Molecular biology

Manipulation just off target

R. Michael Liskay

PROGRESS has been made in several laboratories towards gene targeting — recombinational interactions between an introduced DNA molecule and its homologous sequence residing in a chromosome — in mammalian cell systems[1-4]. The targeting studies of Thomas and Capecchi reported on page 34 of this issue[5] are provocative in several respects. They indicate that DNA introduced into the nucleus by microinjection can induce mutations at high frequency in the target gene in a totally unexpected manner. The observations both sound a note of caution for some applications of gene targeting and provide new methodologies that could facilitate the efficient introduction of mutations into a particular region of a gene.

High-efficiency targeting is a potentially powerful approach for manipulating the genomes of higher organisms in a manner analogous to the molecular gymnastics that are possible in yeast, including the creation or removal of mutations in any chromosomal sequence for which one possesses a molecular clone.

Thomas and Capecchi derived recipient mouse cell lines containing a mutated *neo* gene whose normal activity confers resistance to the toxic compound G418. This particular mutant form of the gene, here called the resident gene, is an 'amber' mutation that causes premature termination of translation and a resultant truncated protein. The amber mutation (a single base-pair change) is located in the initial portion of the coding region of *neo* and its presence can easily be detected by molecular techniques. Thomas and Capecchi attempted to correct the resident *neo* gene by microinjecting a different mutant form of this gene, a deletion that removes the last 52 amino acids from the protein, directly into the nucleus of the recipient cell. At a frequency of approximately 1 per 1,000 injections they recovered cells resistant to G418. Roughly half of these lines now contained, as expected, an apparently normal *neo* gene — that is, the amber mutation had been removed. Surprisingly, in the rest of the $G418^r$ cells the amber mutation was maintained. Eight of these exceptional, corrected genes, each representing an independent occurrence, were cloned into bacteria and their nucleotide sequences near the amber mutation were determined. Each gene now harboured a compensating mutation. Three different insertion mutations were found: insertion of G, of T or of GGCT. These insertions now allow a normally silent and out-of-phase protein synthesis start signal to be used. A functional protein is produced that is incorrect for the first few amino acids but normal for the rest. Furthermore, the mutations occurred within the first of two directly repeated sequences (GGCTAT). That each insertion results in a tandem duplication of at least a portion of the repeat sequence leads the authors to conclude that direct repeats are a prerequisite for this phenomenon.

Control experiments, including microinjection of either the same amber mutant of *neo* or an unrelated DNA molecule, did not yield any $G418^r$ cells. From these control injections the authors conclude that the process requires both homology and at least one nucleotide difference, or mismatch, between the interacting genes. Because of this apparent requirement for a mismatch, Thomas and Capecchi invoke the formation of heteroduplex DNA between the resident gene and the incoming plasmid molecule as an intermediate and therefore refer to this newly observed process as heteroduplex-induced mutagenesis. They point out that because restoration of a functional gene was intrinsic to their experiments, many (even most) of the mutations induced by these interactions would be undetectable, making the actual frequency of mutations much higher than the 1 in 1,000 seen.

Thomas and Capecchi also analyse how the compensating mutations result in functional protein. These experiments provide a glimpse at how the ribosome can apparently reinititate protein-chain elongation after encountering a single stop signal, provided there exists another start signal in the immediate vicinity.

The main issue, however, is the impact that these findings have on various applications of and strategies for gene targeting in mammalian cells. One might first ask why others studying targeting in mammalian cells have not observed error-prone events. The answers seem to be relatively simple. In one recent study[2] direct selection for targeted integration was not applied and there was no reason to examine the targeted gene at the nucleotide level. In other experiments involving direct selection in a manner similar to that of Thomas and Capecchi, the nature and/or positions of the mutation in the resident sequence were such that the detection of error-prone events would often be precluded[1,3,4]. For example, extensive terminal or internal deletions of the resident gene would not be prone to correction by the proposed mutagenic mechanism. Note that in the bacterium *Salmonella* there exists a possibly related phenomenon termed selfing whereby one mutation can be rescued by the very same mutant DNA[6]. Unfortunately, these events have not to my knowledge been analysed at the molecular level. On the other hand, gene conversion, a form of nonreciprocal recombination, has been shown to act with fidelity in the several cases examined in fungi[7-9].

There are two general implications of heteroduplex-induced mutagenesis. First, the results provide a caution to those who wish to manipulate mammalian genomes by gene targeting. It is possible that unpredicatable and unwanted mutations could be introduced into a gene intended to be the target for correction. For gene therapy in humans these potential side effects should be considered serious. Second, and more positive, Thomas and Capecchi's observations suggest new methods for directly mutating mammalian chromosomal sequences that differ from the approaches that are standard fare in the yeast *Saccharomyces cerevisiae*. In this yeast, specific mutations are engineered into a cloned sequence one-by-one and used to replace one-by-one the normal sequence by a targeted interaction. Based on heteroduplex-induced mutagenesis, a single appropriately constructed plasmid might prove sufficient to induce a library of mutations in a particular region of the gene. This approach might be used to advantage for mutating the control region (for example, promoter) of genes to define the location and function of *cis*-acting regulatory sequences. Such regions frequently contain repeated sequence elements and therefore should be prime targets for the proposed mutagenesis by incoming DNA.

One obstacle preventing the widespread use of targeting procedures is the propensity of mammalian cells to incorporate DNA into the genome in an essentially random fashion. The efficiency of homologous relative to nonhomologous interactions is presently being achieved at a ratio of about 1 to 100. This ratio must be improved before the mammalian genome becomes as amenable to genetic tinkering as its smaller cousin, the yeast. □

1. Smith, A.J.H. & Berg, P. *Cold Spring Harb. Symp. quant. Biol.* **49**, 171-181 (1984).
2. Smithies, O. *et al. Nature* **317**, 230-234 (1985).
3. Lin, F.L., Spearle, K. & Sternberg, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1391-1395 (1985).
4. Thomas K.R., Folger, K.R. & Capecchi, M.R. *Cell* **44**, 419-428 (1986).
5. Thomas, K.R. & Capecchi, M.R. *Nature* **324**, 34-38 (1986).
6. Demerec, M. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1696-1704 (1962).
7. Roman, H. *Cold Spring Harbor Symp. quant. Biol.* **21**, 175-185 (1956).
8. Case, M. & Giles, N. *Genetics* **49**, 529-540 (1964).
9. Fogel, S. & Mortimer, R. *Molec. gen. Genet.* **109**, 177-185 (1970).

R. Michael Liskay is in the Departments of Therapeutic Radiology and Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510-8040, USA.