to solve a breeding problem if other means are possible. Although complicated, transgenesis may be simpler than changing breeding methods. Indeed, modification of habits is often much slower than the implantation of techniques requiring no change in work management. It may thus appear easier to solve the problem of phosphate pollution generated by pig breeding by using transgenic pigs expressing a bacterial phytase gene in their salivary gland than by modifying farm management. It may also be simpler to generate pigs expressing a desaturase gene that lowers the level of saturated lipids in order to decrease the incidence of atherosclerosis than to change consumer habits. Clearly, transgenesis and the improvement of breeding methods are not opposed but complementary.

Both transgenesis and cloning may have a strong impact on humans. To reduce the chance of these techniques being used for unacceptable purposes, research institutes might restrict their use to persons having received specific training in ethics.

It is widely acknowledged that biology and biotechnology are more and more closely linked (Fields, 2001). It is also clear that biotechnologies as well as other activities based on innovation rely on techniques that may generate unknown side-effects (Hoffmann-Riem and Wynne, 2002).

It is often claimed that the hostility of the public towards biotechnologies is due to its ignorance of the challenge. Researchers have certainly neglected this fact. It is not so clear, however, whether the lack of information is the only problem. The poor quality of the information delivered by the media has probably had a strong negative impact (Bucchi and Neresini, 2002).

It seems reasonable to inform the public of research developments and to provoke debates even long before a technique is applied. Some

companies support this view (Lanza, Cibelli and West, 2002b). Scientific information should only be made available to the public when it has reached a sufficient level of credibility. This is not always the case. It was probably premature to publish preliminary data on the first human cloning. The public is often surprised and upset today to discover that scientists disagree among themselves, although this is a basis of the scientific activity. Researchers should restrict their public debates to relevant scientific data.

A majority of people seems to consider that the control of scientific activity is declining, whereas the opposite is obviously true. It is striking to observe, for example, that many debates about reproductive cloning are currently taking place, whereas it is still not certain that this technique can be applied to humans.

The present confusion regarding the debate about biotechnologies and particularly about transgenesis and cloning may have a significant negative impact on research. Researchers have to make a special effort to explain what they do to the public in order to create a better climate and win back the public's trust.

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

References

Adler, S.H., Bensinger, S.J. and Turka, L.A., 2002. 'Stemming the tide of rejection'. *Nat. Med.* **8**, 107–108.

Albuquerque-Silva, J., Vassart, G., Lavinha, J. and Abramowicz, M.J., 2001. 'Chimeraplasty validation'. *Nat. Biotechnol.* **19**, 1011.

Alexeev, V., Igoucheva, O., Domashenko, A., Cotsarelis, G. and Yoon, K., 2000. 'Localized in vivo genotypic and phenotypic correction of the albino mutation in skin by RNA–DNA oligonucleotide'. *Nat. Biotechnol.* **18**, 43–47.

Altuvia, S. and Wagner, E.G., 2000. Switching on and off with RNA. *Proc. Natl Acad. Sci.* USA **97**, 9824–9826.

Amado, R.G. and Chen, I.S., 1999. Lentiviral vectors – the promise of gene therapy within reach? *Science* **285**, 674–676.

Ambros, V., 2001. MicroRNAs: tiny regulators with great potential. *Cell* **107**, 823–826.

Araki, K., Araki, M. and Yamamura, K., 1997. Targeted integration of DNA using mutant lox sites in embryonic stem cells. *Nucleic Acids Res.* **25**, 868–872.

Attal, J., Stinnakre, M.G., Théron, M.C., Terqui, M. and Houdebine, L.-M., 1997. 'The use of episomal vectors for transgenesis'. *Transgenic Animals: Generation and Use*. Amsterdam: Harwood, pp. 251–255.

Baccetti, B. and Spadafora, C., 2000. 'Conclusions'. *Mol. Reprod. Dev.* **56**, 329–330.

Baer, A. and Bode, J., 2001. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr. Opin. Biotechnol.* **12**, 473–480.

Bartek, J. and Lukas, J., 2001. Are all cancer genes equal? *Nature* **411**, 1001–1002.

Bass, B.L., 2001. NA interference. The short answer. *Nature* **411**, 428–429.

Bawden, C.S., Dunn, S.M., Mclaughlan, J., Nesci, A., Powell, B.C., Walker, S.K., Rogers, G.E., 1999. Transgenesis with ovine keratin genes: expression in the sheep wool follicle for fibres with new properties. *Transgenic Res.* **8**, 459–461.

Bell, A.C. and Felsenfeld, G., 1999. Stopped at the border: boundaries and insulators. *Curr. Opin. Genet. Dev.* **9**, 191–198.

Bell, A.C., West, A.G. and Felsenfeld, G., 2001. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* **291**, 447–450.

Ben Mepham, T., Combes, R.D., Balls, M., Barbieri, O., Blokhuis, H.J., Costa, P., Crilly R.E., de Cock Buning, T., Delpire, V.C., O'Hare, M.J., Houdebine, L.M., van Kreijl, C.F., van der Meer, M., Reinhardt, C.A., Wolf, E. and van Zeller, A.M., 1998. The use of tansgenic animals in the European Union. *ATLA* **26**, 21–43.

Bernards, A. and Hariharan, I.K., 2001. Of flies and men – studying human disease in *Drosophila*. *Curr. Opin. Genet. Dev*. **11**, 274–278.

Berns, A., 2001. Cancer. Improved mouse models. *Nature* **410**, 1043–1044.

Billings, P.R., 1999. *In utero* gene therapy: the case against. *Nat. Med*. **5**, 255–256.

Billy, E., Brondani, V., Zhang, H., Muller, U. and Filipowicz, W., 2001. Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc. Natl Acad. Sci*. USA **98**, 14 428–14 433.

Blau, H.M. and Rossi, F.M., 1999. Tet B or not tet B: advances in tetracycline-inducible gene expression. *Proc. Natl Acad. Sci*. USA **96**, 797–799.

Bleck, G.T., White, B.R., Miller, D.J. and Wheeler, M.B., 1998. Production of bovine alpha-lactalbumin in the milk of transgenic pigs. *J. Anim. Sci*. **76**, 3072–3078.

Bösze, S., Hiripi, L., Baranyi, M., Szabo, L., Toth, S., Devinoy, E., 2001. Altering milk quality by transgenesis. In: Toutant, J.P., Balazs, E. (Eds), *Molecular Farming Proceeding of the OCED*, workshop held in La Grande Motte (France), INRA, pp. 29–38.

Bramlage, B., Luzi, E. and Eckstein, F., 1998. Designing ribozymes for the inhibition of gene expression. *Trends Biotechnol*. **16**, 434–438.

Brousseau, M.E. and Hoeg, J.M., 1999. Transgenic rabbits as models for atherosclerosis research. *J. Lipid Res*. **40**, 365–375.

Brower, V., 2001. Gene therapy revisited. In spite of problems and drawbacks, gene therapy moves forward. *EMBO Rep*. **2**, 1064–1065.

Brown, S.D. and Balling, R., 2001. Systematic approaches to mouse mutagenesis. *Curr. Opin. Genet. Dev*. **11**, 268–273.

Brüggerhoff, K., Zakhartchenko, V., Wenigerking, H., Reichenbach, H.D., Prelle, K., Schernthaner, W., Alberio, R., Küchehoff, H., Stojkovic, M., Brem, G., Hiendleder, S. and Wolf, E., 2002. Bovine somatic cell nuclear transfer using recipient oocytes recovered by ovum pick-up: effect of maternal lineage of oocyte donors. *Biol. Reprod*. **66**, 367–373.

Brummelkamp, T.R., Bernards, R. and Agami, R., 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550–553.

Bucchi, M. and Neresini, F., 2002. Biotech remains unloved by the more informed. *Nature* **416**, 261.

Butler, D., 2002. Xenotransplant experts express caution over knockout piglets. *Nature* **415**, 103–104.

Butler, J.E. and Kadonaga, J.T., 2001. Enhancer–promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev*. **15**, 2515–2519.

Campbell, K.S.H., McWhir, J., Ritchie, W.A., Wilmut, I., 1996. Sheep cloned by nuclear transfer from a cultured cell lines. *Nature* **380**, 64–66.

Carvallo, G., Canard, G. and Tucker, D., 1997. Standardization of transgenic lines: from founder to an established animal model. In: Houdebine, L.M. (Ed.), *Transgenic Animal. Generation and Use*. Amsterdam: Harwood, pp. 403–410.

Casadevall, A., 2002. Antibodies for defense against biological attack. *Nat. Biotechnol*. **20**, 114.

Castilla, J., Pintado, B., Sola, I., Sanchez-Morgado, J.M. and Enjuanes, L., 1998. Engineering passive immunity in transgenic mice secreting virus-neutralizing antibodies in milk. *Nat. Biotechnol*. **16**, 349–354.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L., Bousso, P., Deist, F.L. and Fischer. A., 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669–672.

Cecconi, F. and Meyer, B.I., 2000. Gene trap: a way to identify novel genes and unravel their biological function. *FEBS Lett*. **480**, 63–71.

Cervantes, R.B., Stringer, J.R., Shao, C., Tischfield, J.A. and Stambrook, P.J., 2002. Embryonic stem cells and somatic cells differ in mutation frequency and type. *Proc. Natl Acad. Sci*. USA **99**, 3586–3590.

Chan, A.W., Chong, K.Y., Martinovich, C., Simerly, C. and Schatten, G., 2001. Transgenic monkeys produced by retroviral gene transfer into mature oocytes. *Science* **291**, 309–312.

Chan, A.W., Homan, E.J., Ballou, L.U., Burns, J.C. and Bremel, R.D., 1998. Transgenic cattle produced by reverse-transcribed gene transfer in oocytes. *Proc. Natl Acad. Sci*. USA **95**, 14 028–14 033.

Chanat, E., Martin, P. and Ollivier-Bousquet, M., 1999. Alpha(S1)-casein is required for the efficient transport of beta- and kappa-casein from the endoplasmic reticulum to the Golgi apparatus of mammary epithelial cells. *J. Cell Sci*. **112**, 3399–3412.

Chapman, P.F., Falinska, A.M., Knevett, S.G. and Ramsay, M.F., 2001. Genes, models and Alzheimer's disease. *Trends Genet*. **17**, 254–261.

Charreau, B., Tesson, L., Soulillou, J.P., Pourcel, C. and Anegon, I., 1996. Transgenesis in rats: technical aspects and models. *Transgenic Res*. **5**, 223–234.

Chesné, P., Adenot, P.G., Viglietta, C., Baratte, M., Boulanger, L. and Renard, J.P., 2002. Cloned rabbits produced by nuclear transfer from adult somatic cells. *Nat. Biotechnol*. **20**, 366–369.

Cibelli, B.J., Kiessling, A.A., Cunniff, K., Richards, C., Lanza, R.P. and West, M.D., 2001. Somatic cell nuclear transfer in humans: pronuclear and early embryonic development. *J Regenerative Med*. **2**, 25–30.

Cibelli, J.B., Grant, K.A., Chapman, K.B., Cunniff, K., Worst, T., Green, H.L., Walker, S.J., Gutin, P.H., Vilner, L., Tabar, V., Dominko, T., Kane, J., Wettstein, P.J., Lanza, R.P., Studer, L., Vrana, K.E. and West, M.D., 2002a. Parthenogenetic stem cells in nonhuman primates. *Science* **295**, 819.

Cibelli, J.B., Lanza, R.P., West, M.D. and Ezzell, C., 2002b. The first human cloned embryo. *Sci. Am*. **286**, 44–51.

Clarke, A.R., 2002. 'Transgenesis techniques'. *Principles and Protocols*. Second Edition. *Methods in Molecular Biology*. Vol. **180**. Totowa, NJ: Human.

Co, D.O., Borowski, A.H., Leung, J.D., van der Kaa, J., Hengst, S., Platenburg, G.J., Pieper, F.R., Perez, C.F., Jirik, F.R. and Drayer, J.I., 2000. Generation of transgenic mice and germline transmission of a mammalian artificial chromosome introduced into embryos by pronuclear microinjection. *Chromosome Res*. **8**, 183–191.

Cohen, J., 2001. Building a small-animal model for AIDS, block by block. *Science* **293**, 1034–1036.

Cohen-Tannoudji, M., Babinet, C. and Morello, D., 2000a. lacZ and ubiquitously expressed genes: should divorce be pronounced? *Transgenic Res*. **9**, 233–235.

Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D. and Jaisser, F., 1998. I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Mol. Cell. Biol*. **18**, 1444–1448.

Cohen-Tannoudji, M., Vandormael-Pournin, S., Drezen, J., Mercier, P., Babinet, C. and Morello, D., 2000b. lacZ sequences prevent regulated expression of housekeeping genes. *Mech. Dev*. **90**, 29–39.

Colman, A., 2002. Turning back the developmental clock. *Nat. Biotechnol*. **20**, 348–349.

Comeron J.M., 2001. What controls the length of noncoding DNA? *Curr. Opin. Genet. Dev*. **11**, 652–659.

Couzin, J., 2002. Wanted: pig transplants that work. *Science* **295**, 1008.

Dagle, J.M. and Weeks, D.L., 2001. Oligonucleotide-based strategies to reduce gene expression. *Differentiation* **69**, 75–82.

Dai, Y., Vaught, T.D., Boone, J., Chen, S.H., Phelps, C.J., Ball, S., Monahan, J.A., Jobst, P.M., McCreath, K.J., Lamborn, A.E., Cowell-Lucero, J.L., Wells, K.D., Colman, A., Polejaeva, I.A. and Ayares, D.L., 2002. Targeted disruption of the alpha1,3–galactosyltransferase gene in cloned pigs. *Nat. Biotechnol*. **20**, 251–255.

Day, C.D., Lee, E., Kobayashi, J., Holappa, L.D., Albert, H. and Ow, D.W., 2000. Transgene integration into the same chromosome location can produce alleles that express at a predictable level, or alleles that are differentially silenced. *Genes Dev*. **14**, 2869–2880.

Dean, N.M., 2001. Functional genomics and target validation approaches using antisense oligonucleotide technology. *Curr. Opin. Biotechnol*. **12**, 622–625.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E. and Reik, W., 2001. Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. *Proc. Natl Acad. Sci*. USA **98**, 13 734–13 738.

Delvin, R.H., 1997. Transgenic salmonids. In: Houdebine L.M. (Ed.), *Transgenic Animal. Generation and Use*. Amsterdam: Harwood, pp. 105–118.

Denning, C., Burl, S., Ainslie, A., Bracken, J., Dinnyes, A., Fletcher, J., King, T., Ritchie, M., Ritchie, W.A., Rollo, M., de Sousa, P., Travers, A., Wilmut, I. and Clark, A.J., 2001. Deletion of the alpha(1,3)galactosyl transferase (GGTA1) gene and the prion protein (PrP) gene in sheep. *Nat. Biotechnol.* **19**, 559–562.

Di Bernardino, M.A., 2001. Animal cloning – the route to new genomics in agriculture and medicine. *Diffferentiation* **68**, 67–83.

Dunn, C.S., Mehtali, M., Houdebine, L.M., Gut, J.P., Kirn, A. and Aubertin, A.M., 1995. Human immunodeficiency virus type 1 infection of human CD4-transgenic rabbits. *J. Gen. Virol.* **76**, 1327–1336.

Dupuy, A.J., Clark, K., Carlson, C.M., Fritz, S., Davidson, A.E., Markley, K.M., Finley, K., Fletcher, C.F., Ekker, S.C., Hackett, P.B., Horn, S. and Largaespada, D.A., 2002. Mammalian germ-line transgenesis by transposition. *Proc. Natl Acad. Sci.* USA **99**, 4495–4499.

Durbach, A., Angevin, E., Poncet, P., Rouleau, M., Chavanel, G., Chapel, A., Thierry, D., Gorter, A., Hirsch, R., Charpentier, B., Senik, A. and Hirsch, F., 1999. Antibody-mediated endocytosis of G250 tumor-associated antigen allows targeted gene transfer to human renal cell carcinoma *in vitro*. *Cancer Gene Ther.* **6**, 564–571.

Edwards, R.G., 2001. IVF and the history of stem cells. *Nature* **413**, 349–351.

Eggan, K., Rode, A., Jentsch, I., Samuel, C., Hennek, T., Tintrup, H., Zevnik, B., Erwin, J., Loring, J., Jackson-Grusby, L., Speicher, M.R., Kuehn, R. and Jaenisch, R., 2002. Male and female mice derived from the same embryonic stem cell clone by tetraploid embryo complementation. *Nat. Biotechnol.* **20**, 455–459.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T., 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498.

Elbashir, S.M., Lendeckel, W. and Tuschl, T., 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200.

Encell, L.P., Landis, D.M. and Loeb, L.A., 1999. Improving enzymes for cancer gene therapy. *Nat. Biotechnol.* **17**, 143–147.

Enright, B.P., Taneja, M., Schreiber, D., Riesen, J., Tian, X.C., Fortune, J.E., Yang, X., 2002. Reproductive characteristics of cloned heifers derived from adult somatic cells. *Biol. Reprod.* **66**, 291–296.

Eszterhas, S.K., Bouhassira, E.E., Martin, D.I. and Fiering, S., 2002. Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol. Cell. Biol.* **22**, 469–479.

Evans, M.J., Gurer, C., Loike, J.D., Wilmut, I., Schnieke, A.E. and Schon, E.A., 1999. Mitochondrial DNA genotypes in nuclear transfer-derived cloned sheep. *Nat. Genet.* **23**, 90–103.

Fan, H., Lin, Q., Morrissey, G.R. and Khavari, P.A., 1999a. Immunization via hair follicles by topical application of naked DNA to normal skin. *Nat. Biotechnol.* **17**, 870–872.

Fan, J., Challah, M. and Watanabe, T., 1999b. Transgenic rabbit models for biomedical research: current status, basic methods and future perspectives. *Pathol. Int*. **49**, 583–594.

Fausto, N., 2001. A mouse model for hepatitis C virus infection? *Nat. Med*. **7**, 890–891.

Feng, Y.Q., Seibler, J., Alami, R., Eisen, A., Westerman, K.A., Leboulch, P., Fiering, S. and Bouhassira, E.E., 1999. Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. *J. Mol. Biol*. **292**, 779–785.

Ferber, D. 2001. Gene therapy. Safer and virus-free? *Science* **294**, 1638–1642.

Ferrari, S., Moro, E., Pettenazzo, A., Behr, J.P., Zacchello, F. and Scarpa, M., 1997. ExGen 500 is an efficient vector for gene delivery to lung epithelial cells *in vitro* and *in vivo*. *Gene Ther*. **4**, 1100–1106.

Fields, S., 2001. The interplay of biology and technology. *Proc. Natl Acad. Sci*. USA **98**, 10 051–10 054.

Finkel, E., 1999. Does cancer therapy trigger cell suicide? *Science* **286**, 2256–2258.

Fischer, A., 2000. Cautious advance. Gene therapy is more complex than anticipated. *EMBO Rep*. **1**, 294–296.

Fong, Y.W. and Zhou, Q., 2001. Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929–933.

Forster, K., Helbl, V., Lederer, T., Urlinger, S., Wittenburg, N. and Hillen, W., 1999. Tetracycline-inducible expression systems with reduced basal activity in mammalian cells. *Nucleic Acids Res*. **27**, 708–710.

Freed, C.R., 2002. Will embryonic stem cells be a useful source of dopamine neurons for transplant into patients with Parkinson's disease? *Proc. Natl Acad. Sci*. USA **99**, 1755–1757.

Fussenegger, M., Morris, R.P., Fux, C., Rimann, M., von Stockar, B., Thompson, C.J. and Bailey, J.E., 2000. Streptogramin-based gene regulation systems for mammalian cells. *Nat. Biotechnol*. **18**, 1203–1208.

Gavin, W.G., 2001. The future transgenics. *Regul. Affairs Focus*, May, pp.13–18.

Giovannangeli, C. and Hélène, C., 2000. Triplex technology takes off. *Nat. Biotechnol*. **18**, 1245–1246.

Giraldo, P. and Montoliu, L., 2001. Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Res*. **10**, 83–103.

Golovan, S.P., Hayes, M.A., Phillips, J.P. and Forsberg, C.W., 2001a. Transgenic mice expressing bacterial phytase as a model for phosphorus pollution control. *Nat. Biotechnol*. **19**, 429–433.

Golovan, S. P., Meidinger, R.G., Ajakaiye, M.C., Wiederkehr, M.Z., Barney, D.J., Plante, C., Pollard, J.W., Fan, M.Z., Hayes, M.A., Laursen, J., Hjorth, J.P., Hacker, R.R., Philipps, J.P. and Forsberg, C.W., 2001b. Pigs expressing salivary phytase produce low-phosphorus manure. *Nat. Biotechnol*. **19**, 741–745.

Goodwin, N.C., Ishida, Y., Hartford, S., Wnek, C., Bergstrom, R.A., Leder, P. and Schimenti, J.C., 2001. DelBank: a mouse ES-cell resource for generating deletions. *Nat. Genet*. **28**, 310–311.

Gordon, K., Lee, E., Vitale, J.A., Smith, A.E., Westphal, H. and Henninghausen, L., 1987. Production of human tissue plasminogen activator in transgenic mouse milk. *Bio/Technol*. **5**, 1183–1187.

Graham, I.R., Manzano, A., Tagalakis, A.D., Mohri, Z., Sperber, G., Hill, V., Beattie, S., Schepelmann, S., Dickson, G. and Owen, J.S., 2001. Gene repair validation. *Nat. Biotechnol*. **19**, 507–508.

Griffith, L.G. and Naughton, G., 2002. Tissue engineering – current challenges and expanding opportunities. *Science* **295**, 1009–1014.

Hackett, P.B., 2002. Genetic engineering: what are we fearing? *Transgenic Res*. **11**, 97–99.

Hackett, P.B., Cui, Z., Geurts, A., Clark, K.J., Yang, G.L., Dupuy, A., Largaespada, D.A. and Ekker, S.C., 2001. 'Structural functional studies with the sleeping beauty transposon system in vertebrates'. *Transgenic Animal Research Conference*. Tahoe City, CA, p. 12.

Hadjantonakis, A.K., and Papaioannou V.E., 2001. The stem cells of early embryos. *Differentiation* **68**, 159–166.

Hakelien, A.M., Landsverk, H.B., Robl, J.M., Skalhegg, B.S. and Collas, P., 2002. Reprogramming fibroblasts to express T-cell functions using cell extracts. *Nat. Biotechnol*. **20**, 460–466.

Hamilton, B.A. and Frankel, W.N., 2001. Of mice and genome sequence. *Cell* **107**, 13–16.

Hammer, R.E., Pursel, V.G., Rexroad, C.E., Jr., Wall, R.J., Bolt, D.J., Ebert, K.M., Palmiter, R.D. and Brinster, R.L., 1985. Production of transgenic rabbits, sheep and pigs by microinjection. *Nature* **315**, 680–683.

Henry, I., Forlani, S., Vaillant, S., Muschler, J., Choulika, A. and Nicolas, J.F., 1999. [LagoZ and LagZ, 2 genes depleted of CpG dinucleotides, derived from the LacZ gene for the study of epigenetic control]. *C. R. Acad. Sci*. III **322**, 1061–1070.

Heppner, F.L., Musahl, C., Arrighi, I., Klein, M.A., Rulicke, T., Oesch, B., Zinkernagel, R.M., Kalinke, U. and Aguzzi, A., 2001. Prevention of scrapie pathogenesis by transgenic expression of anti-prion protein antibodies. *Science* **294**, 178–182.

Herault, Y., Rassoulzadegan, M., Cuzin, F. and Duboule, D., 1998. Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat. Genet*. **20**, 381–384.

Hochedlinger, K. and Jaenisch, R., 2002. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* **415**, 1035–1038.

Hoffmann-Riem, H. and Wynne, B., 2002. In risk assessment, one has to admit ignorance. *Nature* **416**, 123.

Houdebine, L.M., 1998. 'The impact of genetic engineering on milk production'. In: Rasmussen, S. (Ed.). *25th International Dairy Congress* in Aarhus. pp. 127–134.

Houdebine, L.M., 2002. 'Transgenesis to improve animal production'. *Livestock Production Sci.* **74**, 255–268.

Houdebine, L.M., 2000. Transgenic animal bioreactors. *Transgenic Res.* **9**, 305–320.

Houdebine, L.M. and Attal, J., 1999. Internal ribosome entry sites (IRESs): reality and use. *Transgenic Res.* **8**, 157–177.

Houdebine L.M., Attal J. and Vilotte J.L., 2002. Vector design for transgene expression. In: Pinkert, C.A. (Ed.), *Transgenic Animal Technology*, Second Edition. San Diego: Academic, pp. 419–458.

Houdebine, L.M. and Weill, B., 1999. 'The impact of transgenesis and cloning on cell and organ xenotransplantation to humans'. In: Van Broekhoven, A., Shapiro, F. and Anné, J. (Eds), *Novel Frontiers in the Production of Compounds for Biomedical Use*. Dodwecht: Kluwer, pp. 351–362.

Ikawa, M., Tergaonkar, V., Ogura, A., Ogonuki, N., Inoue, K. and Verma, I.M., 2002. Restoration of spermatogenesis by lentiviral gene transfer: offspring from infertile mice. *Proc. Natl Acad. Sci.* USA **99**, 7524–7529.

Ishigaki, Y., Li, X., Serin, G. and Maquat, L.E., 2001. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* **106**, 607–617.

Jackson, I.J., 2001. Mouse mutagenesis on target. *Nat. Genet.* **28**, 198–200.

Jacobs, G.H., Rackham, O., Stockwell, P.A., Tate, W. and Brown, C.M., 2002. Transterm: a database of mRNAs and translational control elements. *Nucleic Acids Res.* **30**, 310–311.

Jaenisch, R. and Wilmut, I., 2001. Developmental biology. Don't clone humans! *Science* **291**, 2552.

Janson, C.G. and During, M.J., 2001. Viral vectors as part of an integrated functional genomics program. *Genomics* **78**, 3–6.

Jenke, B.H., Fetzer, C.P., Stehle, I.M., Jonsson, F., Fackelmayer, F.O., Conradt, H., Bode, J. and Lipps, H.J., 2002. An episomally replicating vector binds to the nuclear matrix protein SAF-A *in vivo*. *EMBO Rep.* **3**, 349–354.

Jiang, Y., Jahagirdar, B.N., Reinhardt, R.L., Schwartz, R.E., Keene, C.D., Ortiz-Gonzalez, X.R., Reyes, M., Lenvik, T., Lund, T., Blackstad, M., Du, J., Aldrich, S., Lisberg, A., Low, W.C., Largaespada, D.A. and Verfaillie, C.M., 2002. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* **418**, 41–49.

Johnson, K.D., Christensen, H.M., Zhao, B. and Bresnick, E.H., 2001a. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol. Cell* **8**, 465–471.

Johnson, L., Mercer, K., Greenbaum, D., Bronson, R.T., Crowley, D., Tuveson, D.A. and Jacks, T. 2001b. Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature* **410**, 1111–1116.

Jost, B., Vilotte, J.L., Duluc, I., Rodeau, J.L. and Freund, J.N., 1999. Production of low-lactose milk by ectopic expression of intestinal lactase in the mouse mammary gland. *Nat. Biotechnol.* **17**, 160–164.

Kahn, A., 2000. Converting hepatocytes to β-cells – a new approach for diabetes? *Nat. Med.* **6**, 505–506.

Kapoor, P., Shire, K. and Frappier, L., 2001. Reconstitution of Epstein-Barr virus-based plasmid partitioning in budding yeast. *EMBO J.* **20**, 222–230.

Kasinathan, P., Knott, J.G., Moreira, P.N., Burnside, A.S., Jerry, D.J. and Robl, J.M., 2001a. Effect of fibroblast donor cell age and cell cycle on development of bovine nuclear transfer embryos *in vitro*. *Biol. Reprod.* **64**, 1487–1493.

Kasinathan, P., Knott, J.G., Wang, Z., Jerry, D.J. and Robl, J.M., 2001b. Production of calves from G1 fibroblasts. *Nat. Biotechnol.* **19**, 1176–1178.

Kawasaki, H., Onuki, R., Suyama, E. and Taira, K., 2002. Identification of genes that function in the TNF-alpha-mediated apoptotic pathway using randomized hybrid ribozyme libraries. *Nat. Biotechnol.* **20**, 376–380.

Kawasaki, H. and Taira, K., 2002. Identification of genes by hybrid ribozymes that couple cleavage activity with the unwinding activity of an endogenous RNA helicase. *EMBO Rep.* **3**, 443–450.

Kayser, K., 1997. Gene transfer in *Drosphila melanogaster*. In: Houdebine, L.M. (Ed.), *Transgenic Animal Generation and Use*. Amsterdam: Harwood, pp. 133–137.

Kelleher, Z.T., Fu, H., Livanos, E., Wendelburg, B., Gulino, S. and Vos, J.M., 1998. Epstein-Barr-based episomal chromosomes shuttle 100 kb of self-replicating circular human DNA in mouse cells. *Nat. Biotechnol.* **16**, 762–768.

Keys, R.A. and Green, M.R., 2001. Gene expression. The odd coupling. *Nature* **413**, 583– 585.

Kolb, A.F., Ansell, R., McWhir, J. and Siddell, S.G., 1999. Insertion of a foreign gene into the beta-casein locus by Cre-mediated site-specific recombination. *Gene* **227**, 21–31.

Kolb, A.F., Pewe, L., Webster, J., Perlman, S., Whitelaw, C.B. and Siddell, S.G., 2001. Virus-neutralizing monoclonal antibody expressed in milk of transgenic mice provides full protection against virus-induced encephalitis. *J. Virol.* **75**, 2803–2809.

Kong, J. and Xu, Z., 2000. Overexpression of neurofilament subunit NF-L and NF-H extends survival of a mouse model for amyotrophic lateral sclerosis. *Neurosci. Lett.* **281**, 72–74.

Kozak, M., 1999. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell. Biol.* **9**, 5134–5142.

Kozak, M., 2001a. Extensively overlapping reading frames in a second mammalian gene. *EMBO Rep.* **2**, 768–769.

Kozak, M., 2001b. New ways of initiating translation in eukaryotes? *Mol. Cell. Biol.* **21**, 1899–907.

Krimpenfort, P., Rademakers, A., Eyestone, W., van der Schans, A., van den Broek, S., Kooiman, P., Kootwijk, E., Platenburg, G., Piper, F. and Strijker, R., 1991. Generation of transgenic dairy cattle using *in vitro* embryo production. *Bio/Technology* **9**, 844–847.

Lai, L., Kolber-Simonds, D., Park, K.W., Cheong, H.T., Greenstein, J.L., Im, G.S., Samuel, M., Bonk, A., Rieke, A., Day, B.N., Murphy, C.N., Carter, D.B., Hawley, R.J. and Prather, R.S., 2002. Production of alpha-1,3-galactosyltransferase knockout pigs by nuclear transfer cloning. *Science* **295**, 1089–1092.

Landsverk, H.B., Hakelien, A.M., Kuntziger, T., Robl, J.M., Skalhegg, B.S. and Collas, P., 2002. Reprogrammed gene expression in a somatic cell-free extract. *EMBO Rep*. **3**, 384–389.

Lanza, R.P., Chung, H.Y., Yoo, J.J., Wettstein, P.J., Blackwell, C., Borson, N., Hofmeister, E., Schuch, G., Soker, S., Moraes, C.T., West, M.D. and Atala, A., 2002a. Generation of histocompatible tissues using nuclear transplantation. *Nat. Biotechnol*. **20**, 689–696.

Lanza, R.P., Cibelli, J.B. and West, M.D., 2002b. Transparency in public relations. *Nat. Biotechnol*. **20**, 117–118.

Lecuit, M., Vandormael-Pournin, S., Lefort, J., Huerre, M., Gounon, P., Dupuy, C., Babinet, C. and Cossart, P., 2001. A transgenic model for listeriosis: role of internalin in crossing the intestinal barrier. *Science* **292**, 1722–1725.

Lee, S.H., Lumelsky, N., Studer, L., Auerbach, J.M. and McKay, R., 2000. Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nat. Biotechnol*. **18**, 675–679.

Lemischka, I., 2002. Rethinking somatic stem cell plasticity. *Nat. Biotechnol*. **20**, 425.

Li, Q., Harju, S. and Peterson, K.R., 1999. Locus control regions: coming of age at a decade plus. *Trends Genet*. **15**, 403–408.

Liu, P., Jenkins, N.A. and Copeland, N.G., 2002. Efficient Cre–LoxP-induced mitotic recombination in mouse embryonic stem cells. *Nat. Genet*. **30**, 66–72.

Liu, Q., Li, M.Z., Leibham, D., Cortez, D. and Elledge, S.J., 1998. The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes. *Curr. Biol*. **8**, 1300–1309.

Loi, P., Ptak, G., Barboni, B., Fulka, J., Jr., Cappai, P. and Clinton, M. 2001. Genetic rescue of an endangered mammal by cross-species nuclear transfer using post-mortem somatic cells. *Nat. Biotechnol*. **19**, 962–964.

Lois, C., Hong, E.J., Pease, S., Brown, E.J. and Baltimore, D., 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295**, 868–872.

Lorens, J.B., Sousa, C., Bennett, M.K., Molineaux, S.M. and Payan, D.G., 2001. The use of retroviruses as pharmaceutical tools for target discovery and validation in the field of functional genomics. *Curr. Opin. Biotechnol*. **12**, 613–621.

Lubon, H., 1998. Transgenic animal bioreactors in biotechnology and production of blood proteins. Biotechnol. *Annu. Rev*. **4**, 1–54.

Luo, M.J. and Reed, R., 1999. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Nat. Acad. Sci*. USA **96**, 14 937–14 942.

Manolakou, K., Beaton, J., McConnell, I., Farquar, C., Manson, J., Hastie, N.D., Bruce, M. and Jackson, I.J., 2001. Genetic and environmental factors

modify bovine spongiform encephalopathy incubation period in mice. *Proc. Natl Acad. Sci*. USA **98**, 7402–7407.

Marsh-Armstrong, N., Huang, H., Berry, D.L. and Brown, D.D., 1999. Germline transmission of transgenes in *Xenopus laevis*. *Proc. Natl Acad. Sci.* USA **96**, 14 389–14 393.

Martin, D.I., 2001. Transcriptional enhancers – on/off gene regulation as an adaptation to silencing in higher eukaryotic nuclei. *Trends Genet*. **17**, 444–448.

Mattick, J.S., 2001. Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep*. **2**, 986–991.

McCreath, K.J., Howcroft, J., Campbell, K.H., Colman, A., Schnieke, A.E. and Kind, A.J., 2000. Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature* **405**, 1066–1069.

McDowell, N., 2002. Stream of escaped farm fish raises for wild salmon. *Nature*, **416**, 571.

Medico, E., Gambarotta, G., Gentile, A., Comoglio, P.M. and Soriano, P., 2001. A gene trap vector system for identifying transcriptionally responsive genes. *Nat. Biotechnol*. **19**, 579–582.

Mendell, J.T. and Dietz, H.C., 2001. When the message goes awry: disease-producing mutations that influence mRNA content and performance. *Cell* **107**, 411–414.

Miller, M.W. and Rubin, E.M., 1997. *Transgenic Animals in Atherosclerosis Research*. Amsterdam: Harwood, pp. 445–448.

Mills, A.A., 2001. Changing colors in mice: an inducible system that delivers. *Genes Dev*. **15**, 1461–1467.

Mitalipov, S.M., Nusser, K.D. and Wolf, D.P., 2001. Parthenogenetic activation of rhesus monkey oocytes and reconstructed embryos. *Biol. Reprod*. **65**, 253–259.

Moll, J., Barzaghi, P., Lin, S., Bezakova, G., Lochmuller, H., Engvall, E., Muller, U. and Ruegg, M.A., 2001. An agrin minigene rescues dystrophic symptoms in a mouse model for congenital muscular dystrophy. *Nature* **413**, 302–307.

Mollard, R., Denham, M. and Trounson, A., 2002. Technical advances and pitfalls on the way to human cloning. *Differentiation* **70**, 1–9.

Moore, A., 2001. Of mice and Mendel. The predicted rise in the use of knock-out and transgenic mice should cause us to reflect on our justification for the use of animals in research. *EMBO Rep*. **2**, 554–558.

Moore, R.C. and Melton, D.W., 1997. Transgenic analysis of prion diseases. *Mol. Hum. Reprod*. **3**, 529–544.

Muir, W.M. and Howard, R.D., 2002. Assessment of possible ecological risks and hazards of transgenic fish with implications for other sexually reproducing organisms. *Transgenic Res*. **11**, 101–114.

Mukherjee, D., Gao, M., O'Connor, J.P., Raijmakers, R., Pruijn, G., Lutz, C.S. and Wilusz, J., 2002. The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *EMBO J*. **21**, 165–174.

Müller, M., 2000. Increasing disease resistance in transgenic domestic animals. In: Toutant, J.P. and Balazs, E. (Eds), *Molecular Farming*. Paris: INRA, pp. 87–98.

Mulsant, P., Lecerf, F., Fabre, S., Schibler, L., Monget, P., Lanneluc, I., Pisselet, C., Riquet, J., Monniaux, D., Callebaut, I., Cribiu, E., Thimonier, J., Teyssier, J., Bodin, L., Cognie, Y., Chitour, N. and Elsen, J.M., 2001. Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Merino ewes. *Proc. Natl Acad. Sci.* USA **98**, 5104–5109.

Nagano, M., Brinster, C.J., Orwig, K.E., Ryu, B.Y., Avarbock, M.R. and Brinster, R.L., 2001. Transgenic mice produced by retroviral transduction of male germ-line stem cells. *Proc. Natl Acad. Sci.* USA **98**, 13090–13095.

Nagy, A., 2000. Cre recombinase: the universal reagent for genome tailoring. *Genesis* **26**, 99–109.

No, D., Yao, T.P. and Evans, R.M., 1996. Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl Acad. Sci.* USA **93**, 3346–3351.

Nottle, M.B., Nagashima, H., Verma, P.J., Du, Z.T., Ashman, C.G., Grupen, C.G., Malcifatrick, S.M., Harding, M.P., Wilgley, P.L., Lyons, I.G., Crawford, R.J., Harrison, D.T., Luxford, B.G., Campwell, R.G., Robins, A.J., 1997. Production and evaluation of growth hormone transgenic pigs. *Transgenic Animals in Agriculture Conference*, Tahoe City, CA, p. 10.

Novacek, M.J., 1996. Taxonomy. Where do rabbits and kin fit in? *Nature* **379**, 299–300.

Ogonuki, N., Inoue, K., Yamamoto, Y., Noguchi, Y., Tanemura, K., Suzuki, O., Nakayama, H., Doi, K., Ohtomo, Y., Satoh, M., Nishida, A. and Ogura, A., 2002. Early death of mice cloned from somatic cells. *Nat. Genet*. **30**, 253–254.

Ohkawa, J. and Taira, K., 2000. Control of the functional activity of an antisense RNA by a tetracycline-responsive derivative of the human U6 snRNA promoter. *Hum. Gene Ther*. **11**, 577–585.

Oldstone, M.B., Lewicki, H., Thomas, D., Tishon, A., Dales, S., Patterson, J., Manchester, M., Homann, D., Naniche, D. and Holz, A., 1999. Measles virus infection in a transgenic model: virus-induced immunosuppression and central nervous system disease. *Cell* **98**, 629–640.

Paddison, P.J., Caudy, A.A. and Hannon, G.J., 2002. Stable suppression of gene expression by RNAi in mammalian cells. Proc. *Natl Acad. Sci.* USA **99**, 1443–1448.

Palmiter, R.D., Brinster, R.L., Hammer, R.E., Trumbauer, M.E., Rosenfeld, M.G., Birnberg, N.C. and Evans, R.M., 1991. Dramatic growth of mice that develop from eggs microinjected with metallothionein–growth hormone fusion genes. *Nature* **300**, 611–615.

Park, K.W., Lai, L., Cheong, H.T., Cabot, R., Sun, Q.Y., Wu, G., Rucker, E.B., Durtschi, D., Bonk, A., Samuel, M., Rieke, A., Day, B.N., Murphy, C.N.,

Carter, D.B. and Prather, R.S., 2002. Mosaic gene expression in nuclear transfer-derived embryos and the production of cloned transgenic pigs from ear-derived fibroblasts. *Biol. Reprod.* **66**, 1001–1005.

Pati, S, 1998. *Genetically Engineering and Cloning Animals*. Park City, Deer Valley, UT.

Pearson, H., 2002. Dual identities. *Nature* **417**, 10–11.

Perry, A.C., Rothman, A., de las Heras, J.I., Feinstein, P., Mombaerts, P., Cooke, H.J. and Wakayama, T., 2001. Efficient metaphase II transgenesis with different transgene archetypes. *Nat. Biotechnol.* **19**, 1071–1073.

Perry, A.C. and Wakayama, T., 2002. Untimely ends and new beginnings in mouse cloning. *Nat. Genet.* **30**, 243–244.

Perry, A.C., Wakayama, T., Kishikawa, H., Kasai, T., Okabe, M., Toyoda, Y. and Yanagimachi, R., 1999. Mammalian transgenesis by intracytoplasmic sperm injection. *Science* **284**, 1180–1183.

Pesole, G., Liuni, S., Grillo, G., Licciulli, F., Mignone, F., Gissi, C. and Saccone, C., 2002. UTRdb and UTRsite: specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs. Update 2002. *Nucleic Acids Res.* **30**, 335–340.

Petitclerc, D., Attal, J., Theron, M.C., Bearzotti, M., Bolifraud, P., Kann, G., Stinnakre, M.G., Pointu, H., Puissant, C. and Houdebine, L.M., 1995. The effect of various introns and transcription terminators on the efficiency of expression vectors in various cultured cell lines and in the mammary gland of transgenic mice. *J. Biotechnol.* **40**, 169–178.

Petters, R.M. and Sommer, J.R., 2000. Transgenic animals as models for human disease. *Transgenic Res.* **9**, 347–351.

Pinkert, C.A., 2002. *Transgenic Animal Technology*. Second Edition. San Diego: Academic.

Pintado, B. and Gutierrez-Adan, A., 1999. Transgenesis in large domestic species: future development for milk modification. *Reprod. Nutr. Dev.* **39**, 535–544.

Platt, J.L., 2002. Knocking out xenograft rejection. *Nat. Biotechnol.* **20**, 231–232.

Powledge, T.M., 2002. Will they throw the bath water out with the baby?: the US Congress is still debating whether to outlaw cloning humans. *EMBO Rep.* **3**, 209–211.

Pursell, V.G., Mitchell, A.D., Wall, M.B., Solomon, M.B., Coleman, M.E. and Schwartz, R.J., 2001. Transgenic research in enhance growth and lean carcass composition in swine. In: Toutant, J.P. and Balazs, E. (Eds), *Molecular Farming*. Paris: INRA, pp. 77–86.

Qian, 2001a. *Transgenic Animal Research Conference* III, Tahoe City, CA. Abstract DD.

Qian, 2001b. *Transgenic Animal Research Conference* III, Tahoe City, CA. Abstract CC.

Ranger, A.M., Malynn, B.A. and Korsmeyer, S.J., 2001. Mouse models of cell death. *Nat. Genet.* **28**, 113–118.

Rideout, W.M. III, Wakayama, T., Wutz, A., Eggan, K., Jackson-Grusby, L., Dausman, J., Yanagimachi, R. and Jaenisch, R., 2000. Generation of mice from wild-type and targeted ES cells by nuclear cloning. *Nat. Genet*. **24**, 109–110.

Reggio, B.C., James, A.N., Green, H.L., Gavin, W.G., Behboodi, E., Echelard, Y. and Godke, R.A., 2001. Cloned transgenic offspring resulting from somatic cell nuclear transfer in the goat: oocytes derived from both follicle-stimulating hormone-stimulated and nonstimulated abattoir-derived ovaries. *Biol. Reprod*. **65**, 1528–1533.

Reid, W., Sadowska, M., Denaro, F., Rao, S., Foulke, J., Jr., Hayes, N., Jones, O., Doodnauth, D., Davis, H., Sill, A., O'Driscoll, P., Huso, D., Fouts, T., Lewis, G., Hill, M., Kamin-Lewis, R., Wei, C., Ray, P., Gallo, R.C., Reitz, M. and Bryant, J., 2001. An HIV-1 transgenic rat that develops HIV-related pathology and immunologic dysfunction. *Proc. Natl Acad. Sci*. USA **98**, 9271–9276.

Reik, W., Dean, W. and Walter, J., 2001. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089–1093.

Ren, R.B., Costantini, F., Gorgacz, E.J., Lee, J.J. and Racaniello, V.R., 1990. Transgenic mice expressing a human poliovirus receptor: a new model for poliomyelitis. *Cell* **63**, 353–362.

Renard, J.P., Zhou, Q., LeBourhis, D., Chavatte-Palmer, P., Hue, I., Heyman, Y. and Vignon, X., 2002. Nuclear transfer technologies: between successes and doubts. *Theriogenology* **57**, 203–222.

Rideout, W.M. III, Eggan, K. and Jaenisch, R., 2001. Nuclear cloning and epigenetic reprogramming of the genome. *Science* **293**, 1093–1098.

Rideout, W.M. III, Hochedlinger, K., Kyba, M., Daley, G.Q. and Jaenisch, R., 2002. Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. *Cell* **109**, 17–27.

Rivera, V.M., Clackson, T., Natesan, S., Pollock, R., Amara, J.F., Keenan, T., Magari, S.R., Phillips, T., Courage, N.L., Cerasoli, F., Jr., Holt, D.A. and Gilman, M., 1996. A humanized system for pharmacologic control of gene expression. *Nat. Med*. **2**, 1028–1032.

Ronfort, C.M., Legras C. and Verdier G., 1997. The use of retroviral vectors for gene transfer into bird embryo. In: Houdebine, L.M. (Ed.), *Transgenic Animal Generation and Use*. Amsterdam: Harwood, pp. 83–95.

Rossant, J., 2002. A monoclonal mouse? *Nature* **415**, 967–969.

Rossi, F.M. and Blau, H.M., 1998. Recent advances in inducible gene expression systems. *Curr. Opin. Biotechnol*. **9**, 451–456.

Saif, L.J. and Wheeler, M.B., 1998. WAPping gastroenteritis with transgenic antibodies. *Nat. Biotechnol*. **16**, 334–335.

Saito, M., Iwawaki, T., Taya, C., Yonekawa, H., Noda, M., Inui, Y., Mekada, E., Kimata, Y., Tsuru, A. and Kohno, K., 2001. Diphtheria toxin receptor-mediated conditional and targeted cell ablation in transgenic mice. *Nat. Biotechnol*. **19**, 746–750.

Sato, M., Ishikawa, A. and Kimura, M., 2002a. Direct injection of foreign DNA into mouse testis as a possible *in vivo* gene transfer system via epididymal spermatozoa. *Mol. Reprod. Dev*. **61**, 49–56.

Sato, M., Ozawa, T., Inukai, K., Asano, T. and Umezawa, Y., 2002b. Fluorescent indicators for imaging protein phosphorylation in single living cells. *Nat. Biotechnol*. **20**, 287–294.

Schneider, H. and Coutelle, C., 1999. *In utero* gene therapy: the case for. *Nat. Med*. **5**, 256–257.

Schnieke, A.E., Kind, A.J., Ritchie, W.A., Mycock, K., Scott, A.R., Ritchie, M., Wilmut, I., Colman, A. and Campbell, K.H., 1997. Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. *Science* **278**, 2130–2133.

Schwertfeger, K.L., Richert, M.M. and Anderson, S.M., 2001. Mammary gland involution is delayed by activated Akt in transgenic mice. *Mol. Endocrinol*. **15**, 867–881.

Shen, C.N., Slack, J.M. and Tosh, D., 2000. Molecular basis of transdifferentiation of pancreas to liver. *Nat. Cell. Biol*. **2**, 879–887.

Shermann A., Dawson, A., Mather, C., Gilooley, H., Li, Y., Mitchell, R., Finnegan, D. and Sang, H., 1998. Transposition of the *Drosophila* element mariner into the chicken germ line. *Nat. Biotech*. **16**, 1050–1053.

Siegel, P.M., Hardy, W.R. and Muller, W.J., 2000. Mammary gland neoplasia: insights from transgenic mouse models. *Bioessays* **22**, 554–563.

Simmons, J.P., McClenaghan, M. and Clark, A.J., 1987. Alteration of the quality of milk by expression of sheep beta-lactoglobulin in transgenic mice. *Nature* **328**, 530–532.

Snyder, E.Y. and Vescovi, A.L., 2000. The possibilities/perplexities of stem cells. *Nature Biotech*. **18**, 827–828.

Solter, D. and Gearhart, J., 1999. Putting stem cells to work. *Science* **283**, 1468–1470.

Somia, N.V., Kafri, T. and Verma, I.M., 1999. Piecing together more efficient gene expression. *Nat. Biotechnol*. **17**, 224–225.

Sonoda, E., Takata, M., Yamashita, Y.M., Morrison, C. and Takeda, S., 2001. Homologous DNA recombination in vertebrate cells. *Proc. Natl Acad. Sci*. USA **98**, 8388–8394.

Studer, L., 2001. Stem cells with brainpower. *Nat. Biotechnol*. **19**, 1117–1118.

Sui, G., Soohoo, C., el Affar, B., Gay, F., Shi, Y. and Forrester, W.C., 2002. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl Acad. Sci*. USA **99**, 5515–5520.

Surani, M.A., 2002. Genetics: immaculate misconception. Nature **416**, 491–493.

Szathmary, E., Jordan, F. and Pal, C., 2001. Molecular biology and evolution. Can genes explain biological complexity? *Science* **292**, 1315–1316.

Taboit-Dameron, F., 1999. 'Obtention de modèles animaux transgéniques pour l'étude des effets anticompléments des protéines humaines DAF et CD59 et de

l'hypodermine A dans la prévention du rejet hyperaigu des xénogreffes discordantes'. Paris XI, Orsay.

Taboit-Dameron, F., Malassagne, B., Viglietta, C., Puissant, C., Leroux-Coyau, M., Chereau, C., Attal, J., Weill, B. and Houdebine, L.M., 1999. Association of the 5′HS4 sequence of the chicken beta-globin locus control region with human EF1 alpha gene promoter induces ubiquitous and high expression of human CD55 and CD59 cDNAs in transgenic rabbits. *Transgenic Res.* **8**, 223–235.

Tamashiro, K.L., Wakayama, T., Akutsu, H., Yamazaki, Y., Lachey, J.L., Wortman, M.D., Seeley, R.J., D'Alessio, D.A., Woods, S.C., Yanagimachi, R. and Sakai, R.R., 2002. Cloned mice have an obese phenotype not transmitted to their offspring. *Nat. Med.* **8**, 262–267.

Tamura, T., Thibert, C., Royer, C., Kanda, T., Abraham, E., Kamba, M., Komoto, N., Thomas J.L., Mauchamp B., Chavancy G., Shirp P., Fraser M., Prudhomme J.C. and Couble P., 1999. Germline transformation of the silkworm *Bombyx mori L.* using a piggyBac transposon-derived vector. *Nat. Biotechnol.* **18**, 81–84.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S. and Jones, J.M., 1998. Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147.

Thomson, J.A. and Odorico, J.S., 2000. Human embryonic stem cell and embryonic germ cell lines. *Trends Biotechnol.* **18**, 53–57.

Timmons, L. and Fire, A., 1998. Specific interference by ingested dsRNA. *Nature* **395**, 854.

Tomizuka, K., Yoshida, H., Uejima, H., Kugoh, H., Sato, K., Ohguma, A., Hayasaka, M., Hanaoka, K., Oshimura, M. and Ishida, I., 1997. Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. *Nat. Genet.* **16**, 133–143.

Trounson, A., 2002. The genesis of embryonic stem cells. *Nat. Biotechnol.* **20**, 237–238.

Tuschl, T., 2002. Expanding small RNA interference. *Nat. Biotechnol.* **20**, 446–448.

Ui-Tei, K., Zenno, S., Miyata, Y. and Saigo, K., 2000. Sensitive assay of RNA interference in *Drosophila* and Chinese hamster cultured cells using firefly luciferase gene as target. *FEBS Lett.* **479**, 79–82.

Upegui-Gonzalez, L.C., Francois, J.C., Ly, A. and Trojan, J., 2000. The approach of triple helix formation in control of gene expression and the treatment of tumors expressing IGF-I. *Adv. Exp. Med. Biol.* **465**, 319–332.

Van den Hout, J.M., Reuser, A.J.J., de Klerk, J.B.C., Arts, W.F., Smeitink, J.A.M. and Van der Ploeg, A.T., 2001. Enzyme therapy for Pompe disease with recombinant human α-glucosidase from rabbit milk. *J. Inherit. Metab. Dis.* **24**, 255–274.

Vasquez, K.M., Marburger, K., Intody, Z. and Wilson, J.H., 2001. Manipulating the mammalian genome by homologous recombination. *Proc. Natl Acad. Sci.* USA **98**, 8403–8410.

Vignon, X., Chesne, P., Le Bourhis, D., Flechon, J.E., Heyman, Y. and Renard, J.P., 1998. Developmental potential of bovine embryos reconstructed from enucleated matured oocytes fused with cultured somatic cells. *C. R. Acad. Sci*. III **321**, 735–745.

Voet, T., Vermeesch, J., Carens, A., Durr, J., Labaere, C., Duhamel, H., David, G. and Marynen, P., 2001. Efficient male and female germline transmission of a human chromosomal vector in mice. *Genome Res*. **11**, 124–136.

Vogelstein, B., Alberts, B. and Shine, K., 2002. Genetics. Please don't call it cloning! *Science* **295**, 1237.

Voinnet, O., 2001. RNA silencing as a plant immune system against viruses. *Trends Genet*. **17**, 449–459.

Vos, J.M., 1998. Mammalian artificial chromosomes as tools for gene therapy. *Curr. Opin. Genet. Dev*. **8**, 351–359.

Wade-Martins, R., Smith, E.R., Tyminski, E., Chiocca, E.A. and Saeki, Y., 2001. An infectious transfer and expression system for genomic DNA loci in human and mouse cells. *Nat. Biotechnol*. **19**, 1067–1070.

Wakayama, T., Perry, A.C., Zuccotti, M., Johnson, K.R. and Yanagimachi, R., 1998. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **394**, 369–374.

Wakayama, T., Shinkai, Y., Tamashiro, K.L., Niida, H., Blanchard, D.C., Blanchard, R.J., Ogura, A., Tanemura, K., Tachibana, M., Perry, A.C., Colgan, D.F., Mombaerts, P. and Yanagimachi, R., 2000. Cloning of mice to six generations. *Nature* **407**, 318–319.

Wang, Y., DeMayo, F.J., Tsai, S.Y. and O'Malley, B.W., 1997. Ligand-inducible and liver-specific target gene expression in transgenic mice. *Nat. Biotechnol*. **15**, 239–243.

Warashina, M., Kuwabara, T., Kato, Y., Sano, M. and Taira, K., 2001. RNA–protein hybrid ribozymes that efficiently cleave any mRNA independently of the structure of the target RNA. *Proc. Natl Acad. Sci*. USA **98**, 5572–5577.

Ward, K.A., 2000. Transgene mediated modifications to animal biochemistry. *TIB Tech*. **18**, 99–102.

Weninger, W.J. and Mohun, T., 2002. Phenotyping transgenic embryos: a rapid 3-D screening method based on episcopic fluorescence image capturing. *Nat. Genet*. **30**, 59–65.

West A.G., Gaszner, M. and Felsenfeld, G., 2002. Insulators: many functions, many mechanisms. *Genes Dev*. **16**, 271–288.

Western, P.S. and Surani, M.A., 2002. Nuclear reprogramming alchemy or analysis? *Nat. Biotechnol*. **20**, 445–446.

Westhusin, M.E., Long, C.R., Shin, T., Hill, J.R., Looney, C.R., Pryor, J.H. and Piedrahita, J.A., 2001. Cloning to reproduce desired genotypes. *Theriogenology* **55**, 35–49.

Wheeler, M.B. and Walters, E.M., 2001. Transgenic technology and applications in swine. *Theriogenology* **56**, 1345–1369.

Whitelaw, B., 1999. Toward designer milk. *Nat. Biotechnol*. **17**, 135–136.
Whitelaw, E. and Martin, D.I., 2001. Retrotransposons as epigenetic mediators of phenotypic variation in mammals. *Nat. Genet*. **27**, 361–365.
Wianny, F. and Zernicka-Goetz, M., 2000. Specific interference with gene function by double-stranded RNA in early mouse development. *Nat. Cell. Biol*. **2**, 70–75.
Wilkinson, M.F. and Shyu, A.B., 2001. Multifunctional regulatory proteins that control gene expression in both the nucleus and the cytoplasm. *Bioessays* **23** 775–787.
Wilmut, I., 2002. Are there any normal cloned mammals? *Nat. Med*. **8**, 215–216.
Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J. and Campbell, K.H., 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810–813.
Wilusz, C.J., Wang, W. and Peltz, S.W., 2001. Curbing the nonsense: the activation and regulation of mRNA surveillance. *Genes Dev*. **15**, 2781–2785.
Wurmser, A.E. and Gage, F.H., 2002. Stem cells: cell fusion causes confusion. *Nature* **416**, 485–487.
Xue, F., Tian, X.C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., Dai, Y., Levine, H., Pereira, L.V. and Yang, X., 2002. Aberrant patterns of X chromosome inactivation in bovine clones. *Nat. Genet*. **31**, 216–220.
Yang, J., Ratovitski, T., Brady, J.P., Solomon, M.B., Wells, K.D. and Wall, R.J., 2001. Expression of myostatin pro domain results in muscular transgenic mice. *Mol. Reprod. Dev*. **60**, 351–361.
Yu, Q., Geng, Y. and Sicinski, P., 2001. Specific protection against breast cancers by cyclin D1 ablation. *Nature* **411**, 1017–1021.
Zallen, T.D., 2000. US gene therapy in crisis. *Trends Genet*. 16, No. **6**, 272–275.
Zambrowicz, B.P. and Friedrich, G.A., 1998. Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse. *Int. J. Dev. Biol*. **42**, 1025–1036.

Animal Transgenesis and Cloning. Louis-Marie Houdebine

ISBNs: 0-470-84827-8 (HB); 0-470-84828-6 (PB)

Index