20 September 1985, Volume 229, Number 4719 SCIENCE

Strategies and Applications of in Vitro Mutagenesis

David Botstein and David Shortle

Summary. The many mutagenesis strategies made possible by the availability of cloned genes have been outlined in the context of how each strategy lends itself to the answering of different kinds of biological questions.

Biochemists have begun to use mutations to probe the relationship between the structure and activity of proteins; cell biologists are using mutations to define the roles of particular proteins and protein assemblies in the cell; developmental biologists are using mutations to determine the logic and order of molecular events during differentiation and morphogenesis; and neurobiologists are beginning to turn to mutations to try to understand the way in which neural networks are formed and, eventually, how they function. These new and expanding applications of mutations in many disciplines of biology represent one of the most important consequences of the revolution in the life sciences that has resulted from the development of recombinant DNA technology. The objective of this review article is to sketch the outlines of the many mutagenesis strategies made possible by the availability of cloned genes. Our emphasis on general principles and applications has required that we gloss over many ideas and technical accomplishments in the field; such information, though, has been reviewed (*1*, *2*).

Before we address the question of how best to isolate or construct mutations, it is important to review the fundamental logic behind the use of gene mutations to analyze biological phenomena. The primary reason for isolating and characterizing a mutation is to assess its consequences, or in genetic terminology, its phenotype. In the ideal mutation experiment, two organisms are submitted to careful comparison, one being mutated at a single known site and the other lacking the mutation. Any observed difference between the two is then attributable to the mutation; by characterization of this difference, inferences can be made about the function of the corresponding wild-type gene, regulatory signal, or nucleotide pair. In other words, the properties of a normal function can be learned from the consequences of perturbing or eliminating a single gene or genetic element.

One approach for designing the ideal mutation experiment is to begin with a phenotype of interest (such as failure to complete mitosis at a given temperature) and search, after classical mutagenesis, for mutant organisms that exhibit this phenotype. The mutations obtained by this strategy would then be characterized by genetic mapping, followed by cloning and sequencing the wild-type and mutant genes. A second approach is to begin with the idea that a particular protein (such as tubulin) is likely to play an important role in mitosis. To establish this point genetically, it would be necessary to clone the gene, induce mutations in the cloned copy in a recombinant DNA molecule, and return these mutant genes to the organism in a way that would allow assessment of phenotype.

Both these approaches might lead to the same experiment, although with unequal likelihood. The first approach, in which the phenotype determines the selection of mutations for analysis, has the advantages that no detailed hypotheses are required to obtain material for study and that each new mutation has a high probability of contributing to an understanding of the phenomena underlying the phenotype. The major advantage of the second approach, in which the gene for a particular protein is cloned initially, is that very specific hypotheses, often arrived at by earlier biochemical or biological studies, can be put to definitive genetic tests.

Although, in principle, every gene that encodes information relevant to a given biological process should be found by the first approach if the "correct" phenotype has been used to search for mutations, the results, in practice, are different. For example, extensive screening for cell-division–cycle mutants in yeast failed to produce mutations in the β-tubulin gene, despite the fact that mutations in that gene isolated by the second approach (*3*) exhibited precisely the same cell-division–cycle phenotype used in the unsuccessful screening. Perhaps the most important feature of in vitro mutagenesis is the ability to efficiently and predictably introduce mutations into a gene of interest.

Classical in Vivo Mutagenesis

The only mutations available for study by the earliest geneticists were natural variants and occasional spontaneous mutations. The discoveries that organisms exposed to x-rays (*4*) and certain chemical compounds (*5*) yield much higher frequencies of mutant progeny led to a revolution in experimental genetics. Because these agents gave geneticists partial control over the process of mutagenesis, they could systematically study biological phenomena by collecting large numbers of different mutations display-

David Botstein is a professor in the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. David Shortle is an assistant professor in the Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205.

ing a characteristic phenotype, which could then be classified on the basis of complementation behavior, map position, and various other criteria. In this way the genetic loci relevant to the phenomena could be identified, counted, positioned on a genetic map, and then functional interactions between loci defined. With the classical studies of eye pigmentation mutations in *Drosophila melanogaster* (*6*) and of auxotrophic mutations in *Neurospora crassa* (*7*), it became apparent that the detailed phenotypic characterization of mutations could be extended to the level of individual molecules.

The power of classical in vivo mutagenesis to provide material for genetic and biochemical analysis is attested to by the progress made over the past few decades in the genetics of eukaryotes such as *D. melanogaster* and the molecular genetics of *Escherichia coli* and other microbes. Nevertheless, the requirement for a specific phenotype in order to identify rare mutations in a mixture of wild-type and irrelevant mutations imposed serious limits on the range of phenomena to which genetics could be applied. It was not feasible to study more than a few mutations in a single gene without the special circumstances of a unique phenotype that is easy to score plus an efficient genetic crossing system for weeding out secondary mutations in irrelevant genes caused by the general mutagenesis. Isolation of mutations in a gene of special biochemical interest by simple assay ("brute force," in the jargon of the geneticist) required not only heroic amounts of labor to screen thousands of mutagenized organisms but also major (even lucky) assumptions about the phenotype. A good example is the isolation of a mutation in DNA polymerase I of *E. coli* (*8*).

In bacteria, some improvements in mutagenesis were afforded by the concept of "localized mutagenesis" (*9*). New mutations could be limited to individual segments of a genome by mutagenizing transducing particles that carry only a small fragment of the bacterial chromosome, followed by generalized transduction with selection for a marker known to be closely linked to the genes targeted for mutagenesis. The development of specialized transducing phages and episomes carrying only a portion of the bacterial genome also allowed the mutagenesis of specific genes without exposure of the entire genome of the host cell to the action of a mutagen. Despite having some of the advantages of the new methods of in vitro mutagenesis, mutagenesis of these naturally occurring recombinants had only limited applicability because many genes of interest in *E. coli* could not be readily isolated on specialized transducing phage or small F′ episomes.

Transposon Mutagenesis

A second revolutionary development in mutagenesis came with the realization that insertion mutations could be induced in virtually any gene of interest by the appropriate manipulation of naturally occurring transposable elements. Although these mobile segments of DNA had been discovered many years earlier in maize and *Drosophila*, it was in bacteria that the advantages of insertion mutagenesis were first systematically exploited (*10*).

Unlike chemical or radiation mutagenesis, transposon insertion mutagenesis results in a single, unique physical alteration in the gene that has been mutated. Mutagenesis can frequently be carried out in such a way as to limit the number of transposons inserted to essentially one per genome. Most importantly, the insertion event can often be selected by means of drug resistance or phage immunity carried by the transposon. Thus, a population of organisms, each of which had a transposon insertion within a gene or intergenic segment, can easily be generated for phenotypic screening. Mutagenesis is therefore extremely efficient, the level of secondary mutations is very low, and most mutations lead to total inactivation of the gene. If the transposon inserts randomly, the probability of finding a mutation in a given gene is the fraction of the total genome that the gene of interest occupies. Thus, for bacteria, about 1 of 3000 to 5000 organisms with a single insertion are expected to be mutated in a given gene. Transposon mutagenesis can be applied to very small genomes, such as plasmids, and is a simple alternative to the use of deletion mutations for defining the extent of a gene of interest (or a small cluster of genes) after it has been cloned [for examples, see (*11*, *12*)].

Probably the single most important advantage of insertion mutations is that they contain an insertion of a known DNA element, the transposon. As a result, transposon mutagenesis can be used to isolate genes by first identifying an insertion mutation in or very near the gene of interest, on the basis of phenotype or genetic linkage, and then cloning a fragment of DNA from the mutant genome that harbors the nucleotide sequences of the transposon. This transposon-tagging technique has made possible the efficient identification and cloning of interesting genes in *Drosophila* (*13*) and also, by means of an RNA tumor virus as the transposon, in mice (*14*).

In Vitro Mutagenesis

Since the introduction of recombinant DNA methodology, genes can be removed from their normal environment in an intact genome and isolated as DNA fragments on cloning vectors. The availability of purified genes in vitro in microgram amounts has dramatically expanded the potential for inducing mutations. In the controlled environment of the test tube, it is now possible to alter, efficiently and systematically, the sequence of nucleotides in a segment of DNA. In the following sections the new methods of in vitro mutagenesis are divided into three broad categories: (i) methods that restructure segments of DNA, (ii) localized random mutagenesis, and (iii) oligonucleotide-directed mutagenesis. This classification emphasizes the practical aspects of each method's application.

The considerable increases in mutagenic efficiency and specificity attainable with the new methods, however, do exact a price. Because these methods are designed for use on isolated DNA molecules, a gene must almost always be removed from its normal genetic context—a unique locus on a large complex chromosome inside a living cell (or virus)—and inserted into the abnormal context of a small cloning vector propagated in *E. coli*. Unlike classical in vivo mutagenesis, in which all mutations are isolated in situ, in vitro mutagenesis invariably yields gene mutations out of their normal context. This is the most radical and most troublesome difference between the classical methods and the powerful in vitro methods.

For some applications, this change in genetic context is relatively unimportant. For other applications, though, inferences about a wild-type gene, on the basis of the phenotype of a mutation construction in vitro, can only be made after the mutant allele has been restored to its normal genetic context. In such situations, the genetic manipulations required to assess the consequences of a mutation in its proper context become the primary challenge to the molecular biologist. Therefore, the last two sections of this review article are devoted to a discussion of the variety of available solutions, some partial and some complete, to this important problem.

Restructuring of DNA Segments

After a gene of interest has been isolated, it is usually necessary to reduce to a minimum the size of the cloned DNA segment carrying the gene and to move it into a small, circular cloning vector. By a form of deletion analysis, extragenic flanking sequences are systematically eliminated from the initially cloned DNA segment and each new deletion is tested, with some assay of structure or function, to determine that the gene is still intact. In this way the ends or boundaries of the gene are roughly defined while, at the same time, smaller subclones are isolated that will simplify subsequent manipulations of the gene, such as DNA sequencing and mutagenesis.

Further reductions in size, which permit very precise definition of the functional boundaries at the 5′ and 3′ ends of the gene, are achieved by generating terminal deletions with one end point located outside of the gene and the other positioned progressively closer to the gene. Collections of deletions, which have been extensively used to identify such regulatory sequences as transcriptional promoters, can be readily constructed by using an exonuclease [either Bal 31 (*15*) or exonuclease III plus S1 nuclease (*16*)] to remove nucleotides starting from a unique site just outside the gene. This site is generated by restriction enzyme cleavage of a circular DNA molecule; in some cases, nuclease digestion can be confined to one of the two ends of the linear DNA molecule, leading to deletions that extend in a single direction (*17, 18*).

Insertion of a small synthetic oligonucleotide that encodes a unique restriction enzyme cleavage site at the end points of these deletions (*16*) or at positions randomly distributed across the cloned DNA segment (*19*) provides a readily available restriction site for use in further restructuring of the cloned DNA segment. These oligonucleotide insertions, termed linker mutations, are especially versatile for the modification of circular DNA molecules. For instance, cleavage at a linker insertion introduced near or within the gene provides a site at which additional nucleotides can be inserted or deleted. It may also provide ends by which the cloned DNA segment itself can be inserted into other types of DNA vectors for procedures such as nucleotide sequencing, production of large quantities of the gene product, and transformation of other types of cells. In addition, by joining restriction fragments isolated from pairs of linker insertion mutations of the appropriate size and position, mutations consisting of four to eight tightly clustered base substitutions can be generated, making it possible to efficiently "scan" a small region of DNA for regulatory sites (*20*).

When changes in the level of expression of a gene are difficult to monitor because of the lack of a convenient assay for the gene product, a common strategy is to construct a gene fusion in which the regulatory elements of the gene of interest control the expression of a gene product that can be readily quantitated. Typically, a second gene specifying an easily assayable enzyme, such as β-galactosidase (*21*) or chloramphenicol acetyltransferase (*22*), is joined downstream of the 5′ end of the first gene, either by fusing the two protein-encoding regions in frame to create a hybrid protein or by fusing the 5′ untranslated sequences of the two genes so that the hybrid gene encodes only the product of the second gene. This strategy, developed in *E. coli* before the availability of recombinant DNA methods (*23*), has been expanded and applied to the study of gene regulation in a large number of divergent organisms, including yeast (*24, 25*) and a number of higher eukaryotes (*26*).

Random Mutagenesis

Several of the mutagens used to randomly mutagenize organisms or viruses in vivo can also induce random point mutations in purified DNA molecules. Hydroxylamine and nitrous acid have been used to generate mutations in genes cloned into a plasmid or phage vector by simply reacting the entire recombinant DNA molecule with mutagen. Although this technically simple approach can be useful when coupled to a simple phenotypic selection or screen, the problem of a background of secondary mutations must be dealt with, either by reducing mutagenesis to a level where secondary mutations are rare or by using recombination to isolate the primary mutation free of any secondary mutations. Furthermore, as is typical of chemical mutagens in general, hydroxylamine and nitrous acid induce a limited number of base substitutions that, instead of being uniformly distributed across a nucleotide sequence, tend to cluster at a small number of "hot spots" (*27*). Although technically not an in vitro method of mutagenesis, brief passage of purified DNA through bacterial strains carrying various mutator activities can also be used to obtain random mutagenesis (*12, 28, 29*).

The potential problem of secondary mutations arising because of random mutagenesis can be avoided in two ways. First, transposon mutagenesis of a recombinant plasmid or phage produces single-insertion mutations that are often essentially random in distribution. These types of random mutations, which usually result in complete loss of function, can be isolated easily by propagating the cloned gene in an appropriate *E. coli* strain and then retransforming purified DNA with selection for the drug-resistance marker on the transposon (*12, 30*). A second, more general approach applicable to small circular DNA molecules is to introduce a single nick at a random site with deoxyribonuclease I plus ethidium bromide (*31, 32*). After conversion of the nick into a short single-stranded gap, deletions and insertions (*33*) or a variety of types of base-substitution mutations (as described in the next section) can be efficiently targeted to the sequence of nucleotides exposed in the single-stranded gap.

Localized Random Mutagenesis

In many applications involving cloned genes, random mutagenesis of the purified gene is still not a feasible approach for mutant isolation, usually because the assay for phenotype is too laborious to apply to the hundreds, if not thousands, of candidates. In response to the need for more direct routes to mutant isolation, in vitro methods have been developed that result in random mutations confined to small segments of large DNA molecules. To understand the principles behind these methods and to compare their relative advantages for a particular mutagenesis project, it is necessary to consider three properties: (i) site specificity, (ii) efficiency, and (iii) complexity.

Site specificity. All the methods discussed in this section can induce a variety of different point mutations localized within a defined interval of a DNA molecule. Exactly how a site or segment is targeted for mutagenesis and how large or small it can be are two of the principal variables that distinguish the different methods. Obviously, the optimal strategy for identifying the genetic information relevant to a particular phenotypic property is to target mutagenesis to the smallest interval in which all the information is likely to reside.

Efficiency. Every method of mutagenesis, whether targeted to the entire genome of a living organism or to a single nucleotide position in a purified DNA molecule, yields a mixture consisting of one or more mutant forms and wild type.

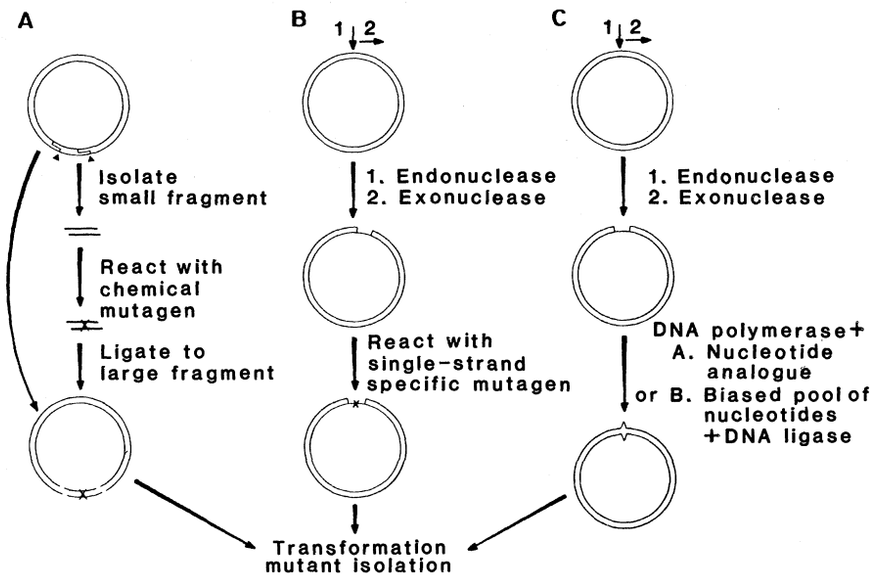

Fig. 1. Outline of three strategies for localized random mutagenesis. (A) Chemical mutagenesis of a restriction fragment isolated from a larger DNA molecule, followed by reassembly of complete molecules with DNA ligase. (B) Mutagenesis of a short single-stranded gap (or loop) in an otherwise duplex DNA molecule by means of a chemical mutagen that is specific for single-stranded DNA (for example, sodium bisulfite). (C) Nucleotide analogue incorporation, or nucleotide misincorporation, by DNA polymerase during synthesis from a primer-template.

As a measure of the accomplishment of mutation induction, efficiency can be defined as the fraction or percentage of all output DNA molecules (candidates) that do not have the starting wild-type sequence. In practical terms that are of the utmost importance, an increase in the efficiency of mutagenesis means that fewer candidates must be screened in order to find phenotypic mutations. Many of the new methods of in vitro mutagenesis have very high efficiencies and, consequently, new opportunities are presented for applying even very labor-intensive, "brute force" screens for mutants with new and rare phenotypes.

Complexity. In many situations, highly efficient mutagenesis is of value only when a large variety of different mutations is generated. Complexity represents a relative measure of the number of different types of nucleotide sequence changes that can be induced. With regard to base-substitution mutations, each of the four types of base pairs in DNA can substitute for each of the other three. Therefore, when mutagenesis is confined to one strand of DNA, there are 12 distinguishable types of base substitution. When both DNA strands are mutable, these 12 reduce to only six types. A low complexity, which is typical of most chemical mutagens used either in vivo or in vitro, translates into the induction of a subset of all possible substitutions at a subset of the different nucleotide positions in a segment of DNA.

One simple strategy for limiting the action of chemical mutagens to a specific DNA sequence is first to isolate a small restriction fragment containing the sequence and then, after in vitro reaction with the mutagen, to recombine the mutagenized fragment back into the DNA of origin (Fig. 1A). Nitrous acid, hydroxylamine, methoxyamine, and other mutagens have been successfully used to obtain localized random mutagenesis of acceptable efficiency, albeit of low complexity (*34–37*). One technical problem that frequently limits the applicability of this strategy is the requirement for efficient reconstruction of the original DNA molecule from the mutagenized fragment plus a larger untreated fragment. If instead of using DNA ligase to rejoin the correct fragments (which usually limits mutagenesis to those fragments generated with two single-cut restriction enzymes) the mutagenized fragment is reannealed to a single-stranded form of the original molecule, mutations can be induced by in vitro synthesis of the second strand and DNA transformation (*38, 39*), in much the same way as with synthetic oligonucleotides (see below).

To increase both the complexity and efficiency of chemical mutagenesis of isolated DNA fragments, Meyers and Maniatis (*40*) have developed a system for in vitro DNA synthesis that generates mutations by the use of a damaged DNA strand as template. After inducing a very limited number of lesions in single-stranded DNA that destroy the base-pairing ability of one or more of the four types of bases, avian myeloblastosis virus (AMV) reverse transcriptase is used to synthesize a second, complementary strand. Apparently at sites in the template where there is a damaged base, or no base at all, this highly error-prone polymerase will insert one of the four nucleotides in a nonspecific manner, leading to a very large variety of types of base substitution. After recombining the fragment back into a specially constructed cloning vector and transforming into *E. coli*, restriction fragments containing single-point mutations can be physically separated from the original wild-type fragment by electrophoresis on urea-gradient gels at high temperature (*41*). Although it requires a considerable number of steps, this approach is particularly attractive in situations in which it is preferable to isolate and sequence mutant forms before determining their phenotypic consequences (the phenotypic screen is very labor-intensive) and in situations in which phenotypically silent mutations are of interest.

A second strategy for achieving site specificity is to modify the local structure of a sequence on a DNA molecule in such a way that the sequence becomes more sensitive to a mutagenic reaction than the flanking sequences. If such mutagen-sensitive target sites can be introduced without reducing the DNA molecule to fragments, the problem of reassembly is avoided altogether. The chemical reactivity of a particular sequence or segment of a duplex DNA molecule can be drastically modified by converting the nucleotides in the segment to be mutagenized into a single-stranded state. By the appropriate enzymatic manipulation of circular DNA molecules in vitro (*32*), each of the three types of single-stranded structures—gaps, deletion loops, and displacement loops—can be introduced at more or less precisely defined sites in duplex DNA molecules and then can be used as targets for localized mutagenesis (Fig. 1B).

For example, the simple inorganic salt sodium bisulfite efficiently induces GC to AT mutations by catalyzing the deamination of cytosine residues in single-stranded, but not double-stranded, DNA (*42*). Consequently, each of the cytosine residues on one strand of a segment of DNA can be made susceptible to mutation by exposing the sequence in a single-stranded gap or loop. The extent of deamination of such reactive cytosines can be carefully controlled by the concentration of bisulfite and the time of the reaction, giving in vitro mutagenesis that is both highly localized and efficient (*32*,

43–46). However, the only mutations induced are C to T substitutions (or G to A substitutions as seen from the opposite strand), a level of complexity too low for many types of mutational analysis.

DNA polymerases, somewhat surprisingly, are emerging as the most versatile of the in vitro mutagens for achieving localized random mutagenesis that is both efficient and high in complexity. As mentioned earlier, polymerases will, in some cases, insert nucleotides nonspecifically onto a growing DNA chain opposite sites of damage in the template strand. Errors in the choice of nucleotide incorporated from an undamaged template during in vitro synthesis can be increased, by many orders of magnitude over the in vivo error rate, either by using nucleotide analogues (*47, 48*) or simply by biasing the pool of nucleoside triphosphates available to the DNA polymerase (Fig. 1C).

For instance, when incubated with a DNA molecule containing a short single-stranded gap plus one α-thiophosphate nucleoside triphosphate, the large fragment of *E. coli* DNA polymerase I can be made to misincorporate a single nucleotide with an efficiency that approaches 100 percent (*49*). The normally very low frequency at which the polymerase makes mistakes is amplified in this in vitro reaction because only one of the four nucleotide substrates is present. Once the enzyme misincorporates a nucleotide, the normal correction mechanism involving the polymerase's intrinsic 3′ to 5′ nuclease activity is blocked by the sulfur atom present in the terminal phosphodiester bond. In a second reaction, the remainder of the single-stranded gap is correctly filled in by extension from the mispaired 3′OH-terminus. Eleven of the 12 types of base substitution have been induced at roughly comparable frequencies by means of this strategy of gap-misrepair mutagenesis with each of the four α-thiophosphate nucleotides (*50*).

A second method for enhancing the rate of errors made by polymerase during the repair of a short single-stranded gap is to leave out one of the four deoxynucleoside triphosphate substrates (*49*). The gap-filling reaction can be driven to completion by addition of DNA ligase plus adenosine triphosphate, which seals the resulting nick and prevents subsequent proofreading. Although this method gives base-substitution mutagenesis of high complexity (plus frequent insertions and deletions), it has the serious disadvantage of frequently inducing multiple, clustered mutations (*51*).

Alternatively, in vitro repair synthesis with a DNA polymerase lacking an editing 3′ to 5′ exonuclease can be used to efficiently misincorporate the conventional nucleoside triphosphates (*52*). After generating unique sites for priming by annealing a short oligonucleotide to circular single-stranded DNA, it is possible to use AMV reverse transcriptase to misincorporate nucleotides at downstream positions. Again, by proper choice of reaction conditions and time of incubation, a large variety of types of base-substitution mutation can be induced with a high efficiency (*53–55*).

The principal technical problem associated with the routine use of DNA polymerases as in vitro mutagens is the general problem of positioning the 3′OH-terminus, which determines the site for misincorporation, at any specified nucleotide or at sites uniformally distributed across any specified interval. As mentioned earlier, completely random mutagenesis of a small circular plasmid can be achieved by nicking with deoxyribonuclease I plus ethidium bromide followed by limited 3′ to 5′ exonucleolytic digestion (which further randomizes the 3′OH-terminus) to produce a short gap (*50*). To obtain mutagenesis localized to a defined sequence, it is sometimes possible to move the 3′OH-terminus to a more or less defined position from a unique site nearby by means of controlled digestion with a 3′ to 5′ exonuclease (*56*) or controlled nick translation with DNA polymerase I (*44*). In addition, modifications in the chemical structure (*57*), genetic information (*53*), or replication pathway (*58*) of the primer-template often permit further improvements in the overall efficiency of mutagenesis by selecting specifically against the template strand or for the strand synthesized in vitro.

Oligonucleotide-Directed Mutagenesis

With the methods of localized random mutagenesis, the desired product is a pool of many different mutations that can be searched, by use of a selection or a screen, for the subset of mutations that cause specific phenotypic changes. However, when only one or a few mutations of defined DNA sequence (genotype) are of special interest, searching through pools of mutations becomes a very inefficient strategy of mutant isolation. In these situations, synthetic oligonucleotides can be used to construct mutant alleles of a cloned gene with base substitutions, insertions, or deletions, either singly or in any combination (*1, 59, 60*).

The enormous versatility of oligonucleotide mutagenesis derives from the fact that the mutation (the part of a gene sequence that differs from the wild type) is actually synthesized de novo as a short oligonucleotide. Because the chemistry involved in coupling nucleotides imposes no constraints on the base sequence of the product oligonucleotides, there are no constraints on the types of mutation that can be induced by this method.

There have been several recent reviews of the technical advances in the chemistry of oligonucleotide synthesis (*59, 61, 62*), and the availability of commercial automated synthesizers has made the actual synthesis of mutant oligonucleotides the most routine and reliable step in the overall protocol. Therefore, this section concentrates on the somewhat more problematic subsequent steps in which the mutation encoded in the sequence of a synthetic oligonucleotide is introduced into a cloned gene.

In the first strategy developed for incorporating a short mutant oligonucleotide into a longer segment of DNA, DNA polymerase is used to copy wild-type sequences onto both ends of the oligonucleotide. The simplest way this can be done is to make a single-stranded oligonucleotide that is identical in sequence, except for the mutation to be induced, to a site on the wild-type gene (the target site for mutagenesis). On annealing the oligonucleotide to a single-stranded form of the wild-type gene, a short heteroduplex forms that can serve as a primer for extension by DNA polymerase with the wild-type sequence as a template. If the single-stranded template is circular, the DNA polymerase may extend the newly synthesized strand all the way around the template until it reaches the 5′ end of the oligonucleotide to form a single-stranded nick. Covalent closure of the nick by DNA ligase generates a duplex circular molecule that has a wild-type sequence in the template strand and a mutant sequence in the newly synthesized strand (Fig. 2A). On transformation into a cell in which the DNA molecule undergoes replication, the two genotypes segregate and subsequently can be isolated as pure clones.

Since the initial application of this approach to the site-specific mutagenesis of single-stranded phage ϕX174 (*63, 64*), a number of extensions have resulted in its general applicability to any gene cloned onto a circular vector. For example, the requirement that the target sequence be in a single-stranded state can be satisfied (i) by recloning the gene into an M13 single-stranded vector (*65*); (ii) or into a plasmid vector that contains an

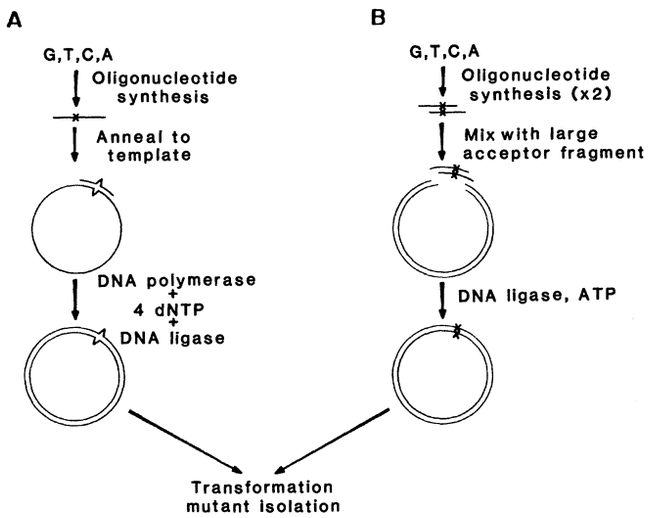

Fig. 2. Outline of two strategies for oligonucleotide-directed mutagenesis. (A) A single mutant oligonucleotide is synthesized and used as a primer for second-strand synthesis by DNA polymerase plus DNA ligase. (B) Two synthetic oligonucleotides (×2) of complementary sequence anneal to form a short duplex, which is joined to a large restriction fragment by DNA ligase to generate a complete duplex DNA molecule. ATP, adenosine triphosphate; dNTP, deoxynucleoside triphosphates.

M13 replication origin [which leads to packaging of single-stranded plasmid DNA into virus particles after superinfection with M13 (*66*)]; (iii) by enzymatically degrading part or all of one strand of a duplex plasmid with exonuclease III (*67*); (iv) by forming a heteroduplex molecule with an appropriate single-stranded gap (*68*); or (v) by direct annealing of the oligonucleotide to duplex DNA to form a D-loop (*69*). Instead of requiring that the polymerase complete the entire strand to attach wild-type sequences to the 5′ end of the mutant oligonucleotide, a second oligonucleotide can be used to prime DNA synthesis from a site upstream of the mutant oligonucleotide, yielding significantly higher efficiencies of mutagenesis (*70, 71*).

Perhaps the most important innovation has been the use of the mutant oligonucleotide, after end labeling with ^{32}P, as a hybridization probe to physically screen for mutant alleles among the progeny molecules generated on replication of the mutagenized DNA (*67*). First, the labeled oligonucleotide is hybridized with the DNA of candidate molecules that have been immobilized on nitrocellulose under conditions of very low stringency. Autoradiography is then performed after each in a series of successive washes at progressively higher stringency. This procedure can reliably discriminate between mutant and wild-type DNA molecules, either as plasmids in *E.coli* colonies (*72*) or as single-stranded M13 DNA in phage plaques (*65*). This technique permits the isolation of mutant alleles in the absence of a scoreable phenotype and solely on the basis of nucleotide sequence, even when the efficiency of mutagenesis is considerably less than 1 percent.

With the reduction in the time and expense of oligonucleotide synthesis, a second strategy requiring two unique oligonucleotides to incorporate a mutant sequence into an intact gene has become a feasible alternative to the oligonucleotide primer strategy. If a short DNA duplex containing a mutant sequence is formed by annealing two complementary oligonucleotides, it can be inserted into a much larger DNA molecule that contains all the remaining gene sequences, provided that the ends of both molecules have complementary single-stranded "sticky" ends. The DNA is circularized by joining of the two molecules with DNA ligase, and an intact, complete gene is formed that contains the mutant duplex segment (*73*) (Fig. 2B).

The synthetic mutant duplex can be made with any type of cohesive ends by appropriately extending the nucleotide sequence of each of the two synthetic oligonucleotides beyond the complementary region. However, to arrange for the appropriate sticky ends at precisely the position where the short mutant duplex is to be inserted almost invariably requires the introduction of unique restriction sites at the boundaries of the insertion site, either by the oligonucleotide primer method (*74, 75*) or through the total chemical synthesis of a gene with silent codon changes to create new restriction sites (*76*). Although this strategy requires considerably more oligonucleotide synthesis plus an initial partial or total redesign of the nucleotide sequence of the gene to be mutagenized, the efficiency of mutagenesis can approach 100 percent and the labor involved in recovering mutations in a double-stranded vector can be significantly less than with other procedures. Consequently, this strategy is likely to be most applicable for relatively small genes or DNA segments that are to be very intensively analyzed by site-specific mutagenesis.

Additional technical improvements and refinements are continuing to be made in both strategies of oligonucleotide-directed mutagenesis, particularly with regard to reducing the time and effort involved in assembling large collections of mutations of known sequence distributed over small intervals. At least two approaches have been described that permit the simultaneous synthesis of multiple oligonucleotides with different sequences. Small paper disks can be used as the support matrix for solid-phase synthesis; by simply sorting each disk (on which one individual chain is synthesized) into groups for condensation with A, G, C, or T before each coupling, as many as 100 different oligonucleotides can be synthesized in four reaction vessels (*77–79*). Alternatively, mixtures of oligonucleotides having the same nucleotide sequence, except for different base substitutions at one or more designated positions, can be synthesized in one reaction vessel by adding a mixture of several activated nucleotides to the coupling reaction when the appropriate point in the synthesis is reached (*74*). By coupling mixtures of nucleotides at several points in the synthesis, complex pools of oligonucleotides encoding different single- and multiple-base substitutions can be generated (*74*) for use in either the oligonucleotide primer or short duplex methods of mutagenesis.

The tactic of synthesizing a collection of mutant oligonucleotides to induce many different mutations in one experiment can be viewed as a form of localized random mutagenesis. It seems reasonable to expect that, in the future, oligonucleotide-directed mutagenesis will evolve from its current status as a reasonably efficient method for altering single nucleotide positions into a highly efficient method for saturation mutagenesis of DNA segments varying in length from 2 to perhaps as many as 50 base pairs. However, for mutagenesis of longer DNA segments, the methods of localized random mutagenesis discussed earlier are likely to continue to provide a more efficient means of generating pools of mutations. Thus, in many respects, these two different classes of methods for in vitro mutagenesis are distinctly complementary. Which strategy is optimal for a given mutational analysis depends primarily on the size of the target sequence that must be covered to extract all the genetic information relevant to the phenomenon under study.

Finally, all the methods of in vitro

mutagenesis, including those that require synthetic oligonucleotides (*80*), can generate nucleotide sequence changes at sites other than those targeted for mutations. Therefore, even though a specific mutation of defined nucleotide sequence has been isolated and sequenced, an observed change in phenotype cannot immediately be attributed to the constructed mutation. Except for determining the entire nucleotide sequence of the gene or genetic element that has undergone mutagenesis, genetic mapping of the mutation is the only conclusive means of establishing the connection between a change in nucleotide sequence and its phenotypic consequences.

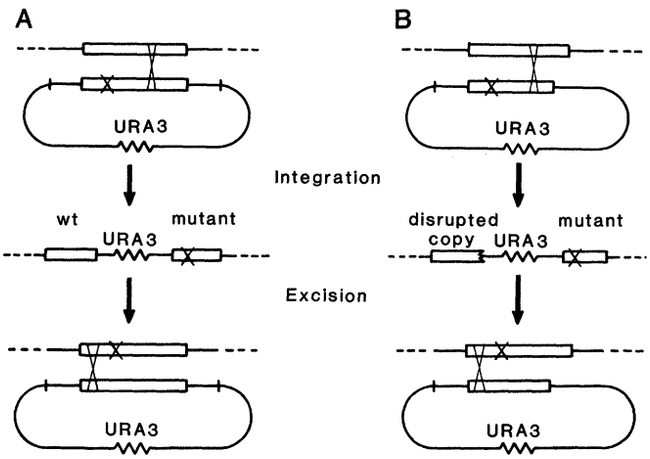

Fig. 3. Two strategies for recombining mutations constructed in vitro into the cellular gene. The *URA3* gene is a selectable marker in the yeast *S. cerevisiae*. The small x indicates the mutant, the large X indicates the sites of recombination.

Determining Phenotype and the Context Problem

The methods described make possible the generation of a vast array of mutant forms of a cloned gene; DNA sequence technology makes possible the rapid determination and/or confirmation of the alterations in sequence, even before any attempt to assess phenotypic consequence. Up to this point, mutagenesis can be viewed as an exercise in nucleic acid biochemistry. To use mutations to draw inferences about underlying biological phenomena, the next step, and in almost all studies the most critical step, is the determination of the functional consequences of the mutations.

Although it is biochemically correct to define a gene mutation as a change in nucleotide sequence, from the point of view of genetic analysis it is more useful to define a mutation as a change in a unit of genetic information. As with all information, the meaning conveyed is dependent on the context in which it is expressed. Whether or not a mutation manifests itself by a change in phenotype can depend very significantly on the genetic context. For example, a recessive lethal mutation may lead to the immediate death of a haploid cell and yet be totally without observable effect when combined with a wild-type allele in a diploid cell. A mutant gene isolated after in vitro mutagenesis is, by definition, recovered in a foreign context. Before the functional consequences of such a mutation can be accurately determined and inferences correctly drawn about the wild-type gene, it may be necessary to partially, or perhaps totally, restore the mutant gene to the same genetic context in which the wild-type form occurs in nature.

The severity of this context problem varies with the system under study. For example, if biochemists are using in vitro mutagenesis to test hypotheses about the catalytic roles of certain amino acid residues in an enzyme of known three-dimensional structure, they will usually not be concerned that the organism in which the mutant gene was expressed or the transcriptional control sequences at its 5′ end will influence the conclusions concerning the properties of the purified protein. However, molecular biologists studying the details of expression of a mouse α-globin gene can draw no inferences at all from the level of gene transcription in *E. coli* from a cloned complementary DNA inserted into a λ phage vector. In this situation the ideal experiment would be to return the mutant mouse gene to its normal genetic context so that any alterations in phenotype could be confidently attributed to the mutation, rather than to possible artifacts arising from the abnormal environment in which gene has been placed.

To meet the challenge of studying mutations in their proper context, molecular biologists have devised a number of experimental solutions. One partial solution that permits study in a genetic context closer to that of the wild-type gene in nature is to examine the expression of a cloned gene, side by side with a collection of mutant forms, in a convenient eukaryotic cell. Genes can be injected into *Xenopus* oocytes and, in a number of cases, the transcriptional and translational behavior of mammalian genes in this context is comparable to the same processes in the cells of the parent mammal (*81*). Alternatively, it may be feasible to reintroduce the mutant genes into a cell line derived from the source organism or from a closely related species by DNA-mediated transformation and then to assay expression in either transiently or in stably transformed cells (*82*). In studies of this type with cultured cells, it is accepted practice to define "wild type" as the genotype and phenotype of the unmodified cloned gene under the same assay conditions. However, in the interpretation of phenotypes of mutant genes, it must be remembered that inferences with regard to the natural "wild type"—that is, the intact organism and the genetic context it provides for the wild-type gene—cannot be made without introducing untestable assumptions.

A complete solution to the context problem would require replacing the wild-type gene at its normal genetic locus in a germline cell (a cell that can give rise to an intact organism) with the mutant allele constructed by in vitro mutagenesis. As discussed in the next section, such allele replacements have become routine genetic manipulations in prokaryotes and several fungi. In a number of higher eukaryotes, cloned genes can be efficiently introduced into germline cells, albeit not at their normal chromosomal location. Microinjection of DNA into oocytes of mice (*82*) and *Caenorhabditis elegans* (*83*), for example, can lead to the stable germline inheritance of some of the injected DNA sequences. In *Drosophila* (*84*) and some dicotyledonous plants (*85*), stable inheritance is greatly facilitated by the presence of transposon sequences in the injected DNA molecules. Initial characterization of the expression of *Drosophila* genes integrated at different sites in the genome suggest that, in some cases, expression and regulation of the inserted gene is not greatly altered by the new and slightly different genetic context (*86–88*).

Allele Replacement

The idea of inducing a mutation into a small portion of a genome and using a cell's recombinational machinery to transfer the mutation onto the genome is not new. In *E. coli*, cellular genes carried on specialized transducing phages (or F′ episomes) can be mutagenized and transferred to an unmutagenized cell, where

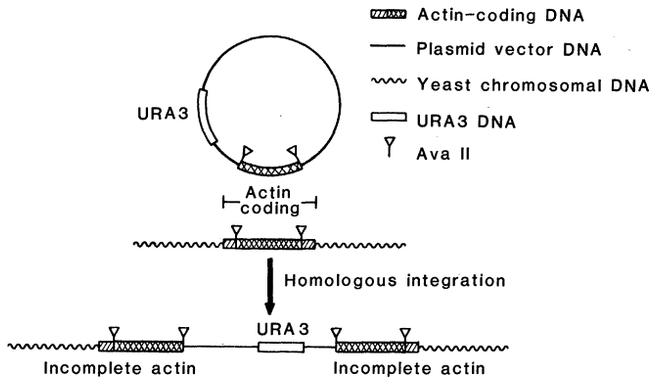

Fig. 4. Disruption of the yeast actin gene by integrative transformation (*99*).

homologous recombination between the mutagenized copy and the chromosomal copy of the gene can result in the desired replacement of wild type by mutant. For this replacement of one allele by another to occur, the equivalent of two crossovers must take place, one on either side of the mutation.

To carry out a similar sort of allele replacement between a mutation constructed in vitro in a cloned gene copy on a circular plasmid and its normal genetic locus requires essentially the same sequence of events (Fig. 3A). The recombinant plasmid, which also carries a selectable genetic marker, is introduced into the host cell by DNA transformation. Recombination between gene sequences on the plasmid and the homologous sequences on the genome must occur. If the only pathway for stable transformation of the genetic marker is through integration into the cell's genome (that is, the plasmid is unable to replicate autonomously in the host cell), the first crossover can be obtained by applying selection for the marker. Integration of the circular DNA results in a tandem duplication of the sequences of the cloned gene, a structure that is unstable to intramolecular recombination between the two repeats. One consequence of this instability is that the second crossover can usually be obtained by selecting, or in some cases simply by screening, for spontaneous loss of the selectable marker. When the two crossovers occur on opposite sides of the mutation, the end result is replacement of the wild-type gene with the mutant copy. Alternatively, both crossovers can be selected in a single event by transformation with a linear DNA fragment that includes a selectable marker closely linked to the mutation, a situation very similar to generalized transduction in bacteria.

Strategies involving circular and linear DNA have been successfully used to replace normal genes with mutant alleles made in vitro in a variety of bacteria [*Salmonella typhimurium* (*89*), *E. coli* (*90*), *Bacillus subtilis* (*91*)], as well as in the yeast *Saccharomyces cerevisiae* (*92*) and the filamentous fungus *Aspergillus nidulans* (*93*). In these bacteria and in yeast, homologous recombination appears to be the dominant mode for stable integration of selectable markers. In *Aspergillus*, however, many of the integrative events appear to involve nonhomologous recombination, thereby resulting in DNA insertions at improper sites in the genome (*93*). Regrettably, in *N. crassa* (*94*), as well as in nearly all higher eukaryotic cells (*95, 82*), integrative transformation by exogenous DNA molecules occurs predominantly via nonhomologous pathways, making replacement of alleles at the chromosomal locus problematic.

Two modified forms of allele replacement deserve brief mention. When mutations are induced in a cloned gene that has a deletion at one end, integration by homologous recombination results in a tandem duplication in which only one of the two copies is intact (Fig. 3B). As a result, recessive mutations can be expressed in the transformed cell, there being no second, wild-type copy to mask the phenotypic consequences. This permits a selection or screen to be applied directly to the initial transformants (instead of transformants that subsequently excised the plasmid), thereby simplifying the identification of recessive mutations. This strategy has been used to isolate conditional lethal mutations in the yeast genes encoding actin (*96*), β-tubulin (*97*), and topoisomerase II (*98*).

Similarly, integration of a gene copy deleted at both ends (an internal DNA fragment) generates a tandem duplication in which neither of the two copies is intact (Fig. 4). This "integrative-disruption" technique has been used to determine the phenotype of null mutations in yeast (*3, 99*), *Aspergillus* (*92*), and several bacterial species (*100*). Alternatively, a gene can be disrupted by DNA transformation with a linear DNA fragment containing a selectable marker inserted within a cloned copy of the gene (*101, 102*), giving rise to a very stable null mutation.

In summary, methods for precise allele replacement are now routine for some microorganisms, making the ideal mutation experiment readily achievable. It is possible to clone a gene from these organisms, make a single mutation, replace the normal gene in its proper genetic context with the mutant allele, and then study the consequences, even if they are quite subtle or altogether unexpected. The ability to do such experiments in *S. cerevisiae* is a major reason for the current enthusiasm for studying gene regulation and basic problems of cell biology in this simple eukaryotic organism. It is hoped that methods for gene replacement will be developed for higher eukaryotes as well in the near future, at which time the genes and genetic regulatory elements involved in developmental pathways and neurological phenomena may become amenable to systematic analysis with gene mutations.

Conclusions

New in vitro methods for altering the nucleotide sequence of DNA molecules have emerged during the past few years and have greatly expanded the range of problems to which mutational analysis can be applied. To a considerable degree, the armamentarium the geneticist needs to collect gene mutations for studying phenomena at all levels of biological organization and complexity is now complete: Mutagenesis can be targeted to the entire genome, to individual genes and gene clusters, to structural and regulatory segments of genes, and to single nucleotide positions. The challenge for the future will not be in the isolation of greater and greater numbers of mutations, but rather in the detailed reading and interpretation of the story each mutation has to tell.

References and Notes

1. M. Smith, *Annu. Rev. Genet.*, in press.
2. D. Shortle, D. DiMaio, D. Nathans, *ibid.* **15**, 265 (1981).
3. J. H. Thomas, P. Novick, D. Botstein, in *Molecular Biology of the Cytoskeleton*, G. G. Borisy, D. W. Cleveland, D. B. Murphy, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1984).
4. H. J. Muller, *Science* **66**, 84 (1927).
5. C. Auerbach and J. M. Robson, *Proc. R. Soc. Edinburgh B* **62**, 279 (1947).
6. G. W. Beadle and B. Ephrussi, *Genetics* **21**, 225 (1936).
7. G. W. Beadle and E. L. Tatum, *Proc. Natl. Acad. Sci. U.S.A.* **27**, 499 (1941).
8. P. deLucia and J. Cairns, *Nature (London)* **224**, 1164 (1969).

9. J. S. Hong and B. N. Ames, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 3158 (1971).
10. N. Kleckner, J. Roth, D. Botstein, *J. Mol. Biol.* **116**, 125 (1977).
11. W. G. Shanabruch and G. C. Walker, *Mol. Gen. Genet.* **179**, 289 (1980).
12. G. B. Ruvkun and F. M. Ausubel, *Nature (London)* **289**, 85 (1981).
13. P. M. Bingham, R. Levis, G. M. Rubin, *Cell* **25**, 693 (1981).
14. M. Breindl, K. Harbers, R. Jaenisch, *ibid.* **38**, 9 (1984).
15. R. J. Legerski, J. L. Hodnett, H. B. Gray, *Nucleic Acids Res.* **5**, 1445 (1978).
16. S. Sakonju, D. Bogenhagen, D. D. Brown, *Cell* **19**, 13 and 27 (1980).
17. S. D. Putney, S. J. Benkovic, P. R. Schimmel *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7350 (1981).
18. S. Henikoff, *Gene* **28**, 351 (1984).
19. F. Heffron, M. So, B. J. McCarthy, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 6012 (1978).
20. S. L. McKnight and R. Kingsbury, *Science* **217**, 316 (1982).
21. M. J. Casadaban, J. Chou, S. N. Cohen, *J. Bacteriol.* **143**, 971 (1980).
22. C. M. Gorman, L. F. Moffatt, B. H. Howard, *Mol. Cell. Biol.* **2**, 1044 (1982).
23. P. Bassford *et al.*, in *The Operon*, J. H. Miller and W. S. Reznikoff, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1978).
24. M. Rose, M. J. Casadaban, D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2460 (1981).
25. L. Guarente and M. Ptashne, *ibid.*, p. 2199.
26. G. M. Weinstock, M. L. Berman, T. J. Silhavy, in *Expression of Cloned Genes in Prokaryotic and Eukaryotic Cells*, T. S. Papas *et al.*, Eds. (Elsevier/North-Holland, New York, 1983).
27. S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.* **47**, 403 (1961).
28. E. C. Cox, *Annu. Rev. Genet.* **10**, 135 (1976).
29. P. Youderian, S. Bouvier, M. S. Susskind, *Cell* **30**, 843 (1982).
30. J. R. Broach, V. R. Guarascio, M. Jayaram, *ibid.* **29**, 227 (1982).
31. L. Greenfield, L. Simpson, D. Kaplan, *Biochim. Biophys. Acta* **407**, 365 (1975).
32. D. Shortle and D. Botstein, *Methods Enzymol.* **100**, 457 (1983).
33. D. Shortle, *Gene* **22**, 181 (1983).
34. D. Solnick, *Nature (London)* **291**, 508 (1981).
35. C. T. Chu, D. S. Parris, R. A. F. Dixon, F. E. Farber, P. A. Schaffer, *Virology* **98**, 168 (1979).
36. S. Busby, M. Irani, B. deCrombrugghe, *J. Mol. Biol.* **154**, 197 (1982).
37. J. T. Kadonaga and J. R. Knowles, *Nucleic Acids Res.* **13**, 1733 (1985).
38. W. E. Borrias, I. J. C. Wilschut, J. M. Vereijken, P. J. Weisbeek, G. A. van Arkel, *Virology* **70**, 195 (1976).
39. S. Hirose, K. Takeuchi, Y. Suzuki, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7258 (1982).
40. R. M. Meyers and T. Maniatis, *Science* **229**, 242 (1985).
41. S. G. Fischer and L. S. Lerman, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1579 (1983).
42. H. Hayatsu, *Prog. Nucleic Acids Res. Mol. Biol.* **16**, 75 (1976).
43. D. Shortle and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2170 (1978).
44. D. DiMaio and D. Nathans, *J. Mol. Biol.* **140**, 129 (1980).
45. K. W. C. Peden and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7214 (1982).
46. D. Kalderon, B. A. Oostra, B. K. Ely, A. S. Smith, *Nucleic Acids Res.* **17**, 5161 (1982).
47. W. Muller, H. Weber, F. Meyer, C. Weissmann, *J. Mol. Biol.* **124**, 343 (1978).
48. H. Weber *et al.*, in *The ICN-UCLA Symposia on Molecular and Cellular Biology* (Academic Press, New York, 1982), vol. 12.
49. D. Shortle, P. Grisafi, S. J. Benkovic, D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1588 (1982).
50. D. Shortle and B. Lin, *Genetics* **110**, 539 (1985).
51. K. M. Overbye and D. Botstein, unpublished observations.
52. R. A. Zakour and L. A. Loeb, *Nature (London)* **295**, 708 (1982).
53. C. Traboni, R. Cortese, C. Ciliberto, G. Cesareni, *Nucleic Acids Res.* **11**, 4229 (1983).
54. R. A. Zakour, E. A. James, L. A. Loeb, *ibid.* **12**, 6615 (1984).
55. J. J. Champoux, *J. Mol. Appl. Genet.* **2**, 45 (1984).
56. P. T. Englund, S. S. Price, P. H. Weigel *Methods Enzymol.* **29**, 273 (1974).
57. T. A. Kunkel, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488 (1985).
58. P. Abarzua and K. J. Marians, *ibid.* **81**, 2030 (1984).
59. K. Itakura, J. J. Rossi, R. B. Wallace, *Annu. Rev. Biochem.* **53**, 323 (1984).
60. C. S. Craik, *Biotechniques* **3**, 12 (1985).
61. M. H. Caruthers, S. L. Beaucage, C. Becker, W. Efcavitch, E. F. Fisher, in *Genetic Engineering: Principles and Methods*, J. K. Setlow and A. Hollaender, Eds. (Plenum, New York, 1982), vol. 4.
62. E. Ohtsuka, M. Ikehara, D. Soll, *Nucleic Acids Res.* **10**, 653 (1982).
63. C. A. Hutchinson III *et al.*, *J. Biol. Chem.* **253**, 6551 (1978).
64. A. Razin, T. Hirose, K. Itakura, A. D. Riggs, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4268 (1978).
65. M. J. Zoller and M. Smith, *Nucleic Acids Res.* **10**, 6487 (1982).
66. L. Dente, G. Caesareni, P. Cortese, *ibid.* **11**, 1645 (1983); R. J. Zagursky and M. L. Berman, *Gene* **27**, 183 (1984); A. Levinson, D. Silver, B. Seed, *J. Mol. Appl. Genet.* **2**, 507 (1984).
67. R. B. Wallace *et al.*, *Science* **209**, 1396 (1980).
68. W. Kramer, V. Drutsa, H. W. Jansen, B. Kramer, M. Pflugfelder, H. J. Fritz, *Nucleic Acids Res.* **12**, 9441 (1984).
69. M. Schold, A. Colombero, A. A. Reyes, R. B. Wallace, *DNA* **3**, 469 (1984).
70. K. Norris, F. Norris, L. Christiansen, N. Fiil, *Nucleic Acids Res.* **11**, 5103 (1983).
71. M. J. Zoller and M. Smith, *DNA* **3**, 479 (1984).
72. R. B. Wallace, M. Schold, M. J. Johnson, P. Dembek, K. Itakura, *Nucleic Acids Res.* **9**, 3647 (1981).
73. K. M. Lo, S. S. Jones, N. R. Hackett, H. G. Khorana, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2285 (1984).
74. M. D. Matteucci and H. L. Heynecker, *Nucleic Acids Res.* **11**, 3113 (1983).
75. J. A. Wells, M. Vasser, D. B. Powers, *Gene*, **34**, 315 (1985).
76. K. P. Nambiar *et al.*, *Science* **223**, 1299 (1984).
77. R. Frank, W. Heikens, G. Heisterber-Montsis, H. Blocker, *Nucleic Acids Res.* **11**, 4366 (1983).
78. H. W. D. Mathes *et al.*, *EMBO J.* **3**, 801 (1984).
79. J. Ott and F. Eckstein, *Nucleic Acids Res.* **12**, 9137 (1984).
80. K. A. Osinga, A. M. van der Blick, G. van der Hoist, M. J. A. G. Koerkamp, H. F. Tabak, *ibid.* **11**, 3595 (1983).
81. J. B. Gurdon and D. A. Melton, *Annu. Rev. Genet.* **15**, 189 (1981).
82. R. Kucherlapati and A. I. Skoultchi, *Critical Rev. Biochem.* **16**, 349 (1984).
83. D. T. Stinchcomb, J. Shaw, D. Hirsh, in preparation.
84. A. C. Spradling and G. M. Rubin, *Science* **218**, 341 (1982).
85. K. A. Barton, A. N. Binns, A. J. M. Matzke, M. D. Chilton, *Cell* **32**, 1033 (1983).
86. S. B. Scholnick, B. A. Morgan, J. Hirsh, *ibid.* **34**, 37 (1983).
87. A. C. Spradling and G. M. Rubin, *ibid.*, p. 47.
88. D. A. Goldberg, J. W. Posakony, T. Maniatis, *ibid.*, p. 59.
89. N. I. Gutterson and D. E. Koshland, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **80**, 489 (1983).
90. M. Jasin, L. Regan, P. R. Schimmel, *Cell* **36**, 1089 (1984).
91. P. Youngman *et al.*, *Science* **228**, 285 (1985).
92. S. Scherer and R. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4951 (1979).
93. R. Morris, personal communication.
94. C. Yanofsky, personal communication.
95. D. M. Robins, S. Ripley, A. S. Henderson, R. Axel, *Cell* **23**, 29 (1981).
96. D. Shortle, P. Novick, D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4889 (1984).
97. T. Huffaker and D. Botstein, unpublished observations.
98. C. Holm, T. Goto, J. C. Wang, D. Botstein, *Cell*, in press.
99. D. Shortle, J. E. Haber, D. Botstein, *Science* **217**, 371 (1982).
100. F. A. Ferrari, E. Ferrari, J. A. Hoch, *J. Bacteriol.* **152**, 780 (1982).
101. R. J. Rothstein, *Methods Enzymol.* **101**, 202 (1983).
102. S. C. Winans, S. J. Elledge, J. H. Krueger, G. C. Walker, *J Bacteriol.* **161**, 1219 (1985)

103. Supported by grants from the National Institutes of Health (D.B. and D.S.) and from the American Cancer Society (D.B.). We would like to thank M. Smith for making a copy of his review available before publication and our many colleagues, current and past, for helpful discussions.