Key animal models for the identification and validation of drug targets

Patrick Nef

Target identification and validation procedures represent a bottleneck in the drug discovery process owing to the low-throughput of most animal models. This review will describe the key model organisms of bacteria, yeast, worm, fly, mouse and human, their relevant uses, and the associated technologies available for each model.

Patrick Nef
Vice Director
Molecular Neurobiology
F. Hoffmann-La Roche
Pharmaceutical Division
Preclinical Research
CH-4070 Basel
Switzerland
tel: +41 61 688 08 28
fax: +41 61 688 17 20
e-mail:
patrick.nef@roche.com

▼ The human genome contains ~30,000 genes that could encode >1,000,000 different proteins via RNA editing, alternative splicing, and post-translational modifications. To date, only 500 gene products have been identified as molecular drug targets to treat human illnesses. A theoretical number of at least 5,000–15,000 potential gene products (or molecular drug targets) has been proposed that could lead to more effective or selective therapies[1]. The pharmaceutical industry and biotechnology companies are now heavily focussed on using tools that can provide a better understanding of the function or product of a gene, and that enable the rapid identification and validation of a human drug target among numerous potential candidates. Potential therapeutics could be not only small chemical drug molecules that modulate the function of a protein but also the gene products themselves.

The use of phylogenetically lower model organisms to mimic human diseases has become very popular as it enables either the identification of a human gene product (or pathway) that is directly involved in a disease state, or the development of biological screens for molecules or gene products that suppress the disease or stop its progression.

The mouse, despite its very low throughput, remains the organism of choice for many close functional parallels with human diseases. However, emerging and promising alternatives include the complementary use of other key organisms such as *Caenorhabdidis elegans*, *Drosophila*, *Escherichia coli* or yeast. These organisms are amenable to higher functional throughput for the validation of several gene products or of different molecular drug targets in parallel, as well as for the identification of signaling pathways involved in a pathological condition.

Speeding up target identification using animal models

From the simplest bacteria to the highly complex primate, animal models have traditionally been used to understand biological functions such as transcription, translation, replication, viral infection, development, or genetic transmissions. Recently, several models have been developed to mimic complex human diseases and to help in the rapid identification and validation of potential drug targets. The key model organisms such as bacteria, yeast, worm, fly, mouse and human are widely used in pre-clinical drug research discovery to enable a better understanding of gene-product function, and to identifiy and validate potential molecular drug targets or pathways. Alternative species with narrower interests are not discussed within the scope of this review, but include: *Arabidopsis thaliana* (mustard plant), *Chlamydomonas reinhardtii* (unicellular algae), cyanobacteria, *Dictyostelium* (slime mold), maize, *Aplysia californica* (snail), *Ascidians* (primitive chordate), *Cnidarians/Coelenterates* (jellyfish, hydra), sea urchin, mosquito, *Fugu rubripes* (pufferfish), *Medaka* (rice fish), *Danio rerio* (zebra fish), *Xenopus laevi* (frog), avian species (chick, quail), and other mammals (rabbit, dog, pig, sheep, cow, horse, cat, monkey).

1359-6446/01/$ – see front matter PII: S1471-1931(01)00165-3

Several dozen animal models have been developed in total. However, compared with the high number of novel genes and proteins that have been recently identified, current model organisms lack the high-throughput necessary to handle thousands of potential targets simultaneously. This situation is unlikely to change, despite the high-throughput generation of mutant mice using ENU mutagenesis coupled with rapid mutation detection methods. It is therefore essential to develop strategies that will enable fast and rapid prioritization of numerous potential targets. The use of complementary and higher-throughput model organisms should help to reduce this bottleneck and enable the early phase of drug discovery to be accelerated.

Key animal models for target identification and validation

Here, each key model is discussed in order of evolutionary biological complexity. Each organism will be described according to its traditional use as a model, and its specific characteristics regarding comparative genomics, expression profiling, transgenic and knockout technologies, genetic tools, and availability of disease models.

Escherichia coli (bacterium)

Representing one of the simplest unicellular systems, *E. coli* is composed of 4.6 million nucleic acid base pairs (4,403 genes located on 1 chromosome[2]; see Table 1), and has been sequenced in its entirety. It has a fast rate of replication (<1 h) and can be genetically engineered to overproduce proteins, vitamins, or DNA. In certain situations, the organism can be used to study virulence and pathogenicity targets. Perhaps more importantly, it can become a human pathogen. It is used for the development of antibiotics, or to determine the early mutagenesis rate or genetic toxicity of chemical molecules using the Ames test[3]. To date, 45 microbial genomes have been sequenced (37 bacterial species with 1–5 $\times$ 10^6 base pairs each) and have been used for comparative pathogenesis studies and vaccine development[4]. Expression profiling between pathogenic and non-pathogenic species enables the identification of gene products or pathways directly involved in resistance or metabolic defects[5]. Transgenesis via transformation or viral infection, and gene targeting via homologous recombination using the pKO3 vector strategy[6] are fast, efficient, cost-effective and simple. Recently, a bacterial system has been developed for the mapping of protein–protein interactions, a technique first developed in yeast[7]. Examples of conserved functions between prokaryotes and eukaryotes include RNA polymerases, rRNA involved in protein synthesis, transport of ions, and catabolic and anabolic enzymes[8]. Despite its simplicity (90% of its genome encodes proteins), about 50% of *E. coli* gene products have unknown functions[9]. The model is therefore of limited use for mimicking human diseases, although key proteins (i.e. MutH, MutS, MutL) involved in DNA mismatch repair systems are highly conserved from bacteria to humans[10], indicating that some basic mechanisms of mammal cell biology can be readily studied in simple organisms.

Saccharomyces cerevisiae (yeast)

S. cerevisiae is one of the simplest unicellular eukaryotic cells, and together with *Schizosaccharomyces pombe*, which is more closely related to man, are the most commonly used yeast species, particularly as they have a fast rate of replication. The *S. cerevisiae* genome is composed of 12 million base pairs (6,190 genes located on 16 chromosomes; see Table 1) and was entirely sequenced in 1996 (Ref. 11). Of these 6,190 potential gene-products, 30% encode proteins with unknown functions[12]. One-third of yeast genes have a robust homology to human genes, and two-thirds have at least one domain with a significant homology. Of 48 human genes identified by positional cloning, 25% had significant homology with yeast[13]. DNA microarrays containing all potential yeast genes have been used to study their level of expression[14]. Large-scale or genome-wide knockout and phenotypic approaches (using transposon-tagging and gene disruptions) are currently being used to generate high-throughput mutants for each yeast gene or gene product[15]. Yeast has become a major tool for mapping protein–protein interactions, which enables the classification of signaling pathways and the segregation of interacting partners among thousands of gene products. A global, large-scale initiative, using the yeast two-hybrid system has been used to determine the network of protein–protein interactions[16] and biological pathways[17]. As a model organism, yeast will probably provide basic knowledge for universal intracellular mechanisms (e.g. metabolism, signal transduction, membrane regulation, gene activation, oxidative stress, drug metabolism, cell cycle regulation)[18], but not for human diseases involving multiple cellular networks or complex pathways. Two examples of human genes with identified yeast homologs are described below.

PI3 kinase (termed *ATM* in humans and *Tel1* in yeast), which causes ataxia telangiectasia (AT), a human disorder characterized by cerebellar ataxia and oculocutaneous defects. Yeast cells with mutations in the *Tel1* and related gene pathway show a similar phenotype to cells from AT patients[19].

Metal Resistance Protein (termed *CFTR* in human, and *Ycf1* in yeast), which causes cystic fibrosis. Two *CFTR* defective alleles cause chronic lung problems and digestive disorders. *Ycf1* is very similar to the cystic fibrosis transmembrane regulator (CFTR) and multidrug resistance protein (MRP). Mutations associated with *CFTR* alleles, once introduced into *Ycf1*, elicit defects analogous to those seen with CFTR (Ref. 20).

Caenorhabditis elegans (worm)

C. elegans, the simplest multicellular organism with a well-understood developmental biology, was introduced as a model by Sydney Brenner in the 1960s (Ref. 21). Worms replicate clonally, are easy to grow, and possess a simple nervous system and other mechanisms that make them readily available for genetic studies. All common cell types found in vertebrates are also found in *C. elegans*[22]. The complete sequence of the *C. elegans* genome (97 Mb, 19,730 genes located on 6 chromosomes; see Table 1) was finalized in early 1999 (Ref. 23). In total, 19,730 potential gene products have been identified and 50% of them have unknown functions[23]. As shown in Table 1, comparative genomic studies indicate that the *C. elegans* genome contains three times more genes than yeast, five times more genes than *E. coli*, and approx. half of the number genes in humans. Forty percent of the *C. elegans* proteins are homologous to proteins in other organisms. Around 20% of all worm open reading frames (ORF) are present in yeast, whereas 40% of yeast ORF have significant orthologs in the worm[24]. These indicate core functions necessary for life such as metabolism, protein folding or trafficking, and protein degradation. Unique worm sequences relate to special functions of multicellular organisms such as signal transduction and regulatory control (e.g. adhesion molecules, growth factors). The vast majority of molecular, structural and cellular signaling pathways are very similar in man and worm[25]. The exploding field of apoptosis has been driven by initial observations made in *C. elegans* (Ref. 26). The essential components of exocytosis are also conserved between mammals and nematodes.

Table 1. Comparative genomic analysis of key model organisms

	Genome size (Mb)	Gene number	Haploid chromosome number
Bacterium (*Escherichia coli*)	~4	4,403	1
Yeast (*Saccharomyces cerevisiae*)	~12	6,190	16
Worm (*Caenorhabditis elegans*)	97	19,730	6
Fruit Fly (*Drosophila melanogaster*)	120	13,601	4
Mouse (*Mus Musculus*)	3,454	~50,000 (estimated)	20
Human (*Homo sapiens*)	2,910	33,609	23

Current technologies available in this model include: forward genetics (finding genes based on phenotypic observations), reverse genetics (finding genes without knowledge of function), antisense technologies, RNA interference (RNAi; the process whereby the introduction of double stranded RNA into a cell inhibits gene expression in a sequence-dependent fashion)[27], transgenic and knockout animals, and finally the use of the green fluorescent protein (GFP) for protein and cell localizations[28]. It is anticipated that this model organism will yield major breakthroughs in biological processes[29]. Large-scale, genome-wide production of null alleles from all known *C. elegans* genes has been initiated by a worldwide consortium (The *C. elegans* Gene Knockout Consortium; http://elegans.bcgsc.bc.ca/knockout.shtml), with the task to generate the 19,000 null alleles within five years, utilizing transposon mutagenesis and RNAi technologies. Already more than one-third of the genes of *C. elegans* have been analyzed for phenotypes following RNAi (Ref. 30). Several congenital human diseases, such as spinal muscular atrophy, polycystic kidney disease, muscular dystrophy, and Alzheimer's disease have been studied using *C. elegans* as a genetic model[31]. There is therefore already strong evidence that research using this organism will continue to enhance our understanding of the function of several human disease genes. An example of a human gene function reconstituted in *C. elegans* is described below.

Mutations in the human presenilin gene *PS1* cause the most frequent and aggressive forms of familial Alzheimer's disease. Mutations in the *C. elegans* presenilin genes *sel-12* and *hop-1* result in a defect in thermotaxis memory of the animals. Thermotaxis is an experience-dependant behavior that involves the pairing of food with the temperature of growth. The defect is caused by a loss of presenilin function in two cholinergic interneurons and causes severe morphological alterations of the neurites. The defect can be rescued by expressing the wild-type human presenilin *PS1* but not by the mutated human presenilin *PS1* (A246E), indicating evolutionary conserved control of neural morphology and function by presenilins. PS1 transgenic worms are now used to understand the molecular details leading to these defects[32].

Several biotechnology companies such as DeVGen (http://www.devgen.com/), Elegen (http://www.elegene.com/), and Exelixis (http://www.exelixis.com/) use worms for the identification and validation of targets ranging from high-throughput phenotypic analyses, to screening of chemical libraries. Messenger RNA expression profiling will soon be available to understand molecular responses during and after specific drug treatments (drug profiling, drug mode-of-action). Disadvantages of using the worm as a model are lack of an immune system, and the difficulties involved in recording electrophysiological properties from most cells.

Drosophila melanogaster (fly)

More complex than the nematode, *D. melanogaster* replicates easily and is a multicelluar organism with a complex nervous system. A variety of genetic tools are available with this organism and it possesses the same cell types and biology as vertebrates. Its 120 Mb euchromatic DNA genome located on 4 chromosomes (Table 1) was completely sequenced in 2000 (Ref. 33). It contains 13,601 genes (twice as many as yeast), and somewhat smaller number than *C. elegans*, but with a comparable functional diversity[33]. Genomic comparisons indicate that 77% of the 929 human disease genes involved in neurological, endocrine, cardiovascular, ophthalmic, immunological, hematological and cancer-related disorders, as well as storage or metabolic defects, have a strong homology with *D. melanogaster* (http://homophila.sdsc.edu/).

Transgenic flies can be easily produced, and forward and reverse genetic tools are available, as are RNAi and transposons, to study the biology of the eye or body development, memory, lifespan and aging. Manipulation of the *Drosophila* genome in ways analogous to the mouse or yeast is now possible[34]. A suppressor screen for polyglutamine toxicity associated with Huntington's disease has been performed in *Drosophila*, and two suppressor genes were found as described below.

dHDJ1 is homologous to the human heat shock protein 40/*HDJ1*, and *dTPR2* is homologous to the human tetratricopeptide repeat protein 2 (Ref. 35). Each of these molecules contains a chaperone-related J domain, and suppression of polyglutamine toxicity was verified in transgenic flies, clearly validating these genes as potential drug targets for the suppression of polyglutamine repeat diseases. Disadvantages of using the fly as a model is the relatively tedious process of handling and storing the flies, and the low-throughput for target- or pathway-identification.

Mus musculus (mouse)

Rodents (rats and mice) are widely used as organisms to study basic biology, development, genetics, signaling pathways, drug responses and metabolism, and are especially useful as a model for humans because of the repertoire of complex behaviors that is available. Mice are easy to breed, manipulate and handle, and can be genetically engineered. Transgenesis (the introduction of new DNA sequences into the germ line, resulting in the production of transgenic animals) and gene targeting (the integration by homologous recombination of new DNA sequences into the genome of an organism at sites where its expression can be controlled) technologies are available for the mouse owing to pluripotent embryonic stem cells[36] and germline manipulations[37]. Typically with knockout technology, the gene of interest is no longer functional (null mutant), whereas with knock-in animals, a reporter gene (i.e. *LacZ*) is inserted in place of the gene of interest and its developmental regulation can be monitored. By contrast, with the conditional Cre/loxP recombination system, the excision event to generate a null mutant can be triggered at any time during or after development to avoid premature lethality. Genome-wide approaches to generate knockouts have been initiated (http://www.lexgen.com/omnibank/omnibank.htm), and conditional gene targeting is becoming very popular (see Ref. 38). Reverse and forward genetic studies using chemical mutagens such as ethylnitrourea (ENU) and ethyl-methane sulfonate (EMS)[39], make the mouse one of the most promising animal models[40] for human genetic studies, especially because of the extensive set of natural mouse mutants available.

There is an increasing set of defined DNA markers for the mouse[40]. Linkage, association and polymorphism studies and positional cloning can be carried out on rodents as they exhibit well-defined behaviors such as fear, learning, memory, stress, depression, and aggression[41]. Although the cause of large behavior variations between natural mouse strains or interstrain variability in inbred mouse strains is not obvious, it is probable that rodents will be used extensively in both basic and applied pharmaceutical research. Human–mouse homology mapping is already possible to find syntenic regions (or genome collinearity) on different chromosomes. Complex human diseases have been successfully modeled in mice, ranging from neurodegenerative disorders such as amyotrophic lateral sclerosis (SOD1 overexpression), Alzheimer's disease (APP overexpression), to prion disease (mutated PrP) glutamine expansion disorders [Huntington (huntingtin), spinocerebellar ataxia (SCA1)], spinal muscular atrophy, epilepsy, and osteoporosis. The disadvantages of using the mouse as a model for the study of human diseases are the slow-throughput necessary to generate knockout mice, and the amount of time required for backcrosses of the mutant strain.

Homo sapiens (human)

Obviously, humans are the target organism for drug discovery technologies. There is large heterogeneity in the behavior of patients or between different populations owing to gene polymorphisms, redundant targets, differently regulated pathways and environment challenges. The human genome is composed of 2,910 Mb on 23 chromosomes (Table 1) containing only ~33,609 genes, barely twice as many genes as in *C. elegans* or *Drosophila*[42]. The total number of gene products is, however, predicted to be significantly larger (>1,000,000 proteins) owing to transcriptional and post-translational modifications, such as RNA editing, alternative splicing events, and protein modifications such as phosphorylation, or glycosylation. Messenger RNA and protein expression profiling from clinical tissue specimens have provided a lot of new data and potential applications for diagnostics, disease prevention[43], and fast, individualized

pathological diagnosis. It is probable that individualized DNA microarrays will be used for diagnosis, drug responses, preferred treatment responses and small-scale, proof-of-concept clinical trials. Significant advances in genetic approaches such as positional cloning, association studies, and SNP mapping have been recently achieved. Other genetic approaches (e.g. affected sibling pairs, inbred populations) are being used in combination with genomic DNA approaches to speed up the identification of gene products causing (or linked to) a specific human disease. Recently, the locations of genes contributing to familial essential tremor, multiple sclerosis, narcolepsy, Alzheimer's disease, schizophrenia, stroke, osteoporosis and osteoarthritis have been identified (http://www.decode.com/resources/diseases/). Non-invasive functional imaging technologies (NMR, fMRI) are also likely to contribute to the identification of gene products[44].

Conclusions

At present, mice are the favored models for therapeutic areas such as oncology, diabetes, autoimmunity, inflammation, neurodegenerative and infectious diseases. A recent quantitative trait locus (QTL) analysis of hypertension has been performed using two mouse strains (The Jackson Laboratory; http://www.jax.org/research/documents/research_areas/pdf/paigen.pdf). Eight different loci linked to the disease were discovered within one year, and six of these regions have also being identified in a large, five-year, human study costing US$30 million[45]. The mouse approach was 100 times cheaper, five times faster, and gave the same loci. Fine mapping and positional cloning are currently in progress. However, only a small number of genes can be studied simultaneously with mouse models. Simpler animal models have also contributed to the identification of drug targets. These organisms can be used to complement information from mouse models, as they possess a greater potential for higher-throughput and parallel processing. A complementary approach using a mouse model and *C. elegans*, *Drosophila*, zebrafish, or yeast models will therefore enable the rapid identification and validation of a gene product among several hundred potential targets. None of the existing model organisms is ideal. Limitations such as simplicity, time, cost, speed, low-throughput and lack of parallel processing capability represent major issues, which are unlikely to disappear in the near future.

Use of animal models can reduce the amount of time and money spent on understanding the molecular basis of human disease. Overall, animal models represent an unavoidable bottleneck in pre-clinical drug discovery research. However, establishments of consortiums, maintenance of transgenic animal collections, widespread QTL/ENU phenotypic analyses, large-scale behavioral screens, combined with complementary approaches using simpler animal models could somewhat help to alleviate this problem.

References

1 Drews, I.I. (2000) Drug discovery today – and tomorrow. *Drug Discov.Today* 5, 2–4

2 Blattner, F.R. *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1474

3 Josephy, P.D. *et al.* (1997) Recent advances in the construction of bacterial genotoxicity assays. *Mutat. Res.* 386, 1–23

4 Saunders, N.J. and Moxon, E.R. (1998) Implications of sequencing bacterial genomes for pathogenesis and vaccine development. *Curr. Opin. Biotechnol.* 9, 618–623

5 Keasling, J.D. (1999) Gene-expression tools for the metabolic engineering of bacteria. *Trends Biotechnol.* 17, 452–460

6 Link, A.J. *et al.* (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J. Bacteriol.* 179, 6228–6237

7 Hu, J.C. *et al.* (2000) *Escherichia coli* one- and two-hybrid systems for the analysis and identification of protein–protein interactions. *Methods* 20, 80–94

8 Saier, M.H., Jr (1996) Phylogenetic approaches to the identification and characterization of protein families and superfamilies. *Microb. Comp. Genomics* 1, 129–150

9 Tang, C.M. *et al.* (1998) Microbial genome sequencing and pathogenesis. *Curr. Opin. Microbiol.* 1, 12–16

10 Harfe, B.D. and Jinks-Robertson, S. (2000) DNA mismatch repair and genetic instability. *Annu. Rev. Genet.* 34, 359–399

11 Dujon, B. (1996) The yeast genome project: what did we learn? *Trends Genet.* 12, 263–270

12 Goffeau, A. *et al.* (1996) Life with 6000 genes. *Science* 274, 546, 563–567

13 Bassett, D.E., Jr. *et al.* (1996) Yeast genes and human disease. *Nature* 379, 589–590

14 Wodicka, L. *et al.* (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 15, 1359–1367

15 Ross-Macdonald, P. *et al.* (1999) Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* 402, 413–418

16 Schwikowski, B. *et al.* (2000) A network of protein–protein interactions in yeast. *Nat. Biotechnol.* 18, 1257–1261

17 Fields, S. and Song, O. (1989) A novel genetic system to detect protein–protein interactions. *Nature* 340, 245–246

18 Munder, T. and Hinnen, A. (1999) Yeast cells as tools for target-oriented screening. *Appl. Microbiol. Biotechnol.* 52, 311–320

19 Fritz, E. *et al.* (2000) The yeast TEL1 gene partially substitutes for human ATM in suppressing hyperrecombination, radiation-induced apoptosis and telomere shortening in A-T cells. *Mol. Biol. Cell* 11, 2605–2616

20 Szczypka, M.S. *et al.* (1994) A yeast metal resistance protein similar to human cystic fibrosis transmembrane conductance regulator (CFTR) and multidrug resistance-associated protein. *J. Biol. Chem.* 269, 22853–22857

21 Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94

22 White, J.G. *et al.* (1988) *The nematode Caenorhabditis elegans* (Wood, W.B., ed.), Cold Spring Harbor Laboratory

23 The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018 [published errata appeared in *Science* (1999) 283, 35; *Science* (1999) 283, 2103; and *Science* (1999) 285, 1493]

24 Chervitz, S.A. *et al.* (1998) Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* 282, 2022–2028

25 Bargmann, C.I. (1998) Neurobiology of the *Caenorhabditis elegans* genome. *Science* 282, 2028–2033

26 Metzstein, M.M. *et al.* (1998) Genetics of programmed cell death in *C. elegans*: past, present and future. *Trends Genet.* 14, 410–416

27 Ketting, R.F. and Plasterk, R.H. (2000) A genetic link between co-suppression and RNA interference in *C. elegans*. *Nature* 404, 296–298

28 Chalfie, M. *et al.* (1994) Green fluorescent protein as a marker for gene expression. *Science* 263, 802–805

29 Hodgkin, J. and Herman, R.K. (1998) Changing styles in *C. elegans* genetics. *Trends Genet.* 14, 352–357

30 Kim, S.K. (2001) Functional genomics: the worm scores a knockout. *Curr. Biol.* 11, R85–R87

31 Culetto, E. and Sattelle, D.B. (2000) A role for *Caenorhabditis elegans* in understanding the function and interactions of human disease genes. *Hum. Mol. Genet.* 9, 869–877

32 Wittenburg, N. *et al.* (2000) Presenilin is required for proper morphology and function of neurons in *C. elegans*. *Nature* 406, 306–309

33 Adams, M.D. *et al.* (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195

34 Rong, Y.S. and Golic, K.G. (2000) Gene targeting by homologous recombination in *Drosophila*. *Science* 288, 2013–2018

35 Kazemi-Esfarjani, P. and Benzer, S. (2000) Genetic suppression of polyglutamine toxicity in *Drosophila*. *Science* 287, 1837–1840

36 Muller, U. (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mech. Dev.* 82, 3–21

37 Moreadith, R.W. and Radford, N.B. (1997) Gene targeting in embryonic stem cells: the new physiology and metabolism. *J. Mol. Med.* 75, 208–216

38 Takeda, J. *et al.* (2000) Conditional gene targeting and its application in the skin. *J. Dermatol. Sci.* 23, 147–154

39 Hrabe de Angelis, M. and Balling, R. (1998) Large scale ENU screens in the mouse: genetics meets genomics. *Mutat. Res.* 400, 25–32

40 Blake, J.A. *et al.* (2000) The Mouse Genome Database (MGD): expanding genetic and genomic resources for the laboratory mouse. *Nucleic Acids Res.* 28, 108–111

41 Owen, M.J. and Craddock, N. (1996) Modern molecular genetic approaches to complex traits: implications for psychiatric disorders. *Mol. Psychiatry* 1, 21–26

42 International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921

43 Rudert, F. (2000) Genomics and proteomics tools for the clinic. *Curr. Opin. Mol. Ther.* 2, 633–642

44 Rudin, M. *et al.* (1999) *In vivo* magnetic resonance methods in pharmaceutical research: current status and perspectives. *NMR Biomed.* 12, 69–97

45 Atwood, L.D. *et al.* (2001) Genome-wide linkage analysis of blood pressure in Mexican Americans. *Genet. Epidemiol.* 20, 373–382