

# nature cell biology

## In search of a function

Improved gene  
inactivation  
technologies now  
provide better tools with  
which to address cell  
biological issues in a  
dynamic *in vivo* context,  
rather than in cell  
cultures alone

**B**efore the advent of reverse genetics and genomics, cloning a gene or a complementary DNA was often the most difficult and time-consuming part of a project. This is no longer the rate-limiting step, however, now that a wealth of sequence information from major genome-sequencing projects is being made available. The genomes of two of the principal model systems used in biology, *Caenorhabditis elegans* and *Drosophila melanogaster*, are now fully sequenced; the human genome will be published shortly (see *Nature* 405, 983–984 and 984–985; 2000); and sequencing of those of the mouse, of *Arabidopsis thaliana* and of *Schizosaccharomyces pombe* are well underway. Modern molecular biology is therefore increasingly concerned with ascribing physiological functions to gene products.

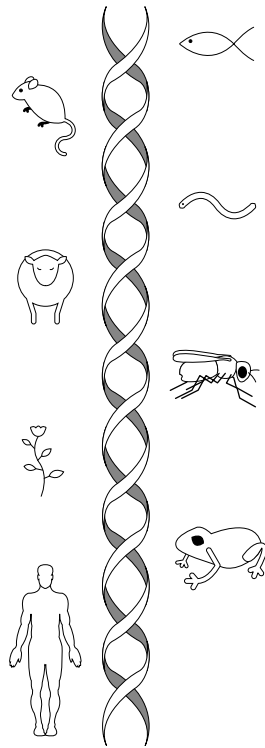
There are essentially two ways to tackle the problem — to overexpress the relevant gene or cDNA, or to inactivate it. Each strategy has its limitations; when expressed at higher-than-physiological levels, proteins can have widespread and non-specific effects, whereas the loss of one gene product can sometimes be compensated for by another gene product, causing an ‘artefactual’ absence of phenotype. Consequently, rigorous gene–function analyses now use a combination of these two techniques, along with the generation of site-directed mutations (analysed *in vivo* either by transfection or in the context of the endogenous gene locus by ‘knock-in’ technology) and the development of conditional (cell-type- or differentiation-stage-specific) knockout animals to avoid problems such as early lethality of a null mutation.

Different strategies have been developed to specifically inactivate genes, each with their own advantages and drawbacks in terms of speed, specificity and cost, and with their limitations in terms of applicability to a given gene or organism. Speed and cost are especially relevant issues in the current climate of post-genomics research, as thousands of new genes need to be screened for function.

Gene targeting by homologous recombination involves replacing one, or in the case of diploid species possibly both, allele(s) of a gene at the endogenous locus with a mutant allele. This technology is now used effectively in haploid eukaryotes such as yeast and *Dictyostelium discoideum* (in which targeting of one allele is sufficient to achieve complete gene inactivation). It is also routinely used in mice, although it is quite costly and time-consuming, and it has recently been reported to work in sheep (*Nature* 405, 1066–1069; 2000).

Until recently, similar strategies had proven unsuccessful in the worm, fly, zebrafish and frog. However, Rong and Golic (*Science* 288, 2013–2018; 2000) have now found a way to adapt homologous-recombination-mediated gene targeting for use in flies. In flies, ‘random’ mutagenesis involving insertion of transposable elements into genes has a venerable history and would, if developed on a large scale, be expected to target most genes. In the new method, directed targeting is achieved by successive crossing of flies that have been engineered to express a site-specific endonuclease, a site-specific recombinase, and a construct bearing a region of the locus to be targeted and target sites for both the nuclease and the recombinase. A linear targeting vector, which is more efficient at homologous recombination than circular vector DNA, is thus generated in germ cells *in vivo*.

Although this method and the efficiency of homologous recombination (1 in 500 in the case of the *yellow* locus used by Rong and Golic) need to be confirmed with other genes at other loci, it is undoubtedly a welcome advance. According to Kristin White of Massachusetts General Hospital, “this technique should be incredibly useful. The method gives [people working with flies] the opportunity to make rapid use of the information that is available in



the genome sequence". The technique has limitations, however. The integration of the linear vector at the targeted gene locus generates a duplication, with one mutant allele being appended to the original wild-type allele. Thus, whereas the gene-targeting strategies most commonly used in mice, for example, result in replacement of the wild-type locus, the technique for flies is not appropriate if the aim is to completely silence the endogenous locus and/or express site-directed mutants in the absence of the wild-type gene product (so-called 'knock-in' mutations). Clearly, more work is required in flies to allow targeting by deletion of the endogenous locus and replacement with a mutant allele. Also, strategies to screen for successful recombination events will have to be developed. Rong and Golic used genetic linkage screens with known visible markers, which could not be used in cases when the expected phenotype of the gene being targeted is not known, as is the case in general genomic screens. Instead, other molecular biology techniques such as PCR, Southern blotting and/or introduction of selectable markers into the targeted allele may have to be used.

### RNA interference as an alternative

An alternative, less time-consuming and cheaper technique to study loss of gene function is RNA interference (RNAi). This was initially carried out by injecting antisense RNA, but further studies revealed that sense RNA and double-stranded RNA (dsRNA) are also effective, and dsRNA is now most commonly used. The injected dsRNA interferes with gene expression by an unknown mechanism, possibly involving specific degradation or translation inhibition of the corresponding mRNA (see *Nature Cell Biol.* 2, E31–E36; 2000 for review). RNAi was initially developed in *C. elegans*, where it holds great sway, as traditional chemical mutagenesis procedures only cover 8% of the genes. With the sequence information now available, systematic targeting of all *C. elegans* genes is underway. It has now been optimized in this organism (amazingly, injection of the worm can be bypassed by simply feeding the worms bacteria engineered to express the relevant dsRNA fragment (*Nature* 395, 854; 1998)) and expanded to other species such as trypanosomes, flies and plants. There have also been reports of effective RNAi in vertebrates, such as the mouse (*Nature Cell Biol.* 2, 70–75; 2000), frog (*Nature* 405, 757–763; 2000) and zebrafish (*Biochem. Biophys. Res. Commun.* 263, 153–161, 1999; *Dev. Biol.* 217, 394–405, 2000). Other laboratories, however, have failed to obtain effective, specific gene inactivation in these species, and it is unclear whether these discrepancies are the result of differences in the stage of development of the embryos used or in the particular RNA sequences that were injected. Many laboratories working with vertebrates are returning to the original principles of RNAi, using antisense single-stranded RNA or oligonucleotide fragments of antisense RNA. The new trick, however, is to incorporate modified bases (phosphorothioates or morpholinos) into the oligonucleotides to increase their stability, an approach that seems to be effective in both the frog (*Dev Biol.* 222, 124–134; 2000) and the zebrafish.

Another strategy currently in the pipeline combines the transgenic and RNAi approaches, using vector DNAs encoding palindromic 'foldback' sequences to express dsRNA. This method allows tissue- and stage-specific gene targeting and is expected to be much less laborious than traditional gene inactivation by homologous recombination. However, it does not result in complete gene inactivation, as the level of the mRNA is thought to be reduced only three-to-fivefold. Depending on the gene, this modest change may be either a drawback (if insufficient to affect the function of the gene product) or an advantage (for dosage-sensitive genes, in cases in which null alleles are lethal, or for screening for suppressors and/or enhancers of a phenotype in the same pathway).

The availability of genomic sequence information offers several alternatives to the traditional 'mutagenesis and genetic screen' approach to identifying the principal regulators of a given biological pathway. Scientists are increasingly using gene arrays to simultaneously monitor the expression profiles of thousands of genes. With these data, they can identify candidate genes before devising gene-inactivation strategies.

Last, but not least, there is chemical genetics, a powerful tool with which to study protein function. On one hand, screens can be carried out using libraries of small synthetic or natural compounds to generate a given phenotype or disrupt a particular cellular process, allowing the subsequent identification of the relevant target. This approach is particularly relevant to fields such as membrane traffic, as few chemical inhibitors are known that can be used to disrupt and study critical transport pathways. On the other hand, silent mutations can be designed — often on the basis of structural information about the target protein — that render the mutant, but not the wild-type, allele of the protein in question sensitive to inhibition by small, highly specific, cell-permeable chemical compounds. The expression of such an analogue-sensitive allele and the conditional and reversible usage of an allele-specific inhibitor opens new avenues to analyse protein function in the context of a living cell or organism (*Curr. Biol.* 8, 257–266; 1998).

Gene inactivation is only the first step in understanding the function of a gene. The next step involves analysing the cellular processes that are affected in mutant animals, and the biochemical bases for its phenotype, at the levels of both organism and cell. Improved gene-inactivation technologies now provide better tools with which to address cell-biological questions in a dynamic *in vivo* context.