Transgenic systems in drug discovery: from target identification to humanized mice

Jan Törnell and Michael Snaith

Pharmaceutical companies are faced with the challenge that only ~10% of compounds tested in costly clinical trials eventually become a new drug. Investment in early discovery research can decrease this attrition in late-stage R&D and focus resources on the best targets. Transgenic technology influences decision-making in target identification, target validation, and can also provide better models for human diseases, as well as models designed to alert researchers early about potential issues with drug metabolism and toxicity. Here we review how transgenic technology can reduce the late-stage attrition by increasing the quality of both the target and the compound.

*Jan Törnell
Michael Snaith
AstraZeneca Transgenics and Comparative Genomics
AstraZeneca R&D
S-431 83 Mölndal
Sweden
tel: +46 31 776 10 00
fax: +46 31 776 38 20
*e-mail: jan.tornell@astrazeneca.com

▼ The first transgenic mouse was generated using the microinjection technology by Gordon and co-workers in the early 1980s [1,2]. To date, >21,000 papers have been published in Medline-indexed journals using transgenic and knockout technologies. This demonstrates the extent to which these technologies have become an essential part of basic biological research. Their impact is further supported by the fact that >1500 of these were published in high profile journals such as *Cell, Nature* and *Science*. The first gene addition experiments (using viral gene transfer) were described by Jaenish and co-workers in 1976 [3], but it was not until the spectacular publications from the laboratories of Palmiter and Brinster in 1982 [4], describing the giant mice overexpressing growth hormone, that the technology gained general awareness. Since then, transgenic technology has facilitated great advances in research fields such as immunology [5], oncology [6], developmental biology [7], neuroscience [8], cardiovascular research [9] and many others.

Transgenic technology does, however, have several limitations and mutation of genes does not always result in the 'anticipated' phenotype or disease model. This might result from compensatory mechanisms, poor correlation between mouse and human pathophysiology, not applying the proper challenges to the animals to discover the phenotype, not having the appropriate technology to make difficult phenotypic measurements or sometimes the working hypothesis can turn out to be invalid. In trying to understand human pathology it is important to know the limitations of the model. No model can be expected to give all the answers. It is, however, important to note the features that make the mouse a unique species for modelling and understanding human disease (i.e. mammalian, DNA sequence determined, collection of inbred strains with defined phenotypes available, gene targeting possible, sufficiently small for large-scale random mutagenesis studies). No other model organism has all these features.

Challenges for the pharmaceutical industry

In many disease areas there is great unmet medical need for new effective drugs and pharmaceutical companies are acutely aware of this, but they also recognize that any new drug treatments must be as safe as possible to be acceptable to society and regulatory authorities alike. At the same time, there is pressure on pharmaceutical companies from investors to decrease the time and cost needed to generate these new, safe medicines. In the future, one way drug companies might be able to satisfy some of these demands is to reduce the late-stage attrition rate that is an integral part of the current general approach to drug development.

1359-6446/02/$ – see front matter PII: S1359-6446(02)02234-1

As previously mentioned, only ~10% of compounds tested in clinical trials eventually make it to the market, and of these only a small proportion will generate a significant profit. The cost of identifying a new drug is immense – a rough estimate would be ~US$800 million, and 80% of this cost is spent on clinical trials and development. To improve this situation it will be necessary for the pharmaceutical industry to adopt new technologies and working strategies. In place of the 'high-quantity' model used widely today (e.g. expression profiling, HTS of large compound libraries; overcoming high attrition with high volume), a 'high-quality' model must be implemented. This could include some of the following: disease-driven target identification, thorough *in vivo* validation of selected targets, HTS of focussed chemical libraries, rational drug design, which should lead to a reduction in attrition by increasing the quality of both targets and compounds entering the clinical trials phase.

Need for *in vivo* models

Normal physiology in humans is dependent on the interactions between different cell types via cell–cell communication and signalling through many pathways and systems. This makes it crucial that diseases and associated therapies are studied in the *in vivo* setting. Genetic and/or environmental factors can perturb the body's natural homeostasis and thereby lead to the development of disease. Indeed, it is generally accepted that most diseases are caused by the complex interplay between our genes and the environment. Over time, somatic mutations accumulate in our cells, therefore, in reality, our genetic make-up is continuously varying, both temporally and spatially. The environment of an individual is also continuously changing. Thus, our ability to retain the homeostasis is constantly being challenged both from the outside and from within. To discover novel treatments for human diseases all of these aspects must be considered. Challenging experimental animals with specific environmental factors has been possible for a long time. With the development of transgenic technology it is now also possible to induce specific mutations in somatic or germ cells to mimic genetic alterations.

Transgenic technology has the potential to influence the attrition rate in pharmaceutical R&D by increasing the quality of both targets and compounds. One way it can increase the quality of the target is by demonstrating that the specific target is rate limiting for disease development or progression *in vivo* and at the same time not essential for vital physiological functions. Similarly, for the compound, transgenic technology can be used to generate tools (e.g. animals containing the human target molecule or improved models of human disease that are better at mimicking specific human pathologies) to demonstrate that the new potential drug has the desired effect *in vivo*. In addition, for the compound, animals humanized for drug metabolism pathways and toxicology reporter mice can be used to investigate whether the potential drug has overt metabolism or toxicological problems.

Overview of the drug discovery process

The process of generating new drugs needs to be described in some detail to be able to explain where transgenic technology can make an impact. The objective of the first step, target identification, is to discover a molecular target that might be appropriate for pharmaceutical intervention to prevent, stabilize or reverse disease. The next step of drug discovery, target validation, aims to determine whether the target identified in target identification is rate limiting for a given disease's progression or induction, while not being essential for vital physiological functions. The next phase after target validation is turning hits into leads. HTS is often carried out in this phase [10]. An HTS assay compatible with automation is required and must be developed to identify small molecular-weight compounds that interact in the desired way with the proposed drug target. Hundreds of thousands (or millions) of compounds are tested rapidly and hits are identified. These hits are then evaluated by chemists to give a subset of leads that form the starting point for chemistry. In the next phase, lead optimization, the lead molecules are modified to enhance specificity, safety, efficacy and pharmacokinetic properties. This stage requires testing the substance in systems that mimic the human context as closely as possible. After an effective and safe compound has been identified (based on *in vitro* studies and *in vivo* animal data), testing in humans can begin. First, small numbers of healthy volunteers will receive the substance (Phase I), followed by a small number of patients (Phase II) and finally a large patient cohort (Phase III).

The likelihood of success for a new potential drug is related to the quality of the target and the quality of the chemical compound. The quality of the target and the compound are equally important and both of these can be enhanced by the use of appropriate transgenic technologies.

Impact on target identification

There are several approaches to discover new potential drug targets. Two conceptually different approaches are the genomic approach and the genetic approach. The genomic approach [11] looks at the differential expression of mRNA in, for example, normal versus disease tissue or gene expression changes after a compound has been added to the culture medium of cells or given to an animal. Today,

many groups use microarray technology [12] and proteomics [13] to identify the changes in gene expression. This approach assumes that the expression of relevant targets will be different in diseased tissue or after treatment. The challenges with this technology are many [14], for example, several current drug targets are not up- or down-regulated in disease, only changes of a certain magnitude can be detected because of practical constraints, and the change in expression might be a consequence or a cause of the disease (which adds to the problem of 'false positives' that plagues this technique). In practice, the downstream analysis after array experiments is a major and expensive challenge.

One interesting approach in the post human genome sequencing era [15,16] is the genetic analysis of human and model organisms. The genetic approach is more long-term and labour intensive, but has greater potential (resource intensive but with the potential of yielding high returns). One of the main hurdles for this approach is that most major diseases are polygenic. For monogenic diseases, many of the underlying genetic alterations have been published [17–22]. Encouragingly, progress has now also been reported for several polygenic human diseases where some causative genes have been identified. For example, one of the loci for the polygenic Crohn's disease is mapped to chromosome 16. Further work revealed that the NOD2 gene was located in this region and mutations in this gene were found in patients with the disease [23]. This is one of the first examples where a gene contributing to a polygenic human disease has been identified using linkage analysis.

Genetic studies in humans are costly in the long term. An interesting complement to human genetic studies is genetic analysis in more amenable model organisms. A challenge is to make sure that the phenotype identified in the chosen model system is relevant to the human disease and that the biological pathways are similar in humans and the experimental species. For these reasons, the mouse is often selected as the model organism of choice. Most major human diseases can be mimicked in the mouse and most pathways are likely to be conserved between humans and mice. Recently, an interesting publication appeared that compared loci linked to hypertension identified in mice and humans. Six loci were found in mice and four of them were concordant with the human loci [24]. This finding becomes even more interesting when comparing the costs and times for human and mouse genetic studies. The mouse study could be performed five-times faster and cost just 1% of that of the human study [25]. The huge advantage of using genetic approaches instead of expression profiling genomic approaches is that there is a clear link between a genetic alteration and a disease phenotype. The gene identified is clearly a cause, not a consequence of disease. The disadvantage is the time required to generate the data.

In genetic studies a major challenge is to narrow down the genetic region to a size that will enable proper unbiased analysis of all the candidate genes in the region. In practice, this currently means that the genetic study has to deliver a region containing only five candidate genes. One means of reducing these regions to manageable sizes is to perform complementation studies using appropriate transgenic animals [26]. The concept is to introduce the large genomic region, defined by the genetic study, as a series of smaller fragments (isolated from bacterial artificial chromosomes; BACs) into the susceptible or resistant mouse strain, as appropriate for the phenotype followed in the genetic study. Fragments of up to 200 kb can readily be isolated from BAC clones and then used to make transgenic animals. The phenotype of these BAC transgenics is then assessed to see which of the genomic fragments carried the crucial gene(s) capable of influencing the phenotype being monitored. To increase the resolution still further, overlapping genomic fragments isolated from BACs can be used to narrow regions to less than the size of a single BAC insert.

One approach for increasing the genetic diversity and inducing random mutations in the genome is *N*-ethylinitrosourea (ENU) mutagenesis [27]. For many diseases, this could offer an ideal way of reducing polygenic diseases to monogenic traits. Using this approach enables researchers to select those genetic loci having the greatest impact on the polygenic disease. This approach has been used recently on a large scale in the mouse and interesting phenotypes have been discovered [28,29].

Impact on target validation

Target validation is a continuous critical process which begins early in the discovery phase. Target validation data to support a chosen target are generated using many different approaches often performed in parallel, but the ultimate validation of the target (along with the compound selected to interact with it) occurs only when the new potential drug is actually tested in patients (Phase IIB). Different validation methods have to be used during the different stages of the drug discovery process. In the early phases, when selections have to be made between a large number of targets, information from bioinformatics and data from experiments in cultured cells often proves useful. However, there rapidly becomes a need for *in vivo* target validation data before large amounts of resource are invested in the potential target. Before entering the lead optimization phase, when large resources are engaged in chemistry, results obtained from the modulation of target function

in vivo should ideally be available. Considerable efforts are being made to identify rapid, reliable and general tools for *in vivo* target validation but so far only transgenic animals have proved to be robust on a wide range of targets.

Genetically modified animals that either overexpress (gene addition) or no longer possess the target (knockout animals) are extremely useful in the target validation phase of the drug discovery process, often providing crucial *in vivo* functional data on a potential target. The alteration of expression of the target gene in genetically modified animals can be achieved with different levels of sophistication. The most rapid transgenic animal that can be produced is one in which a transgene of interest is overexpressed ubiquitously or in a tissue-specific manner. A large number of tissue-specific promoters have now been used to direct transgene expression to different cell types [30–34]. Two options for inhibiting gene function in a relatively rapid (compared with knockout technology) and specific way are to either overexpress a dominant-negative transgene or an antisense cDNA [35]. One example where the dominant-negative approach was used successfully is the model for epidermolysis bullosa [36]; however, this approach, as well as the antisense one, has only been successful in a limited number of cases.

Conditional systems

Several different systems are now available to regulate the expression of transgenes using external inducers. The first inducible systems used in transgenic animals made use of the methallothionein [4] and phosphoenolpyruvatecarboxykinase promoters [37]. These systems were difficult to regulate and there were problems with high levels of background expression. More recently, tertiary systems have become the inducible regulatory systems of choice. Of these, the most widely used, and therefore validated, is the tetracycline system [38–40], which has been employed in several successful studies [41–43]. Recently, the ecdysone system [44] and systems for the artificial dimerization of molecules [45] have been developed, but their broad utility for regulating transgenes still has to be proved. Also fusion proteins between an effector protein (e.g. Cre) and a ligand binding domain of a steroid receptor have been used to control the activity of transgenes (see section on Cre–lox later).

Initially, all knockouts produced were complete knockouts or germ-line knockouts. These mutant animals lack a given gene function from conception and throughout their embryonic and adult lives. This work has yielded great insights into the *in vivo* functions of a wide range of genes including developmental genes [46,47], growth factors [48–50], molecules involved in signalling transduction [51] and many more. However, more recently, the need to be able to make tissue restricted and/or inducible knockouts has increased. These animals can often enable researchers to overcome an embryonic lethal phenotype of a homozygous-null animal or concerns surrounding the possible induction of compensatory mechanisms resulting from the chronic absence of a gene-encoded function. Other justifications often cited for inducible knockouts include avoidance of developmental phenotypes and to better mimic the type of gene expression changes occurring in late-onset human pathologies.

The most widely used system today for the production of conditional knockouts is the Cre–lox system from bacteriophage P1 [52,53] with the flp/frt system from yeast [54] being the second choice, having also been deployed successfully. Both systems involve expression of a site-specific recombinase coupled to recombinase recognition sites strategically inserted into target genes. Control is achieved at the level of recombinase expression with the use of transient delivery methods or tissue-specific or inducible promoters. Several tissue-specific and inducible Cre mouse lines have been generated and a comprehensive list can be found at http://www.mshri.on.ca/nagy/. The Cre–lox system has now been used in several successful studies to generate null mutations in specific genes in particular cell types (Table 1) [55,56] and at certain time points (Table 2) [57–59]. A great advantage with the Cre–lox system is that it is possible to mutate your gene of interest in specific tissues by breeding 'target' animals containing floxed alleles of a particular gene with animals expressing the Cre enzyme in the desired cell type. Floxed alleles are genes with two loxP sequences (34 bp) inserted into them in a way that does not influence normal gene function if Cre is not expressed (e.g. within an intron). This means that in cells that are not expressing Cre, the floxed genes will be silenced. In this way, it is possible to mutate the gene in specific regions of the CNS, different endocrine and exocrine glands, skin, liver, and many more tissues (Table 1).

To generate knockouts that are regulated both spatially and temporally, one approach that has proved successful is to fuse the Cre protein with a mutant form of the ligand binding domain (LBD) of the oestrogen receptor. A tissue-specific promoter can then be used to drive expression of the fusion protein in a particular subset of cells restricted by the characteristics of the promoter. This hybrid protein remains sequestered and inactive until tamoxifen is given to the animals. Tamoxifen-induced Cre activity then deletes the floxed region of the target gene, thus removing its function, but only in those cell types where the LBD–Cre is expressed [60–62]. As described previously for the spatially controlled knockouts, animals in which the gene of interest is floxed can be mated to animals with a temporal and

Table 1. Spatial regulation of gene targeting

Cell type	Promoter/enhancer
Adipocytes	aP2
Apical ectodermal ridge	Retinoic acid receptor (RAR)
Astrocytes and subpopulation of neural precursors	Glial fibrillary acidic protein (GFAP)
B cells	CD19
Basal cells of epidermis	K5
Brain	CamKII-α
Cardiac myocytes	α-Myosin heavy chain
Cerebellar granule cell	GluRepsilon3
Cerebellar Purkinje cells	L7/pcp-2
Cerebral cortex and hippocampus	Emx1
Chondrocyte	Collagen type II
Epiblast	Mox2
Extra-embryonic mesoderm, lateral plate and limb mesoderm, midbrain–hindbrain junction	Hoxb6
Floorplate, notochord, gut epithelium	Hoxa1
Gut epithelium	Fatty acid binding protein (FABP)
Hepatocyte	Albumin
Hindbrain, bone, peripheral nervous system	Krox20
Immature T cells	lck
Keratinocytes of epidermis, tongue, and oesophagus	K14
Mammary epithelium	β-Lactoglobulin
Mammary epithelium	Whey acidic protein (WAP)
Mid-hindbrain	Engrailed-2
Myeloid cells	LysM
Neural crest	Pax3
Neural crest	Wnt1
Neural crest	Protein O
Neuronal cell of CNS	KA1
Neuronal lineage	Nestin
Oligodendrocytes	Myelin basic protein
Oocyte	ZP3
Pancreas	Pdx1
Pancreatic alpha	Glucagon
Pancreatic beta	Insulin
Pancreas	Pancreatic polypeptide
Primordial germ cells	Tissue non-specific alkaline phosphatase (TNAP)
Prostate epithelium	Probasin
Retina and ventral forebrain	Six3
Skeletal and cardiac muscle	Muscle creatinine kinase (MCK)
Skeletal muscle only	Myosin light chain 1f
Spermatocytes	Pgk-2
Spermatogenesis	Protamine-1 (Prm1)
Striated muscle	α-Skeletal actin

spatial control of Cre enzyme expression, to enable spatial and temporal gene targeting to be made. In this way, genes have been inactivated at specific time points in lymphocytes, smooth muscle cells, liver and skin. The tetracycline system, interferon induction or RU486 induction have also been used successfully to temporally control Cre-mediated gene inactivation in the mammary gland, gut, liver and CNS (Table 2).

In another interesting approach, a doxycycline-controlled, autoinducible, ubiquitous Cre expression system has been developed for temporally controlled general inactivation of target genes. Although these systems are interesting, the most widely used system today is the tamoxifen-induced LBD–Cre fusion protein. (Table 2). Use of this or similar systems could bring us closer to the generation of the optimal knockout model for many pharmaceutical studies: an inducible, tissue-specific knockout that does not suffer from mosaicism.

Impact on lead optimization phase: new optimal disease models

To minimize the attrition rate in clinical trials, potential drugs need to be tested in experimental systems that resemble human pathology (and physiology) as closely as possible. One way to achieve this is to introduce genetic modifications to make high-fidelity animal models resembling human pathology. The degree of comparability between the models and the human disease is generally determined both functionally and morphologically, but we now have the ability to augment this analysis using microarray technology to compare the mRNA expression 'fingerprints' of pathologic tissue from humans with a given mouse model.

For some diseases it might not be possible to mimic the complete human pathology in the mouse, and for others it might be more desirable to create a model of a specific biological process such as angiogenesis, metastasis or plaque rupture. These are all processes that are believed to be good 'targets' for intervention. It should also be noted that these processes are not solely pathological (e.g. metastasis) but can be physiological (e.g. angiogenesis).

Table 2. Temporal (and spatial) regulation of gene targeting

Tamoxifen inducible expression		Interferon inducible expression		RU486 inducible expression		Tetracycline inducible expression	
Cell type	**Promoter/ enhancer**	**Cell type**	**Promoter/ enhancer**	**Cell type**	**Promoter/ enhancer**	**Cell type**	**Promoter/ enhancer**
B-cell-specific expression	CD19	Primarily liver and lymphocytes	Msx-1	Cortex and hippocampus	CamKII-α and Thy-1	Mammary epithelium	Whey acidic protein (WAP)–rtTA and TetO–Cre
Basal cells of epidermis	K5					Gut epithelium	Fatty acid binding protein (FABP)–rtTA and TetO–Cre
Hepatocyte	α1-Antitrypsin					Ubiquitous (auto-inducible)	Doxycycline-responsive promoter
Keratinocytes of epidermis, tongue, and oesophagus	K14						
Neural crest	Wnt1						
Smooth muscle	SM22						

Transgenic technology has resulted in the development of several disease models and some of these have enabled the testing of new treatment paradigms.

Oncology

In oncology, transgenic and knockout animals have been most important in understanding the function of oncogenes and tumour suppressor genes [6]. In addition to these fundamental studies, transgenic technology has been used in more detailed studies that might be closer to clinical applications.

The fatal outcome of cancer is often caused by metastases from the primary tumour and, therefore, understanding the underlying mechanisms behind metastasis could provide valuable therapeutic opportunities. The role for adhesion molecules in metastasis has been studied in a model with highly defined multistage tumour progression based on the transgenic expression of the SV40 large T-antigen in Langerhan's cells of the pancreas (Rip1Tag2 model) [6]. Transgenic animals expressing a truncated dominant-negative E-cadherin gene in the Rip1Tag2 transgenic mouse showed an essential role for E-cadherin in tumour progression and metastasis [63]. The fundamental role for another cell adhesion molecule, NCAM, in metastasis was demonstrated in a cross between the Rip1Tag2 model and a NCAM knockout [64].

The inhibition of angiogenesis is currently considered to be one of the biggest opportunities for new cancer therapies. It could be possible to inhibit angiogenesis in adults with considerably fewer side effects than are seen, for example, with general cell-cycle inhibitors or agents that block protein synthesis. Inhibitors of angiogenesis have already been identified that act on different mechanisms within the angiogenic pathway, and it has become clear that these different inhibitors could be of importance at different stages of tumour progression. This suggestion is based on further work with the Rip1Tag2 multistage tumour model previously mentioned. The inhibitors all had distinct efficacy profiles at the different tumour stages tested [65] suggesting that different angiogenesis inhibitors should be used at specific disease stages.

Expression of the transgene H-Ras^{V12G} in animals with a mutant *INK4a* tumour suppressor gene results in the formation of malignant invasive melanomas. In this study, the Ras transgene is expressed using a system regulated by the administration of doxycycline (mentioned previously). Withdrawal of the doxycycline results in downregulation of the Ras transgene and regression of the large tumours [66]. This model is being used to study the general mechanisms behind tumour regression as well as the role for Ras in tumour maintenance. In addition, the findings with this model suggest that Ras could be a good drug target itself, at least for advanced melanomas where prognosis is poor.

Cardiovascular disease

There is a great need for relevant disease models of cardiovascular disease because the pathogenic process is prolonged, thus making it difficult to measure any beneficial effects of a potential new therapy. Transgenic animals have been important for studying the pathology of cardiovascular disease [9] and for developing treatments in spite of the challenges faced when using mice for cardiovascular studies because of their small body size. The understanding of hypertension, lipoprotein regulation, obesity, insulin resistance and

minimizing infarction size after post-ischaemic reperfusion have all been advanced by the use of transgenic technology.

Non-transgenic mice only develop blood-vessel lesions that mimic the initial stages of human atherosclerosis (fatty streaks) partly because of their high circulating levels of high-density lipoprotein (HDL) and low levels of low-density lipoprotein (LDL). Mutation of the ApoE gene [67], which is crucial for the uptake of chylomicrons and very-low-density lipoprotein (VLDL) particles, results in a model that develops atherosclerotic lesions that are histologically similar to those found in humans. Deletion of the LDL-receptor gene in isolation [68], or when combined with the ApoE deletion in double knockout animals [69], can further refine the phenotype. The double knockout animals have high VLDL levels and lower HDL levels. However, one difference between this mouse model and the disease in humans is that more lesions are seen in the female animals whereas pathological alterations are much more common in men than in premenopausal women. This inconsistency could be partly a result of the difference between mice and humans in the modification of the apolipoprotein ApoB100. This is carried out by an enzyme called apobec-1 and results in the formation of ApoB48. In humans, apobec-1 is not expressed in the liver, whereas it is in mice. Knocking out the apobec-1 gene and breeding it onto the LDLr$^{-/-}$ and ApoE$^{-/-}$ background to generate triple knockout animals gives an even better lipoprotein profile and the gender distribution of lesions now mirrors that seen in the human situation [70].

Alzheimer's disease

The need for good animal models to study Alzheimer's disease is clear; it is currently extremely difficult to imagine a cell culture system that goes any way towards modelling the complexity of the brain. Alzheimer's disease has complex symptoms and clinical readouts, the disease also progresses slowly and it is difficult to access human pathological and control tissue necessary for research into this disease. No animal models existed for the disease before transgenic technology was employed, but today several transgenic models have been established. These models resemble much of the human pathology and are frequently used in the search for new therapeutic opportunities.

Recently a paper was published by Schenk *et al.* [71] where a new approach for therapy was tested in transgenic mice. Immunization of amyloid precursor protein (APP) transgenic mice with the protein Aβ42 before disease establishment resulted in disease prevention, whereas, if performed on older animals, immunization inhibited disease progression. This study using transgenic mice suggests that vaccination against Alzheimer's disease could have potential as a therapeutic approach.

Safety and pharmacokinetic models

It is not only desirable to mimic human pathology in a model system, but it is also useful to be able to mimic the human physiology and the response to compounds using models. One way of achieving this is by introducing human genes into the mouse. The genes introduced might be the human target genes or human drug-metabolism genes. Several human drug-metabolism genes have been introduced into mice, for example, human cytochrome P450 4B1 (*CYP4B1*) was expressed in mice and the resultant enzyme was biologically active [72]. Also, some endogenous mouse P450 genes have been mutated to help elucidate their function [73] and the *CYP1A2*, *CYP2E1* and *CYP1B1* [74–78] were mutated to enable studies of their specific role in toxicology.

One area where several transgenic models have been evaluated more systematically is in the field of compound toxicity, and in particular carcinogenicity. Several models now exist that can potentially provide early warning of compound carcinogenicity. The benefit of these models is that they respond more rapidly to carcinogenic agents (6–9 months instead of two years) saving both time and also potentially reducing the number of animals used because of fewer animals developing age-related disease in the control group. The value of these models was recently investigated in a large multicentre evaluation co-ordinated by the International Life Sciences Institute (ILSI), where four different transgenic models were evaluated for their potential to predict the carcinogenicity of different compounds in humans [79]. The results of this study showed that two of the models, the RasH2 model and the p53$^{+/-}$ model, were particularly promising although further evaluation is needed. These models were capable of responding to known human carcinogens and so might serve as a complement to the conventional two-year rat carcinogenicity studies.

Revival of the physiologists

During the past 20 years, molecular geneticists and embryologists have developed and refined the methods to modify the genome of mice in an unprecedented way, modifying gene expression both in time and space. The success has been enormous and the reward in terms of publications has been great. The next 20 years will belong to the physiologists, analyzing what has been, and still is being, generated. The impact of these analyses on our understanding of biology, particularly human, will be great.

The mouse is, for several reasons, the most attractive model species to be used in the functional analysis of genes. It is the only mammalian species in which genetically-modified animals can be routinely generated with precise alterations to their genome and it is the second

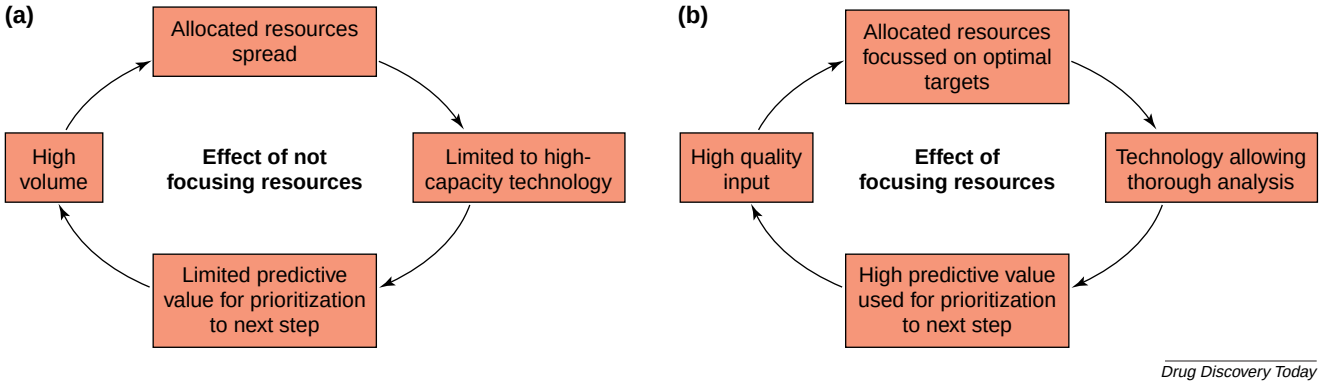

Figure 1. The impact of choosing a high-quantity model or a high-quality model on the type of tools that can be used in the R&D process. Working with a high-quantity model **(a)** will restrict the use to high-throughput techniques with low predictive capability. Decisions cannot be made to select the targets or compounds that are most likely to eventually succeed, which means that a large number of targets or compounds must be analyzed in the following step. This will again restrict the types of techniques available and a 'domino' effect might be evident. If a model of high-quality is adopted **(b)**, tools that enable the selection of the best targets or compounds in each step can be used. This will impact on the number of targets or compounds in succeeding steps making it possible to use techniques to make firm decisions.

mammalian species (after humans) in which we will know the complete DNA sequence. Furthermore, it shares the majority of pathological conditions that occur in humans. There are, however, challenges with using the mouse as an experimental species, many related to its small body size. The miniaturization of methods for analysis of complex physiological functions and analysis of small sample volumes is currently being developed to meet these challenges.

The speed at which the analytical tools have been generated for mice has been extremely impressive. Analyses that were considered impossible just a few years ago are today mastered to perfection by skilled groups. There are several beautiful studies on mice where methods have been adopted from larger species. Examples are found in the cardiovascular field [80,81], in behavioural research [82], in studies on pulmonary function [83,84], in metabolism [85] and in many more areas of research.

Future perspectives

Improving the quality of the targets and the compounds progressing along the R&D pipeline is a priority for drug companies to reduce and focus the money they have to spend on bringing high-quality medicines to market. One way of addressing this quality issue is to fully use the power of transgenic technology in the context of the invaluable post-genomic information. Using this technology, one can generate an exciting array of animal models specifically designed to address the function of targets and the action of compounds.

Currently, most pharmaceutical companies operate to a 'high-quantity' model where relatively large numbers of targets, mainly identified through bioinformatics and expression profiling efforts, are validated to varying extents and progressed to the HTS stage resulting in multiple lead series. Optimized lead compounds are then used in clinical trials where the attrition rate is generally high. Success in this type of operation is achieved by keeping both the numbers of screened targets and the number of compound series entering the clinical trial phase at a high level. This model restricts the technical arsenal to high-throughput methods that usually have low predictive value for the subsequent step in the process (Fig. 1). This can result in a 'domino effect'; high numbers require the use of techniques with lower predictive value, this leads to an inability to make firm selections for progression to the next step, and so higher numbers have to be used, and so on, at each stage in the process. This reasoning can be made for each and every step of the R&D process. This model is extremely costly because of the in-built high rates of attrition and it is our belief that greater use of transgenic models could reduce the required throughput for achieving success and thereby significantly impact on costs.

References

1 Gordon, J.W. *et al.* (1980) Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci. U. S. A.* 77, 7380–7384
2 Gordon, J.W. and Ruddle, F.H. (1981) Integration and stable germ line transmission of genes injected into mouse pronuclei. *Science* 214, 1244–1246
3 Jaenisch, R. (1976) Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proc. Natl. Acad. Sci. U. S. A.* 73, 1260–1264
4 Palmiter, R.D. *et al.* (1982) Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* 300, 611–615

5 Jafarian-Tehrani, M. and Sternberg, E.M. (1999) Animal models of neuroimmune interactions in inflammatory diseases. *J. Neuroimmunol.* 100, 13–20
6 Hanahan, D. and Weinberg, R.A. (2000) The hallmarks of cancer. *Cell* 100, 57–70
7 Betsholtz, C. *et al.* (2001) Developmental roles of platelet-derived growth factors. *BioEssays* 23, 494–507
8 Pepeu, G. (2001) Overview and perspective on the therapy of Alzheimer's disease from a preclinical viewpoint. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 25, 193–209
9 Carmeliet, P. and Collen, D. (2000) Transgenic mouse models in angiogenesis and cardiovascular disease. *J. Pathol.* 190, 387–405
10 Kenny, B.A. *et al.* (1998) The application of high-throughput screening to novel lead discovery. *Prog. Drug Res.* 51, 245–269
11 Debouck, C. and Goodfellow, P.N. (1999) DNA microarrays in drug discovery and development. *Nat. Genet.* 21(Suppl.), 48–50
12 Lockhart, D.J. and Winzeler, E.A. (2000) Genomics, gene expression and DNA arrays. *Nature* 405, 827–836
13 Pandey, A. and Mann, M. (2000) Proteomics to study genes and genomes. *Nature* 405, 837–846
14 Miles, M.F. (2001) Microarrays: lost in a storm of data? *Nat. Rev. Neurosci.* 2, 441–443
15 Venter, J.C. *et al.* (2001) The sequence of the human genome. *Science* 291, 1304–1351
16 Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
17 Futreal, P.A. *et al.* (1994) BRCA1 mutations in primary breast and ovarian carcinomas. *Science* 266, 120–122
18 Albertsen, H.M. *et al.* (1994) A physical map and candidate genes in the BRCA1 region on chromosome 17q12-21. *Nat. Genet.* 7, 472–479
19 Group, T.H.s.D.C.R. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 72, 971–983
20 Miki, Y. *et al.* (1994) A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266, 66–71
21 Wooster, R. *et al.* (1994) Localization of a breast cancer susceptibility gene, *BRCA2,* to chromosome 13q12-13. *Science* 265, 2088–2090
22 Riordan, J.R. *et al.* (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245, 1066–1073
23 Hugot, J.P. *et al.* (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411, 599–603
24 Sugiyama, F. *et al.* (2001) Concordance of murine quantitative trait loci for salt-induced hypertension with rat and human loci. *Genomics* 71, 70–77
25 Paigen, K. and Eppig, J.T. (2000) A mouse phenome project. *Mamm. Genome* 11, 715–717
26 Probst, F.J. *et al.* (1998) Correction of deafness in shaker-2 mice by an unconventional myosin in a BAC transgene. *Science* 280, 1444–1447
27 Vitaterna, M.H. *et al.* (1994) Mutagenesis and mapping of a mouse gene, *Clock*, essential for circadian behavior. *Science* 264, 719–725
28 Nolan, P.M. *et al.* (2000) A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat. Genet.* 25, 440–443
29 Hrabe de Angelis, M.H. *et al.* (2000) Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nat. Genet.* 25, 444–447
30 Morimoto, S. *et al.* (2001) Elevated blood pressure in transgenic mice with brain-specific expression of human angiotensinogen driven by the glial fibrillary acidic protein promoter. *Circ. Res.* 89, 365–372
31 Lin, F. *et al.* (2001) Targeted alpha(1A)-adrenergic receptor overexpression induces enhanced cardiac contractility but not hypertrophy. *Circ. Res.* 89, 343–350
32 Milano, C.A. *et al.* (1994) Enhanced myocardial function in transgenic mice overexpressing the beta 2-adrenergic receptor. *Science* 264, 582–586
33 Oberdick, J. *et al.* (1990) A promoter that drives transgene expression in cerebellar Purkinje and retinal bipolar neurons. *Science* 248, 223–226
34 Field, L.J. (1988) Atrial natriuretic factor-SV40 T antigen transgenes produce tumors and cardiac arrhythmias in mice. *Science* 239, 1029–1033
35 King, L.B. *et al.* (1995) A targeted glucocorticoid receptor antisense transgene increases thymocyte apoptosis and alters thymocyte development. *Immunity* 3, 647–656
36 Fuchs, E. and Coulombe, P.A. (1992) Of mice and men: genetic skin diseases of keratin. *Cell* 69, 899–902
37 Lim, I.K. *et al.* (1990) High level, regulated expression of the chimeric P-enolpyruvate carboxykinase (GTP)-bacterial O6-alkylguanine-DNA alkyltransferase (ada) gene in transgenic mice. *Cancer Res.* 50, 1701–1708
38 Ewald, D. *et al.* (1996) Time-sensitive reversal of hyperplasia in transgenic mice expressing SV40 T antigen. *Science* 273, 1384–1386
39 Gossen, M. and Bujard, H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A.* 89, 5547–5551
40 Furth, P.A. *et al.* (1994) Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter. *Proc. Natl. Acad. Sci. U. S. A.* 91, 9302–9306
41 Zhu, Z. *et al.* (2001) Use of the tetracycline-controlled transcriptional silencer (tTS) to eliminate transgene leak in inducible overexpression transgenic mice. *J. Biol. Chem.* 276, 25222–25229
42 Perea, J. *et al.* (2001) Induced myelination and demyelination in a conditional mouse model of Charcot-Marie-Tooth disease type 1A. *Hum. Mol. Genet.* 10, 1007–1018
43 Mansuy, I.M. *et al.* (1998) Inducible and reversible gene expression with the rtTA system for the study of memory. *Neuron* 21, 257–265
44 No, D. *et al.* (1996) Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.* 93, 3346–3351
45 Amara, J.F. *et al.* (1997) A versatile synthetic dimerizer for the regulation of protein–protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* 94, 10618–10623
46 Schneider-Maunoury, S. *et al.* (1993) Disruption of Krox-20 results in alteration of rhombomeres 3 and 5 in the developing hindbrain. *Cell* 75, 1199–1214
47 Studer, M. *et al.* (1996) Altered segmental identity and abnormal migration of motor neurons in mice lacking Hoxb-1. *Nature* 384, 630–634
48 Bostrom, H. *et al.* (1996) PDGF-A signaling is a critical event in lung alveolar myofibroblast development and alveogenesis. *Cell* 85, 863–873
49 Lindahl, P. *et al.* (1997) Pericyte loss and microaneurysm formation in PDGF-B-deficient mice. *Science* 277, 242–245
50 Liu, J.P. *et al.* (1993) Mice carrying null mutations of the genes encoding insulin-like growth factor I (Igf-1) and type 1 IGF receptor (Igf1r). *Cell* 75, 59–72
51 Scita, G. *et al.* (1999) EPS8 and E3B1 transduce signals from Ras to Rac. *Nature* 401, 290–293
52 Sauer, B. (1994) Site-specific recombination: developments and applications. *Curr. Opin. Biotechnol.* 5, 521–527
53 Snaith, M.R. *et al.* (1996) An *Escherichia coli* system for assay of F1p site-specific recombination on substrate plasmids. *Gene* 180, 225–227
54 O'Gorman, S. *et al.* (1991) Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* 251, 1351–1355
55 Gu, H. *et al.* (1994) Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* 265, 103–106
56 Giovannini, M. *et al.* (2000) Conditional biallelic Nf2 mutation in the mouse promotes manifestations of human neurofibromatosis type 2. *Genes Dev.* 14, 1617–1630
57 Kuhn, R. *et al.* (1995) Inducible gene targeting in mice. *Science* 269, 1427–1429
58 Sjogren, K. *et al.* (2001) Liver-derived IGF-I is of importance for normal carbohydrate and lipid metabolism. *Diabetes* 50, 1539–1545
59 Sjogren, K. *et al.* (1999) Liver-derived insulin-like growth factor I (IGF-I) is the principal source of IGF-I in blood but is not required for postnatal body growth in mice. *Proc. Natl. Acad. Sci. U. S. A.* 96, 7088–7092

60 Metzger, D. and Chambon, P. (2001) Site- and time-specific gene targeting in the mouse. *Methods* 24, 71–80

61 Imai, T. *et al.* (2001) Impaired adipogenesis and lipolysis in the mouse upon selective ablation of the retinoid X receptor alpha mediated by a tamoxifen-inducible chimeric Cre recombinase (Cre-ERT2) in adipocytes. *Proc. Natl. Acad. Sci. U. S. A.* 98, 224–228

62 Brocard, J. *et al.* (1997) Spatio-temporally controlled site-specific somatic mutagenesis in the mouse. *Proc. Natl. Acad. Sci. U. S. A.* 94, 14559–14563

63 Perl, A.K. *et al.* (1998) A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* 392, 190–193

64 Perl, A.K. *et al.* (1999) Reduced expression of neural cell adhesion molecule induces metastatic dissemination of pancreatic beta tumor cells. *Nat. Med.* 5, 286–291

65 Bergers, G. *et al.* (1999) Effects of angiogenesis inhibitors on multistage carcinogenesis in mice. *Science* 284, 808–812

66 Chin, L. *et al.* (1999) Essential role for oncogenic Ras in tumour maintenance. *Nature* 400, 468–472

67 Plump, A.S. *et al.* (1992) Severe hypercholesterolemia and atherosclerosis in apolipoprotein E-deficient mice created by homologous recombination in ES cells. *Cell* 71, 343–353

68 Ishibashi, S. *et al.* (1993) Hypercholesterolemia in low density lipoprotein receptor knockout mice and its reversal by adenovirus-mediated gene delivery. *J. Clin. Invest.* 92, 883–893

69 Ishibashi, S. *et al.* (1994) The two-receptor model of lipoprotein clearance: tests of the hypothesis in 'knockout' mice lacking the low density lipoprotein receptor, apolipoprotein E, or both proteins. *Proc. Natl. Acad. Sci. U. S. A.* 91, 4431–4435

70 Powell-Braxton, L. *et al.* (1998) A mouse model of human familial hypercholesterolemia: markedly elevated low density lipoprotein cholesterol levels and severe atherosclerosis on a low-fat chow diet. *Nat. Med.* 4, 934–938

71 Schenk, D. *et al.* (1999) Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* 400, 173–177

72 Imaoka, S. *et al.* (2001) A transgenic mouse expressing human CYP4B1 in the liver. *Biochem. Biophys. Res. Commun.* 284, 757–762

73 Gonzalez, F.J. and Kimura, S. (1999) Role of gene knockout mice in understanding the mechanisms of chemical toxicity and carcinogenesis. *Cancer Lett.* 143, 199–204

74 Liang, H.C. *et al.* (1996) Cyp1a2(–/–) null mutant mice develop normally but show deficient drug metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1671–1676

75 Lee, S.S. *et al.* (1996) Role of CYP2E1 in the hepatotoxicity of acetaminophen. *J. Biol. Chem.* 271, 12063–12067

76 Sinclair, P.R. *et al.* (1998) Uroporphyria produced in mice by iron and 5-aminolaevulinic acid does not occur in Cyp1a2(-/-) null mutant mice. *Biochem. J.* 330, 149–153

77 Buters, J.T. *et al.* (1999) Cytochrome P450 CYP1B1 determines susceptibility to 7,12-dimethylbenz[a]anthracene-induced lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1977–1982

78 Heidel, S.M. *et al.* (2000) Cytochrome P4501B1 mediates induction of bone marrow cytotoxicity and preleukemia cells in mice treated with 7,12-dimethylbenz[a]anthracene. *Cancer Res.* 60, 3454–3460

79 MacDonald, J.R.D. and Pettit, S. (2001) *An Evaluation of Alternatives to Carcinogenicity Testing*, ILSI Press

80 Heymans, S. *et al.* (1999) Inhibition of plasminogen activators or matrix metalloproteinases prevents cardiac rupture but impairs therapeutic angiogenesis and causes cardiac failure. *Nat. Med.* 5, 1135–1142

81 Wikstrom, L. *et al.* (1998) Abnormal heart rate and body temperature in mice lacking thyroid hormone receptor alpha 1. *EMBO J.* 17, 455–461

82 Steele, P.M. *et al.* (1998) Using genetic mutations to study the neural basis of behavior. *Cell* 95, 879–882

83 Hardie, W.D. *et al.* (1997) Postnatal lung function and morphology in transgenic mice expressing transforming growth factor-alpha. *Am. J. Pathol.* 151, 1075–1083

84 Clark, J.C. *et al.* (1997) Decreased lung compliance and air trapping in heterozygous SP-B-deficient mice. *Am. J. Respir. Cell Mol. Biol.* 16, 46–52

85 Naveilhan, P. *et al.* (1999) Normal feeding behavior, body weight and leptin response require the neuropeptide Y Y2 receptor. *Nat. Med.* 5, 1188–1193