Genetica **86**: 99–111, 1992.
© 1992 *Kluwer Academic Publishers. Printed in the Netherlands.*

Natural genetic engineering in evolution

J. A. Shapiro
Department of Biochemistry and Molecular Biology, University of Chicago, 920 E. 58th St., Chicago, IL 60637, USA

Received and accepted 10 March 1992

Abstract

The results of molecular genetics have frequently been difficult to explain by conventional evolutionary theory. New findings about the genetic conservation of protein structure and function across very broad taxonomic boundaries, the mosaic structure of genomes and genetic loci, and the molecular mechanisms of genetic change all point to a view of evolution as involving the rearrangement of basic genetic motifs. A more detailed examination of how living cells restructure their genomes reveals a wide variety of sophisticated biochemical systems responsive to elaborate regulatory networks. In some cases, we know that cells are able to accomplish extensive genome reorganization within one or a few cell generations. The emergence of bacterial antibiotic resistance is a contemporary example of evolutionary change; molecular analysis of this phenomenon has shown that it occurs by the addition and rearrangement of resistance determinants and genetic mobility systems rather than by gradual modification of pre-existing cellular genomes. In addition, bacteria and other organisms have intricate repair systems to prevent genetic change by sporadic physico-chemical damage or errors of the replication machinery. In their ensemble, these results show that living cells have (and use) the biochemical apparatus to evolve by a genetic engineering process. Future research will reveal how well the regulatory systems integrate genomic change into basic life processes during evolution.

Introduction: Three lessons from molecular genetics

Recombinant DNA technology and DNA sequencing have made it possible to test theories about how genomes change in evolution. The results have often been surprising and have raised serious challenges to conventional evolutionary thinking. While it is always possible to adapt existing theories to unexpected observations, it is also useful to ask whether those observations indicate a different way of approaching the evolutionary process. The objective of this paper is to discuss certain kinds of molecular genetic data which, in the author's opinion, raise serious questions about the prevailing evolutionary wisdom based on notions of piecemeal, stochastic genetic change due to replication errors and physico-chemical instabilities. Consequently, the presentation will emphasize the theoretical implications of the results and will be partisan in favor of the need for new perspectives. It is based on my experience of over two decades in bacterial genetics, where the tools of the trade include phages, plasmids and transposons. Working with these agents for rearranging and transferring DNA molecules leads one to see the genome as a dynamic information storage system that is always subject to rapid modification.

There are three broad areas in which molecular observations have consistently suggested the need to alter long-held assumptions about the ways that genomes change during evolution.

Genetic conservation across taxonomic boundaries

There is a surprisingly high degree of conservation of protein structure and function among very different organisms. The excitement created when Lee and Nurse (1987) demonstrated that a human cDNA clone would complement a yeast cell-cycle mutation is a good illustration of how unexpected some of these observations have been. It may be

argued with hindsight that conservation is to be expected in molecules controlling basic cellular housekeeping functions, but the same result has been obtained with the transcriptional regulatory molecules controlling developmental gene expression in plants, insects and vertebrates (Akam, 1989; Goff *et al.*, 1990). In some cases, the developmental regulatory protein from one taxon even functions properly in the heterologous host (McGinnis *et al.*, 1990). Thus, even for processes like morphogenesis which are most subject to change during evolution, conservation is the rule. In some cases, conservation extends from bacteria through plants and animals (e.g. Inouye *et al.*, 1983; Baker & Saier, 1990). Taken together with the fundamental equivalence of physiological processes in all living organisms, these observations suggest that evolutionary novelty often does not reside in the invention of new biochemical processes by the continual modification and selection of individual proteins. Instead, evolution appears to proceed by the utilization of basic biochemical routines in different combinations in different organisms. With few exceptions, the structural proteins of all mammals, for example, are probably interchangeable; what makes a mouse different from an elephant is when and how those molecules are synthesized and assembled during development.

The mosaic structure of genomes

Comparisons of DNA structures within and between different species have revealed an underlying mosaic structure to all genomes. The data can be considered under several different categories.

Protein coding regions. Proteins and their coding regions are composed of domains, and these domains can each be used many times in various combinations with other domains. In eukaryotes, the domains frequently correspond to exons (Blake, 1985), but even in yeast and bacterial protein coding sequences without introns, domain mosaics are commonly observed. The two-component transcriptional regulatory systems in bacteria are an excellent example. These systems consist of environmental or metabolic sensors with protein kinase activities paired with transcriptional regulators whose functions are determined by whether or not they are phosphorylated. Both the sensors and regulators fall into families which share conserved domains (protein kinase, phosphorylation substrate and DNA binding regions) but differ in other domains according to the specific regulatory task to be accomplished (Stock *et al.*, 1990).

Transcriptional regulatory regions. Molecular analysis of transcriptional regulation has uncovered a fundamental mosaic pattern to the organization of 5′ control regions. In both prokaryotes and eukaryotes, these regions are composed of distinct motifs (frequently called consensus sequences or 'boxes') for the binding and action of different transcription factors and RNA polymerases. A particular motif is generally shared among many transcription units, such as a promoter for a particular polymerase or a binding site for a particular transcription factor. These shared motifs are important in building coordinated regulatory systems of multiple unlinked genetic loci that respond to common transcription proteins, such as the *E. coli* heat-shock regulon transcribed by RNA polymerase with the σ^{32} subunit or the SOS regulon repressed by the LexA protein (Hoopes & McClure, 1987).

The way that different motifs are combined affects the nature of the regulatory process. This was first discovered in analyzing the action of the λcI repressor: (i) cooperative binding depends upon the spacing of adjacent operator sequence repeats, and (ii) the relative positions of operator and promoter motifs are critical to whether the repressor blocks or stimulates transcription (Ptashne, 1986). There also appears to be a hierarchical aspect to the mosaic structures of transcriptional regulatory regions. A particular motif may itself have subdomains. Enhancers frequently contain short sequence repeats (Müller *et al.*, 1988), and prokaryotic promoters have two distinct polymerase interaction sites, such as the -10 (Pribnow box) and -35 domains recognized by the canonical σ^{70}-containing polymerase used during rapid exponential growth (Hoopes & McClure, 1987). These various sequence motifs constitute codons for an additional type of genetic code that is used to control when and where transcription occurs (e.g. Rosenfeld *et al.*, 1989).

Repetitive sequence elements. Binding sites in regulatory regions are only one class of repetitive DNA sequence element present in genomes. Other aspects of genetic function (e.g. replication, proofreading, recombination, chromosome mechanics) are also controlled by particular genetic codes, and the motifs which constitute the appropriate codons

are inscribed in DNA molecules (Trifonov & Brendel, 1986). Examples of codons which have meaning for genetic functions other than protein synthesis include: Dam methylation sites for proofreading and *chi* sequences for homologous recombination in *E. coli* (Modrich, 1987; Marinus, 1989; Stahl, 1979), the 11 b.p. species identifiers which distinguish *Haemophilus influenzae* DNA for transformation (Smith *et al.*, 1981), and centromere and telomere repeats in eukaryotes (Blackburn & Szostak, 1984; Zakian, 1989; Blackburn, 1991). In addition, all genomes contain a wide array of repeat elements whose functional significance remains unclear. Some of these are found clustered in tandem arrays, such as the alphoid DNA that constitutes a high percentage of primate genomes (Donehower & Gillespie, 1979), while other repeat elements are dispersed, such as the *Alu* retroposons which are virtually ubiquitous in primate euchromatin (Rogers, 1985; Weiner *et al.*, 1986; Deininger, 1989).

It is significant, as Dover and others have pointed out, that these repetitive sequences show a high degree of taxonomic specificity (Dover, 1982). For example, a very good primate taxonomy can be constructed on the basis of restriction site polymorphisms in alphoid DNA (Donehower & Gillespie, 1979), and various groups of mammals are identifiable by their characteristic dispersed retroposons (Rogers, 1985; Weiner *et al.*, 1986). Indeed, the abundance and distribution of repetitive elements are often among the best available genomic characters for species identification. This is true both of prokaryotes and eukaryotes. It has long been noted that *E. coli* and *Salmonella typhimurium* share extremely similar chromosome maps (Riley & Krawiec, 1987), but they differ widely in their IS elements: *E. coli* has many dispersed all over the genome (Galas & Chandler, 1989), whereas *S. typhimurium* only has a few copies of IS200, an element that is very rarely recovered in new genomic positions (Casedesús & Roth, 1989). Similarly, *Drosophila melanogaster* and *Drosophila simulans* have virtually identical genetic maps, chromosome structures and banding patterns, but the content and distribution of repetitive DNA elements is very different between the two species (Dowsett, 1983). If we think about genomes as complex information storage and retrieval systems (Shapiro, 1991), then it may be helpful to consider the possibility that taxonomically-specific repetitive DNA sequences help define the 'system architecture' of each species, perhaps serving a role in hierarchical organization of different genomic regions.

The mosaic pattern of organization is thus seen at multiple levels: within isolated genetic units (coding sequences, regulatory regions), marking specific locations on larger genetic structures (plasmids, chromosomes) and dispersed throughout the genome (transposable elements, species identifier sequences, recombination sites). A particularly well-documented case at the whole genome level concerns the lambdoid bacteriophages of *E. coli.* These phages all share a common organization of clustered functional regions. However, there are multiple different sequences which can provide each functional module. Interestingly, any one phage is a pastiche of modules from different sources because its DNA shows patchwork homology with several other members of the group (Highton *et al.*, 1990).

The biochemical mechanisms of genetic change

As pointed out by Gilbert and others, the mosaic nature of regions encoding protein structure indicates that rearrangement of DNA sequence elements to construct new genetic units is a fundamental evolutionary process (Gilbert, 1978). The question thus arises as to how such rearrangements can occur. Mechanistically, that question has been answered by the discovery of mobile genetic elements and of multiple biochemical systems for restructuring DNA molecules. In the 50s and 60s, McClintock (1950, 1951, 1956, 1965, 1967) pointed out the role that rearrangements play in creating new patterns of genetic functioning, and the idea of modular construction of genomes was suggested in the 70s based on early molecular observations of genome reorganization by mobile genetic elements (Shapiro *et al.*, 1977; Shapiro, 1977).

In other words, it can be argued that much of genome change in evolution results from a genetic engineering process utilizing the biochemical systems for mobilizing and reorganizing DNA structures present in living cells. This paper will explore that proposition further by reviewing what those biochemical systems are and discussing examples of how they are used to solve adaptive problems. Only enough information on each system to illus-

trate the theoretical point being made will be given here; the reviews listed in the bibliography provide a more thorough description.

Two important concerns that will be addressed are (a) control over the genetic engineering process - i.e. how responsive are systems acting on DNA to signals from outside the cell in which changes occur - and (b) coordination of multiple changes throughout the genome. One major tenet of conventional evolutionary theory is that genetic changes arise sporadically withouth reference to the needs of the organism. As we shall see, the molecular genetic data suggest a different point of view. Examples are accumulating of systems where DNA rearrangements occur in a predictable, biologically significant manner. Moreover, there are several systems where coordinated genome-wide rearrangements are integrated into developmental cycles. Thus, one of the fundamental genetic facts underlying any modern evolutionary theory will have to be cellular capacity for specific and widespread genomic change. At the end of the paper, a contemporary example of genetic engineering at work in evolution will be discussed: the emergence of transmissible resistance determinants as the bacterial response to antibiotic chemotherapy. An appendix will discuss the role of accidental DNA change resulting from replication errors and physico-chemical damage in the context of what we have learned about repair and proofreading systems which anticipate these accidents.

Assembling genome components into larger structures: the tools for cutting and splicing DNA

The kind of genetic engineering that is practiced in research laboratories and biotechnology companies depends upon reagents developed by several decades of research on DNA biochemistry. Virtually all the methods used by molecular geneticists employ enzymes and systems extracted from living cells: nucleases, ligases, polymerases, vectors, packaging extracts, etc. (see the catalogues of biotech firms). The one apparent exception is synthetic oligonucleotide technology based on organic chemistry methods, but even this process has its enzymatic parallel in terminal transferase activity (see the discussion of immune system rearrangements below). It is useful to bear in mind that none of the biochemical activities which work on DNA were known when the prevailing evolutionary theories were formulated (i.e. before the structure of DNA was elucidated). Some of them may have been anticipated without challenging prior notions, such as DNA polymerase and homologous recombination systems. Others, however, were truly unimaginable, and the surprise which has frequently greeted the discovery of specific activities (e.g. reverse transcriptase) testifies to their theoretical as well as practical significance.

Because the basic enzymes of human genetic engineering come from living cells, we may assume that what is possible in the eppendorf tube is also possible in the organism, and molecular genetic studies have documented numerous instances of biologically significant DNA rearrangements (Shapiro, 1983; Berg & Howe, 1989). One generalization which can be based on the study of genetic variation is that there exist a wide array of elaborate biochemical systems for replicating, correcting, repairing, packaging, rearranging and transporting DNA molecules. In other words, living cells have many mechanisms at their disposal for processing DNA according to their needs. Just as we have come to know about the wide range of specificities associated with the nucleases used for *in vitro* genetic engineering, the *in vivo* systems are highly differentiated in how they execute their tasks. Some of these are very specific and normally restricted in their action to certain signals; these systems carry out predictable events in the cell cycle or life history of each organism, such as the use of site-specific inversion systems to regulate gene expression in bacteria and 2μ plasmid replication in yeast (Sadowski, 1986; Glasgow *et al.*, 1989). Other DNA processing systems act with fewer sequence restrictions, and there are many cases where the specificity of a DNA processing system can be related to its utility. DNA uptake in *Neisseria* is highly sequence-specific, thereby restricting normal transformation to DNA from other *Neisseria* cells; this makes sense because intraspecies transformation is utilized in *Neisseria* populations for variation of pilus structure and surface proteins (Goodman & Scocca, 1988; Seifert *et al.*, 1989; Gibbs *et al.*, 1989; Scocca, 1990). V-J and V-D-J joining in the mammalian immune system occurs near certain recombination signals but is flexible in

the exact positions of strand breaks; this flexibility permits a high degree of protein sequence variation in antigen-binding regions of the immunoglobulin molecules (Blackwell & Alt, 1989). The insertion of many transposable elements in bacteria displays a low degree of sequence specificity; thus, IS elements can be utilized for mutagenesis in rare and exceptional circumstances, such as insertions to create new promoters allowing *E. coli* cells to transcribe otherwise inactive coding sequences (Iida *et al.*, 1983; Galas & Chandler, 1989).

Process control: response of DNA rearrangement systems to regulatory inputs

Looking at DNA changes as a biochemical process makes it possible to investigate and understand how genetic variation can be regulated. In addition to identifying systems for reorganizing DNA molecules and their biochemical mechanisms, studies of the mutational process have revealed a wealth of control phenomena that operate at many levels. The evidence for these regulatory systems has been obtained at the molecular level, by genetic studies, and by observing the developmental specificity of DNA changes.

Molecular studies of specific systems

The detailed analysis of individual mobile genetic elements (temperate bacteriophages, transposons, retrotransposons) and of the SOS repair/mutator system have demonstrated control mechanisms that operate at all stages of gene expression and directly on the functions of proteins already synthesized. In *E. coli*, transcription of phages λ and Mu, of Tn3 and related transposons, and of the SOS regulon is controlled in each case by a well-characterized repressor molecule (Shapiro, 1983; Berg & Howe, 1989; Walker, 1987). Transcription of the Ty family of retrotransposons in *Saccharomyces cerevisiae* is controlled by the MAT locus transcription factors (Errede *et al.*, 1981; Boeke & Corces, 1989). Germ line-specific expression of P factor transposase is determined by germ line-specific splicing of the last intron from transposase message (Engel, 1989). Expression of the Tn10 transposase molecule is regulated both at the transcriptional level by promoter methylation and at the translational level by anti-sense RNA (Kleckner, 1989). The activities of Tn5 and Tn10 termini as transposase substrates can be modulated by methylation (Berg, 1989; Kleckner, 1989), and methylation also controls the activity states of several transposable elements in maize (Chandler & Walbott, 1986). The activities of Tn5 and Mu transposases are subject to direct inhibition by interactions with other proteins (Adzumi & Mizuuchi, 1988; Berg, 1989). The UmuCD mutator activity of the SOS system requires proteolytic activation by the RecA function (Walker, 1987).

These molecular examples illustrate two essential features of the control of DNA change: (a) regulation occurs at multiple levels, which means that it can operate in a complex, non-linear fashion, and (b) specific mechanisms (e.g. RecA proteolysis, methylation, MAT regulation) integrate different individual DNA reorganization systems into multivalent cellular control networks (Gottesman, 1984), so that these systems do not act independently of biological inputs.

Genetic studies

In some cases, we know that there is regulation from genetic analysis even though the molecular components have not yet been identified. Two good examples are the 'directed mutation' phenomenon in bacteria and hybrid dysgenesis in *Drosophila*.

(a) There have been numerous published and unpublished reports that prolonged incubation of bacteria under selective conditions triggers mutagenic processes ('directed mutation') that allow the formation of mutant clones capable of growth (Shapiro, 1984; Cairns *et al.*, 1988; Hall, 1988). The results from various systems are quite consistent in showing that the frequencies of mutational events (base substitutions, frameshifts, excisions, fusions) increase by orders of magnitude under selection or related kinds of stress (Mittler & Lenski, 1990). However, little information is yet available on signal transduction pathways operating between the specific stresses induced by selection and activation of DNA reorganization activities (Shapiro & Leach, 1990).

(b) In hybrid dysgenesis, transposable elements become active at high levels in hybrids formed by mating a male from a population harboring active elements with a female from a population free of

active elements. In some cases, more than one transposition event will occur in the germ-line cells of dysgenic individuals (Bregliano & Kidwell, 1983; Engels, 1989). This phenomenon has been observed for elements thought to be active at the DNA level (P factors, Hobo), and for the LINE class of retrotransposons (Finnegan, 1989; see also paper by Di Franco, Galappi and Junakovic, this volume).

Developmental studies

Chromosome changes during development have been known for over a century from light microscopy (Boveri, 1887). Since many of these changes involve the loss of whole chromosomes or chromosome fragments, they affect the underlying structure of the genome. The conclusions from microscopic observations have been confirmed in many cases by molecular studies. The changes are quite precise and reproducible, which means that they involve recognition of specific signals in the DNA, and the timings of the changes are also precise and reproducible, which means that there is regulation of the relevant biochemical activities. Three examples will illustrate these points.

(a) In many invertebrates, embryonic development is accompanied by a process called 'chromatin diminution' in which chromosomes or chromosome regions are deleted. One of the best documented examples involves Copepods in a group of three sibling species of *Cyclops* (Beerman, 1977; Beerman & Meyer, 1980). The three species are morphologically indistinguishable but can be discriminated cytologically. Each species contains large amounts of heterochromatin that is excised from the chromosomes of somatic cells (but not from pre-germ cells) during early development. The species differ in the locations of this heterochromatin and in the timing of excision. Two species (*C. divulsus, C. furcifer*) have large heterochromatin blocks at the ends of their chromosomes, while the other (*C. strenuus*) has smaller blocks interspersed at multiple locations along the chromosomes. In *C. divulsus*, excision occurs at divisions 5 and 6; in *C. furcifer* at divisions 6 and 7; and in *C. strenuus*, circular heterochromatin blocks are excised at divisions 4 and 5. Thus, at least one aspect of genomic differentiation during species formation in the *Cyclops* group affected both the location of repetitive DNA in chromosomes and the control over its developmental removal from somatic nuclei.

(b) Cells of ciliated protozoa have two kinds of nuclei (Gall, 1986). A germ line *micronucleus* contains the entire genome organized into typical eukaryotic chromosomes. The micronucleus is transcriptionally inactive but undergoes mitosis during vegetative growth and meiosis during gametogenesis. A somatic *macronucleus* contains the protein-coding regions of the genome organized into a very large number of highly polyploid, small minichromosomes which are capped by telomeres but contain no known centromeric apparatus. The macronucleus is transcriptionally active and determines the cellular phenotype. Many aspects of miniichromosome replication and segregation during cell division are not understood. During conjugation, the old macronucleus degenerates and a new macronucleus is formed from the zygote micronucleus following exchange of gametes and fertilization. This process involves endoreplication and polytenization of micronuclear chromosomes, massive excisions of non-coding DNA, rearrangements of coding information, and addition of new telomeres at the ends of the minichromosomes (Yao, 1989). We will return to this case later when we consider the potential for rapid large-scale genomic change.

(c) In order to achieve the synthesis of immunoglobulins capable of recognizing a virtually infinite range of antigens with high specificity, the mammalian immune system employs a cascade of DNA changes which occur only in the appropriate cells at the appropriate times in lymphoid development (Blackwell & Alt, 1989). One of the most important characteristics of the immunoglobulin rearrangements is that several different DNA reorganization systems are integrated functionally and developmentally into a single system:

(i) V (variable), J (joining) and D (diversity) regions from the germline chromosomes are connected in pre-B cells to form intact expression units for the light and heavy chains of the immunoglobulin molecule. These connections share a common mechanism that recognizes specific pairs of recombination signals built up of heptamer and nonamer repeats with defined spacings. The actual recombination events occur near these signals but are flexible as to the exact internucleotide positions where

strand cleavages and rejoinings occur, so that one pair of V and J regions, for example, can join at various base-pairs to generate a range of different coding sequences. Since the recombination sites occur at the amino-terminal antigen-binding domains of the immunoglobulin chains, this recombinational flexibility greatly enhances the possible repertoire of antibody specificities.

(ii) Following V-D-J recombination to assemble recombinant heavy chain coding units, novel oligonucleotide sequences (N regions) are frequently found on either side of the germline-derived D sequence. These short N region sequences have no germline equivalents and are thought to arise during the recombination process by untemplated DNA synthesis involving terminal transferase activity (Gough, 1983). If correct, this could be considered a cellular equivalent to the use of synthetic oligonucleotides in laboratory genetic engineering.

(iii)During the maturation of the humoral immune response, a different type of recombinational event occurs in one of the introns to exchange one series of exons encoding the heavy chain carboxy-terminal domains for another. Since the carboxy-terminal domains determine antibody class type (μ, α, γ) but not antigen specificity, this recombination event is called class switching, and it employs a very different kind of recognition signal from V-D-J joining. It is significant, however, that the class switch signals are also composed of repeated DNA sequence elements.

(iv) Also during the maturation of the humoral response, the affinity of a particular antibody molecule for its cognate antigen tends to increase with the accumulation of base substitutions in the region encoding the antigen-binding site ('somatic hypermutation'). This process is remarkable because it only takes place in mature antibody-producing B cells and because the base substitutions only occur within a limited region of the immunoglobulin coding sequence. The mechanism of somatic hypermutation is unknown; it may involve a directed kind of gene conversion similar to mating-type switching in yeast (Klar, 1989) and surface protein variation in *Neisseria gonnorrhoeae* (Meyer, 1987; Swanson & Koomey, 1989).

These three developmental examples illustrate how sophisticated the ability of cells to control the timing of particular DNA rearrangements can be. Although it may be objected that development and evolution are quite different processes, the point is that evolutionary thinking will have to based on what we know cells can do. Thus, if we see examples of highly sophisticated regulation of DNA changes during development, it cannot be realistic to base evolutionary theories on the concept that such regulation is not possible.

Retooling for a new model: the capacity for genome-wide changes

One of the major issues in evolutionary theory is the question of the relative importance of accumulated micromutations versus systemic macromutations. The conventional wisdom is that mutational events are sporadic, undirected and independent, and a tacit assumption is generally made that each new mutation arises in a single individual or single gamete. Nonetheless, a closer consideration of the P-M system of hybrid dysgenesis in *Drosophila* and of macronuclear development in ciliates will show that multiple mutations can occur in clusters of gametes and that thousands of coordinated genome-wide changes can occur in a single cell generation.

When a *Drosophila* fly develops from a hybrid dysgenic egg, the P factors in the paternal chromosomes become active during mitotic development of the germ line (Bregliano & Kidwell, 1983). Multiple transposition and chromosome rearrangement events frequently occur within a single mitotic cell division. The result of this activity is a cell that carries a number of genetic changes in its genome. Since this cell can undergo one or more mitotic divisions before meiosis, its haploid progeny will constitute a cluster of gametes which may share a constellation of different mutations and chromosome rearrangements. In a single mating, these gametes can give rise to several progeny which likewise share these multiple genetic changes and, consequently, form an interbreeding founder population with a new genomic structure. Multiple progeny carrying such pre-meiotic mutational events have repeatedly been reported. Therefore, pre-meiotic events can give rise to groups of individuals capable of propagating large-scale genetic changes which would be lost if they appeared in a single individual of a population. In plants, of course, the germ-line develops from somatic tissues in which

mutations can occur long before meiosis, and certain kinds of stresses are known to be capable of activating systems that induce rapid genome reorganization in plant cells (McClintock, 1978; Peschke *et al.*, 1987).

The case of macronuclear development in ciliates is especially illuminating because it demonstrates how distinct DNA rearrangement activities acting at many locations can reorganize the entire genome within a single cell division. The steps of DNA reorganization in macronuclear development are known to include the following (Gall, 1986; Yao, 1989):

- endoreplication and polytenization of the chromosomes in one of two sister micronuclei;
- enclosure of discrete polytene chromosome domains within vesicles;
- fragmentation of the polytene chromosomes at precise locations, some of these at the termini of transposable elements;
- rejoining of subgenic fragments to reconstitute intact coding units (Greslin *et al.*, 1989);
- addition of telomeres to the ends of each fragment by telomerase, a specialized reverse transcriptase (Blackburn, 1991).

What is remarkable about macronuclear development is that so many different events can take place reliably at thousands of distinct sites within the genome. Some system must be operating in the ciliates to keep the different DNA fragments ordered so that they rejoin appropriately. Given this example, the possibility of rapid, massive fragmentation and rearrangement of the genome can no longer be dismissed as fanciful. Thus, it is possible to envisage similar kinds of genome-wide changes to explain the origin of taxonomic differences in the distribution of dispersed sequence repeats, such as the *Cyclops strenuus* heterochromatin or SINE and LINE elements in various mammalian groups (Deininger, 1989; Hutchinson *et al.*, 1989). The same kind of process could also be invoked to account for the creation of coordinated multi-gene systems by insertion of copies of a specific transcription signal in the 5′ regulatory regions of a number of genetic loci.

An evolutionary case history: bacterial resistance to antibiotic chemotherapy

Since World War II, a major evolutionary event has taken place – the emergence of multiply drug resistant bacteria in response to widespread antibiotic use for chemotherapy and prophylaxis. Since the molecular details of this process have been extensively characterized, we have a unique opportunity to examine how well the facts fit with different theoretical perspectives.

The first major discovery about the bacterial response to antibiotic challenge was the report of transmissible resistance factors in the late '50s (Watanabe, 1963). This was a great surprise at the time; the virtually universal expectation was that resistant bacteria would arise in nature as they were observed to do in the laboratory, by chromosomal mutations modifying antibiotic uptake or altering the targets of antibiotic action. Instead, a whole new class of genetic elements was revealed and found to be linked to infectious agents: plasmids and bacteriophages. Resistance determinants did not arise (as predicted) by gradual modifications of pre-existing cellular machinery but appeared as complete systems that could spread rapidly between distinct clones and species.

Work on the physiology of resistance has revealed a variety of sophisticated biochemical mechanisms for dealing with antibiotics (Foster, 1983):

- inactivation by hydrolysis, acetylation, adenylylation, and phosphorylation;
- removal from the cytoplasm by efflux pumps and volatilization;
- covalent modification of target molecules to render them drug-insensitive without altering their activities;
- substitution of drug-insensitive enzymes for their normal (sensitive) analogues.

How these different resistance determinants originally arose is itself an intriguing evolutionary question. For the present discussion, however, what is important is the observation that evolving bacterial cells incorporated DNA sequences encoding novel biochemical systems; they did not use the slower and less efficient process of accumulating mutations altering basic cellular functions.

Studies of how resistance determinants spread have provided much of the information we now possess about two major agents of *in vivo* prokaryotic genetic engineering: plasmids and transposons. These elements are themselves complex mosaics employing multiple biochemical activities for appropriate DNA processing. Plasmids encode intricate systems for their maintenance in clonal populations. These systems include determinants for replication initiation, copy number control, partitioning, and (in some cases) control of cell division so that all daughter bacteria in a population contain one or more copies of the plasmid (Helinski *et al.*, 1985). In addition, virtually all plasmids also encode systems ensuring their transfer to plasmid-free cells (Willets & Wilkins, 1984). For the so-called conjugative or self-transmissible plasmids, these systems are complete and the presence of the plasmid is sufficient for its own transfer. Typically, conjugative plasmids encode surface structures needed to attach to recipient cells, enzymes for initiating DNA replication to mobilize DNA into the recipient cell, and various regulatory activities to control the mating cycle. Non-conjugative plasmids do not encode a complete transfer system, but they do have sequences for the proteins and DNA signals needed to parasitize the transfer apparatus of one or more self-transmissible plasmids. In essence, resistance plasmids are very sophisticated biochemical machines for maintaining and spreading genetic determinants in bacterial populations.

One of the major ways that antibiotic resistance determinants came to be associated with plasmids was through the activities of transposable elements. The argument has already been made that these elements are basically tools for genetic engineering (Shapiro *et al.*, 1977; Shapiro, 1977). Not only can transposable elements move themselves from one genomic location to another, often between separate DNA molecules; they can also join two unrelated DNA molecules. There is little doubt that transposable elements have played a major role in the proliferation of antibiotic resistance in nature: identical drug resistance transposons have repeatedly been identified in different plasmids and different species as a particular resistance mechanism has spread around the world (Heffron, 1983). Like plasmids, transposable elements are complex systems, generally encoding several biochemical activities and regulatory functions. In some cases (the so-called composite transposons like Tn5 and Tn10), the mechanism for incorporation of resistance determinants into the transposable unit appears to depend upon the basic transposase function of the terminal IS elements. Interestingly, other transposons with a different basic structure (e.g. members of the Tn21 family) appear to utilize a separate non-homologous recombination mechanism (TnpI) to incorporate DNA segments encoding individual resistance mechanisms at a specific site within the transposon (Mercier *et al.*, 1990). Thus, transposons illustrate evolutionary change by genetic engineering at two levels: first, in their ability to distribute resistance markers to different molecules, and secondly, in their own internal evolution.

Conclusions: Thinking about evolution in a more sophisticated way

Molecular genetic results have tremendously expanded our understanding of what living cells can do with their genomes. The examples described above illustrate some of the many ways that different biochemical systems serve to restructure DNA molecules in organisms as diverse as bacteria and mammals. These DNA reorganization systems are subject to cellular regulation, and some of them serve specific adaptative functions in organismal life cycles. It is easier to understand how genetic change can be regulated and used to meet adaptive needs if we think of it as a biochemical process rather than as a blind consequence of physicochemical damage. Such damage does occur, of course, but it is anticipated, and the contribution of purely chemical events to genetic change is kept at a very low level by elaborate repair systems. The Appendix gives a brief summary of how bacterial cells deal with accidental genetic changes. Biochemical systems for proofreading and repair constitute an integral part of natural genetic engineering systems; our knowledge of how efficiently they operate is central to understanding the sources of genetic variation in evolution.

As we have seen, the facts established by molecular genetic studies contradict many of the standard exclusionary arguments used in support of conventional evolutionary theories. Through cellular regulatory systems, genetic change can be linked to

adaptive challenges and can be induced to occur at very high frequency under particular circumstances. The integrated operation of DNA reorganization activities can rapidly bring about widespread change throughout the genome. Interbreeding populations of individuals sharing major genomic changes can arise by a well-known series of events. In certain well-documented cases, DNA reorganizations can occur throughout the genome, and the mitotic development of mutant pre-gametic cells can lead to the formation of small populations with extensively restructured genomes. The examination of bacterial antibiotic resistance as a real-world case-study of evolutionary change supports the contention that natural genetic engineering systems play a key role in nature. Thus, the thinking of evolutionary theorists about conceivable mechanisms of genetic variation should be freed from restrictions imposed when knowledge of genetic mechanisms and DNA biochemistry was still rudimentary.

There are major limits to our knowledge of the informational connections between cellular and organismal sensory inputs and the biochemical mechanisms of cellular genetic engineering. Nonetheless, we already have sufficient evidence to know that such connections exist, and it is hard to imagine that they have not played a significant role in episodes of genome reorganization during evolution. If it is true that genetic engineering has played an important role in genome evolution, then we will need to understand how genome integrity is maintained during episodes of massive reorganization and how biologically appropriate structures result. These questions may seem strange now. However, the recent history of molecular genetics has contained many surprises, such as alternative splicing of transcripts from a single locus (Smith *et al.*, 1989) and retroposons (Rogers, 1985; Weiner *et al.*, 1986; Boeke & Corces, 1989; Brosius, 1991). Thus, we can confidently look forward to additional astonishing lessons about the integration of genomic change into basic life processes.

Appendix

Quality control: preventing accidental genetic change by proofreading and repair systems

Before we learned about biochemical systems for DNA reorganization, physical damage from radiation, chemical damage from alkylating agents and other reactive substances, spontaneous chemical changes, and inevitable errors in the fidelity of replication were thought to be the main agents of genetic change. It is very important to recognize that living cells resemble man-made systems for information processing and communication in their use of mechanisms for error detection and correction. These mechanisms minimize the degradation of information through accidents which are inevitable in all complex systems. Cells illustrate the application of this principle with multiple genomic proofreading and repair systems that anticipate a wide range of accidental genetic changes, including replication errors, chemical changes and radiation damage. Because it has been so extensively investigated, *E. coli* is our best example (Kushner, 1987), and some of the known repair systems are listed:

Proofreading replication errors. The first proofreading level comprises an editing 3′-5′ exonuclease activity. In DNA polymerase III, the major replicative enzyme, this activity is provided by a distinct subunit encoded by the *dnaQ* cistron, also known as *mutD*. Mutants defective in *dnaQ* exhibit a very strong mutator phenotype, displaying mutation frequencies up to four orders of magnitude higher than wild-type.

Methyl-directed mismatch repair. A second postreplicational proofreading level involves the methyl-directed mismatch repair system (Modrich, 1987). This system includes a protein that can recognize mismatched DNA sequences. If a mismatch is present in newly replicated DNA, distinguished by hemimethylated GATC sequences, the protein bound to the mismatch stimulates another protein to cleave the unmethylated (i.e. newly synthesized) strand. Excision of the cleaved strand past the mismatch allows a polymerase to replace the mismatched region with a faithful copy of the parental information.

Spontaneous chemical change. Cytosine spontaneously deaminates to uracil, which base-pairs in the same way as thymine. If this were to occur and the uracil persisted during replication, one daughter duplex would contain a GC $\rightarrow$ AU base-pair substitution. *E. coli* prevents such mutations from occurring by the presence of an enzyme, uracil-N-glycosidase (Ung), which specifically excises uracil from DNA. The empty site then triggers a nearby cleavage, and excision patch repair follows in the same

way as described for mismatches in newly replicated DNA. Some cytosines in *E. coli* are methylated, and these methylcytosines deaminate to thymine, which is not recognized by Ung. However, there is another system, called the very short patch (VSP) repair system, which removes the products of methylcytosine deamination (Marinus, 1989).

Alkylation. Alkylation damage at a low level can be removed by a constitutive methyl-transferase activity, but there is also an inducible system to remove higher levels of damage that might occur when the bacterium comes into an environment rich in akylating agents.

Radiation damage. Research has defined a complex set of cellular responses to UV radiation damage known as the SOS response (Walker, 1989). Three aspects of SOS are especially interesting from a theoretical point of view. (1) SOS is an inducible system. The major regulatory protein is RecA, which serves as a sensory function to monitor the genome; it is activated to derepress the SOS system when replication is inhibited or by-products of DNA damage accumulate. Thus, SOS action is purposeful and anticipatory, because the system responds to appropriate signals and is not active when it is not needed. (2) The SOS system is multivalent and includes a variety of integrated biochemical activities which not only correct DNA damage but also affect cell physiology and division in such a way that daughter cells are not produced until the genome has been repaired. (3) UV mutagenesis occurs as a result of the SOS response, not as an inevitable consequence of photochemical damage to DNA. We know this because mutant strains deficient in either RecA or UmuCD (the SOS mutator function) show no increase in mutant frequencies among survivors following UV irradiation. It is interesting to note that the mutagenic response in the Ames test for genotoxic chemicals also depends upon SOS activation (McCann *et al.*, 1975). This means that many examples of induced mutation are, like most spontaneous mutations, the results of cellular DNA biochemistry.

Acknowledgements

I am grateful to Dan Gottschling, Mark Hochstrasser, Lucia Rothman-Denes and Dan Hartl for their comments on the manuscript. I owe an inestimable debt to Barbara McClintock who opened my eyes to the variety and beauty of many different genetic systems. Research in my laboratory is supported by grants from the National Science Foundation (DMB-8715935 and DCB-8816274).

References

Adzumi, A. & K. Mizuuchi, 1988. Target immunity of Mu transposition reflects a differential distribution of Mu B protein. Cell 53: 257-266.

Akam, M., 1989. *Hox* and HOM: homologous gene clusters in insects and vertebrates. Cell 57: 347-349.

Alt, F. W., T. K. Blackwell & G.D. Yancopoulos, 1987. Development of the primary antibody repertoire. Science 238: 1079-1087.

Baker, M. E. & M. H. Saier, Jr., 1990. A common ancestor for bovine lens fiber major intrinsic protein, soybean nodulin-26 protein, and *E. coli* glycerol facilitator. Cell 60: 185-186.

Beerman, S., 1977. The diminution of heterochromatic chromosomal segments in *Cyclops* (Crustacea, Copepoda). Chromosoma 60: 297-344.

Beerman, S. & G. F. Meyer, 1980. Chromatin rings as products of chromatin diminution in *Cyclops*. Chromosoma 77: 277-283.

Berg, D. E., 1989. Transposon Tn*5*. pp. 185-210 in Mobile DNA (D.E. Berg and M. M. Howe, eds) American Society for Microbiology.

Berg, D. E. & M. M. Howe, 1989. Mobile DNA. American Society for Microbiology, Washington, D.C.

Blackburn, E. H., 1991. Structure and function of telomeres. Nature 350: 569-573.

Blackburn, E. H. & J. W. Szostak, 1984. The molecular structure of centromeres and telomeres. Ann. Rev. Biochem. 53: 163-194.

Blackwell, T. K. & F. W. Alt, 1989. Mechanism and developmental program of immunoglobulin gene rearrangement in mammals. Ann. Rev. Genet. 23: 605-636.

Blake, C. C. F., 1985. Exons and the evolution of proteins. Int. Rev. Cytol. 93: 149-185.

Boeke, J. D. & V. Corces, 1989. Transcription and reverse transcription of retrotransposons. Ann. Rev. Microbiol. 43: 403-434.

Boveri, T., 1887. Über Differenzierung der Zellkerne während der Furchung des Eies von Ascaris megalocephala. Anat. Anz. 2: 688-693.

Bregliano, J.-C., & M. G. Kidwell, 1983. Hybrid dysgenesis determinants. pp. 363-410 in Mobile Genetic Elements (J. A. Shapiro, ed.) Academic Press.

Brosius, J., 1991. Retroposons-Seeds of evolution. Science 251: 753.

Cairns, J., J. Overbaugh & S. Miller, 1988. The origin of mutants. Nature 335: 142-145.

Casadesús, J. & J. Roth, 1989. Absence of insertions among spontaneous mutants of *Salmonella typhimurium*. Molec. Gen. Genet. 216: 210-216.

Chandler, V. & V. Walbott, 1986. DNA modification of a maize

transposable element correlated with loss of activity. Proc. Nat. Acad. Sci. USA 83: 1767-1771.

Deininger, P. L., 1989. SINES: Short interspersed repeated DNA elements in higher eucaryotes. pp. 619-636 in Mobile DNA (D. E. Berg & M. M. Howe, eds) American Society for Microbiology.

Donehower, L. & D. Gillespie, 1979. Restriction site periodicities in highly repetitive DNA of primates. J. Mol. Biol. 134: 805-834.

Dover, G. A., 1982. Molecular drive: a cohesive mode of species evolution. Nature 299: 111-117.

Dowsett, A. P., 1983. Closely related species of *Drosophila* can contain different libraries of middle repetitive DNA sequences. Chromosoma 88: 104-108.

Engels, W. R., 1989. P elements in *Drosophila melanogaster*. pp. 437-484 in Mobile DNA (D. E. Berg & M. M. Howe, eds) American Society for Microbiology.

Errede, B., T. S. Cardillo, G. Wever & F. Sherman, 1981. ROAM mutations causing increased expression of yeast genes: Their activation by signals directed toward conjugation functions and their formation by insertions of Ty1 repetitive elements. Cold Spr. Harb. Symp. Quant. Biol. 45: 593-607.

Finnegan, D. J., 1989. The I factor and I-R Hybrid dysgenesis in *Drosophila melanogaster*, pp. 503-518 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Foster, T. J., 1983. Plasmid-determined resistance to antimicrobial drugs and toxic metal ions in bacteria. Microbiol. Rev. 47: 361-409.

Galas, D. & M. Chandler, 1989. Bacterial insertion sequences. pp. 109-162 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Gall, J., 1986. The Molecular Biology of Ciliated Protozoa. Academic Press, Orlando.

Gibbs, C. P., B.-Y. Reimann, E. Schultz, A. Kaufmann, R. Hass & T. F. Meyer, 1989. Reassortment of pilin genes in *Neisseria gonorrhoeae* occurs by two distinct mechanisms. Nature 338: 651-652.

Gilbert, W., 1978. Why genes in pieces? Nature 271: 501.

Glasgow, A. C., K. T. Hughes & M. I. Simon, 1989. Bacterial DNA inversion systems, pp. 637-660 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Goff, S. A., T. M. Klein, B. A. Roth, M. E. Fromm, K. C. Cone, J. P. Radicella & V. L. Chandler, 1990. Transactivation of anthocyanin biosynthetic genes following transfer of B regulatory genes into maize tissues. EMBO J. 9: 2517-2522.

Gottesman, S., 1984. Bacterial regulation: global regulatory networks. Ann. Rev. Genet. 18: 415-441.

Goodman, S. D. & J. J. Scocca, 1988. Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. Proc. Nat. Acad. Sci. USA 85: 6982-6986.

Gough, N., 1983. Has terminal transferase finally made it as a mutator of antibody genes? Trends Biochem. Sci.: 227-228.

Greslin, A. F., D. M. Prescott, Y. Oka, S. H. Loukin & J. C. Chappell, 1989. Reordering of nine exons is necessary to form a functional actin gene in *Oxytricha nova*. Proc. Nat. Acad. Sci. USA 86: 6264-6268.

Hall, B. G., 1988. Adaptive evolution that requires multiple spontaneous mutations. I. Mutations involving an insertion sequence. Genetics 120: 887-897.

Heffron, F., 1983. Tn*3* and its relatives, pp. 223-260 in Mobile Genetic Elements, edited by J. A. Shapiro. Academic Press.

Helinski, D. R., S. N. Cohen, D. B. Clewell, D. A. Jackson & A. Hollaender, 1985. Plasmids in Bacteria. Plenum, New York.

Highton, P. J., Y. Chang & R. J. Myers, 1990. Evidence for the exchange of segments between genomes during the evolution of lambdoid bacteriophages. Molec. Microbiol. 4: 1329-1340.

Hoopes, B. C. & W. R. McClure, 1987. Strategies in regulation of transcription initiation, pp. 1231-1240, In *Escherichia coli and Salmonella typhimurium* edited by F. C. Neidhardt et al., American Society for Microbiology, Washington.

Hutchison, C. A., S. C. Hardies, D. D. Loeb, W. R. Shehee & M. H. Edgell, 1989. LINES and related retroposons: Long interspersed repeated sequences in the eucaryotic genome, pp. 593-618 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Iida, S., J. Meyer & W. Arber, 1983. Prokaryotic IS Elements, pp. 159-221 in Mobile Genetic Elements, edited by J. A. Shapiro. Academic Press.

Inouye, S., T. Franceschini & M. Inouye, 1983. Structural similarities between the development-specific protein S from a Gram-negative bacterium, *Myxococcus xanthus*, and calmodulin. Proc. Nat. Acad. Sci. USA 80: 6829-6833.

Johnson, P. F. & S. L. McKnight, 1989. Eukaryotic transcriptional regulatory proteins. Ann. Rev. Biochem. 58: 799-839.

Klar, A. J. S., 1989. The interconversion of yeast mating type: *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, pp. 671-692 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Kleckner, N., 1989. Transposon Tn*10*, pp. 227-268 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Kushner, S. R., 1987. DNA repair. In *Escherichia coli and Salmonella typhymurium* (F. C. Neidhardt et al., eds), American Society for Microbiology, 1044-1053.

Laski, F. A., D. C. Rio & G. M. Rubin, 1986. Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. Cell 44: 7-19.

Lee, M. G. & P. Nurse, 1987. Complementation used to clone a human homologue of the fission yeast cell cycle control gene cdc2. Nature 327: 31-35.

Lindquist, S. & E. A. Craig, 1988. The heat-shock proteins. Ann. Rev. Genet. 22: 631-677.

Marinus, M. G., 1987. Methylation of DNA, pp. 697-702 in *Escherichia coli and Salmonella typhimurium*, edited by F. C. Neidhardt *et al.* American Society for Microbiology, Washington.

McCann, J., N. E. Spingarn, J. Kobori & B. N. Ames, 1975. Detection of carcinogens as mutagens: bacterial tester strains with R factor plasmids. Proc. Nat. Acad. Sci. USA 72: 979-983.

McClintock, B., 1950. The origin and behavior of mutable loci in maize. Proc. Nat. Acad. Sci. USA 36: 344-355.

McClintock, B., 1951. Chromosome organization and genic expression. Cold Spr. Harb. Symp. Quant. Biol. 16: 13-47.

McClintock, B., 1956. Controlling elements and the gene. Cold

Spr. Harb. Symp. Quant. Biol. 21: 197-216.

McClintock, B., 1965. The control of gene action in maize. Brookhaven Symp. Biol. 18: 162-184.

McClintock, B., 1967. Genetic systems regulating gene expression during development. Develop. Biol. Suppl. 1: 84-112.

McClintock, B., 1978. Mechanisms that rapidly reorganize the genome. Stadler Symp. 10: 25-47.

McClintock, B., 1984. The significance of responses of the genome to challenge. Science 226: 792-801.

McGinnis, N., M. A. Kuziora & W. McGinnis, 1990. Human Hox-4.2 and Drosophila Deformed encode similar regulatory specificities in Drosophila embryos and larvae. Cell 63: 969-976.

Mercier, J., J. Lachapelle, F. Couture, M. Lafond, G. Vezina, M. Boissinot & R. C. Levesque, 1990. Structural and functional characterization of *tnpI*, a recombinase locus in Tn*21* and related β-lactamase transposons. J. Bacteriol. 172: 3745-3757.

Meyer, T. F. 1987. Molecular basis of surface antigen variation in *Neisseria*. Trends in Genet. 3: 319-324.

Mittler, J. & R. E. Lenski, 1990. Further experiments on excisions of Mu from *Escherichia coli* MCS2 cast doubt on directed mutation hypothesis. Nature 344: 173-175.

Modrich, P., 1987. DNA mismatch correction. Ann. Rev. Biochem. 56: 435-466.

Müller, M. M., T. Gerster & W. Schaffner, 1988. Enhancer sequences and the regulation of gene transcription. Eur. J. Biochem. 176: 485-495.

Peschke, V. M., R. L. Phillips & B. G. Gengenbach, 1987. Discovery of transposable element activity among progeny of tissue culture-derived maize plants. Science 238: 804-807.

Ptashne, M., 1986. A Genetic Switch. Cell/Blackwell, Cambridge MA and Palo Alto.

Riley, M. & S. Krawiec, 1987. Genome organization, pp. 967-981 in *Escherichia coli and Salmonella typhimurium*, edited by F. C. Neidhardt et al. American Society for Microbiology, Washington.

Rogers, J., 1985. The origin and evolution of retroposons. Int. Rev. Cytol. 93: 231-279.

Rosenfeld, M. G., C. K. Glass, S. Adler, E. B. Crenshaw III, X. He, S. A. Lira, H. P. Elsholtz, H. J. Mangalam, J. M. Holloway, C. Nelson, V. R. Albert & H. A. Ingraham, 1989. Response and binding elements for ligand-dependent positive transcription factors integrate positive and negative regulation of gene expression. Cold Spr. Harb. Symp. Quant. Biol. 53: 545-556.

Sadowski, P., 1986. Site-specific recombinases: Changing partners and doing the twist. J. Bacteriol. 165: 341.

Scocca, J. J., 1990. The role of transformation in the variability of the *Neisseria gonorrhoeae* cell surface. Molec. Microbiol. 4: 321-327.

Seifert, H. S., R. A. Ajioka, C. Marchal, P. F. Sparling & M. So, 198. DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. Nature 336: 392-395.

Shapiro, J. A., 1977. DNA insertion elements and the evolution of chromosome primary structure. Trends in Biochem. Sci. 2: 622-627.

Shapiro, J. A., 1983. Mobile Genetic Elements. Academic Press, New York.

Shapiro, J., 1984. Observations on the formation of clones containing *araB-lacZ* cistron fusions. Molec. Gen. Genet. 194: 79-90.

Shapiro, J. A., 1991. Genomes as smart systems. *Genetica* 84: 3-4.

Shapiro, J. A., A. I. Bukhari & Adhya, 1977. New Pathways in the evolution of chromosome structure, pp. 3-13 in DNA Insertion Elements, Plasmids and Episomes, edited by A. I. Bukhari, J. A. Shapiro and S. Adhya. Cold Spring Harbor Press.

Shapiro, J. A. & D. Leach, 1990. Action of a transposable element in coding sequence fusions. Genetics 126: 293-299.

Smith, C. W. J., J. G. Patton & B. Nadal-Ginard, 1989. Alternative splicing in the control of gene expression. Ann. Rev. Genet. 23: 527-577.

Smith, H. O., D. B. Danner & R. A. Reich, 1981. Genetic transformation. Ann. Rev. Biochem. 50: 41-68.

Stahl, F. W., 1979. Special sites in generalized recombination. Ann. Rev. Genet. 13: 7-24.

Stock, J. B., A. M. Stock & J. M. Mottonen, 1990. Signal transduction in bacteria. Nature 344: 395-400.

Swanson, J. & J. M. Koomey, 1989. Mechanisms for variation of pili and outer membrane protein II in *Neisseria gonorrhoeae*, pp. 743-762 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Trifonov, E. N. & V. Brendel, 1986. GNOMIC: A Dictionary of Genetics Codes. Balaban, Philadelphia.

Walker, G. C., 1987. The SOS response of *Escherichia coli*, pp. 1346-1357 in *Escherichia coli and Salmonella typhimurium* edited by F. C. Neidhardt et al. American Society for Microbiology, Washington.

Watanabe, T., 1963. Infective heredity of multiple drug resistance in bacteria. Bacteriol. Rev. 27: 87-115.

Weiner, A. M., P. L. Deininger & A. Efstratiadis. 1986. Nonviral retroposons: Genes, pseudogenes and transposable elements generated by the reverse flow of genetic information. Ann. Rev. Biochem. 55: 631-661.

Willets, N. S. & B. Wilkins, 1984. Processing of DNA during bacterial conjugation. Microbiol. Rev. 48: 24-41.

Yao, M.-C., 1989. Site-specific chromosome breakage and DNA deletion in ciliates, pp. 715-734 in Mobile DNA, edited by D. E. Berg & M. M. Howe. American Society for Microbiology.

Zakian, V., 1989. Structure and function of telomeres. Ann. Rev. Genet. 23: 579-604.