35. H. Bache, *Proceedings of the 2nd International Conference on Superplasticizers in Concrete*, CANMET, Ottawa, Canada (June, 1981), p. 33.
36. L. Hjorth, in (*22*), p. 167.
37. F. F. Radjej, T. Bogen, E. J. Sellevold, R. E. Loeland, in *Proceedings of the First International Conference on the Use of Fly Ash, Silica Fume, Slag and Other Mineral By-Products in Concrete* (meeting sponsored jointly by the Canadian Center for Mineral and Energy Technology and the American Concrete Institute) (*ACI Publ. SP-79*, American Concrete Institute, Detroit, 1983), p. 1135.
38. P.-C. Aitcin, C. Bedard, M. Plumat, G. Haddal, in (*28*), p. 201.
39. S. Diamond, *ibid.*, p. 233.
40. S. Wise *et al.*, *ibid.*, p. 253.
41. J. P. Romualdi and J. A. Mandel, *J. Am. Concr. Inst.* **5**, 657 (1964).
42. D. M. Roy, A. J. Majumdar, S. P. Shah, J. A. Manson, Eds., *Advances in Cement-Matrix Composites, Proceedings of the Materials Research Society*, Pittsburgh, November 1980 (Materials Research Society, University Park, PA, 1980).
43. D. J. Hannant, *Fibre Cements and Fibre Composites* (Wiley, New York, 1978).
44. A. Kelly, in (*42*), p. 3.
45. ———, in (*22*), p. 175.
46. A. J. Majumdar, *Composites* **5**, 7 (January 1975); ——— and J. F. Ryder, *Glass Technol.* **3** (no. 9), 78 (1968); M. A. Ali, A. J. Majumdar, B. Singh, *J. Mater. Sci.* **10**, 1732 (1975).
47. A. J. Majumdar, in (*42*), p. 37.
48. ——— and V. Laws, in (*22*), p. 191.
49. A. M. S. Stucke and A. J. Majumdar, *J. Mater. Sci.* **11**, 1019 (1976).
50. D. J. Hannant, J. J. Zonsveld, D. C. Hughes, *Composites* **9**, 83 (April 1979).
51. K. L. Biryukovich, *Glass-Fibre-Reinforced Cement*, Budivel'nik Kiev (1964) Civil Engineering Research Association, London, 1986).
52. J. Aveston, G. A. Cooper, A. Kelly, in *The Properties of Fibre Composites*, Conference Proceedings of the National Physical Laboratory (U.S.) (NPL) (IPC Science and Technology, Guildford, U.K., 1971), p. 15; J. Aveston, R. A. Mercier, J. M. Silwood, *Composites-Standards Testing and Design*, Conference Proceedings of the NPL (IPC Science and Technology, Guildford, U.K., 1974), p. 91.
53. A. Bentur and S. Diamond, in *Durability of Glass-Fiber Reinforced Concrete Symposium*, Proceedings of the Prestressed Concrete Institute, Chicago, 12 to 15 November 1985, S. Diamond, Ed. (Prestressed Concrete Institute, Chicago, 1986), p. 337; A. Kumar and D. M. Roy, *ibid.*, p. 147; K. D. Litherland and D. R. Proctor, *ibid.*, p. 124; K. L. Litherland, D. R. Oakley, B. A. Proctor, in (*42*), p. 61.
54. D. R. Lankard and D. H. Lease, *Am. Ceram. Soc. Bull.* **61**, 728 (1982).
55. N. McN. Alford and J. D. Birchall, *J. Mater. Sci.* **20**, 37 (1985).
56. J. G. Moore, H. W. Godbee, A. H. Kibbey, *Nucl. Technol.* **32**, 39 (1977).
57. D. M. Roy and G. R. Gouda, *ibid.* **40**, 214 (1978); D. M. Roy, in *Proceedings of the Conference on High-Level Radioactive Solid Waste Forms, Nuclear Regulatory Commission Conference Proceedings, No. 5*, L. Casey, Ed., Denver, 18 to 21 December 1978 (Nuclear Regulatory Commission, Washington, DC, 1979), p. 479.
58. M. W. Barnes, D. M. Roy, C. A. Langton, *Scientific Basis for Nuclear Waste Management VIII*, vol. 44 of *Mater. Res. Soc. Symp. Proc.*, C. M. Jantzen, J. A. Stone, R. C. Ewing, Eds., December 1984 (Materials Research Society, Pittsburgh, 1985), p. 865; W. Fajun, M. W. Grutzeck, D. M. Roy, *ibid.*, p. 899.
59. D. M. Roy, B. E. Scheetz, L. D. Wakeley, M. W. Barnes, *Nucl. Chem. Waste Manage.* **3**, 35 (1982).
60. B. E. Scheetz and D. M. Roy, in (*58*), p. 935.
61. W. A. Ross *et al.*, *Ceram. Bull.* **62**, 1026 (1983); *ibid.*, p. 1035.
62. S. Komarneni and D. M. Roy, *Science* **221**, 647 (1983).
63. M. Perez-Pena, D. M. Roy, A. S. Bhalla, L. E. Cross, *Cem. Concr. Res.* **16**, 951 (1986).
64. J. D. Birchall and A. Kelly, *Sci. Am.* **248**, 104 (May 1983).
65. I appreciate comments on the manuscript by N. H. Macmillan. This research was partially funded under NSF grant DMR-8418160.

Programmed Gene Rearrangements Altering Gene Expression

P. Borst and D. R. Greaves

Programmed gene rearrangements are used in nature to to alter gene copy number (gene amplification and deletion), to create diversity by reassorting gene segments (as in the formation of mammalian immunoglobulin genes), or to control the expression of a set of genes that code for the same function (such as surface antigens). Two major mechanisms for expression control are DNA inversion and DNA transposition. In DNA inversion a DNA segment flips around and is rejoined by site-specific recombination, disconnecting or connecting a gene to sequences required for its expression. In DNA transposition a gene moves into an expression site where it displaces its predecessor by gene conversion. Gene rearrangements altering gene expression have mainly been found in some unicellular organisms. They allow a fraction of the organisms to preadapt to sudden changes in environment, that is, to alter properties such as surface antigens in the absence of an inducing stimulus. The antigenic variation that helps the causative agents of African trypanosomiasis, gonorrhea, and relapsing fever to elude host defense is controlled in this way.

Gene order is a stable characteristic of all organisms analyzed. Each virus, bacterium, plant, and animal has a genetic map, shared by all members of the species. If the β-globin gene is on chromosome 11 in one representative of *Homo sapiens*, it is likely to be there in all. Genes do move around, however. Such DNA rearrangements are of two types, programmed and incidental (also called unprogrammed or mutational) (*1*). The incidental rearrangements arise from errors in DNA replication, repair, or recombination; from the movement of mobile elements, such as transposons; or from the insertion or excision of plasmid DNA, viral DNA, or other immigrant DNA. The exact outcome of these incidental rearrangements is unpredictable, but most of them are deleterious to the individual affected by them.

In contrast, programmed gene rearrangements are part of the normal developmental program of an organism and its progeny. The outcome of these rearrangements is largely predictable; the process is usually carried out by specific recombination enzymes and is developmentally regulated. Such programmed (or developmental) rearrangements can be grouped in three categories: (i) amplification or deletion of genes; (ii) assembly of genes from gene segments; and (iii) DNA rearrangements that alter gene expression.

This review will mainly deal with the third category, the gene

P. Borst is director of research and D. R. Greaves is a postdoctoral fellow at The Netherlands Cancer Institute (Antoni van Leeuwenhoek Huis), Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.

rearrangements that alter the level of expression of genes. As the three categories are related in mechanism and rationale and as the borders between them are sometimes arbitrary, we shall start with a brief discussion of the other two types of programmed gene rearrangements.

Amplification or Deletion of Genes

Programmed amplification or deletion of genes is widespread in nature. Amplification (*2*) may affect most of the DNA, as in the polytenization of chromosomes in insect salivary glands; it may affect a substantial fraction of the genome, as in the formation of macronuclei in ciliates, or only a small set of genes, as in the amplification of the genes for the large ribosomal RNAs (rRNAs) in frog eggs. In most cases the rationale of gene amplification seems clear (*3*): polyploidy allows cells to grow to large size; local gene amplification allows cells to produce a specific gene product in a short period of time.

Although gene amplification leads to the synthesis of more gene product, it is usually not associated with a change in the expression of the amplified genes. Formally, gene amplification is therefore not a gene rearrangement that alters gene expression.

Programmed deletions (*4, 5*), like programmed amplifications, vary in extent. They may be limited to one gene, as in the assembly of immunoglobulin genes, or they may involve the systematic loss of X chromosomes or satellite sequences from chromosomes. Whereas the role of the smaller deletions is in most cases defined, the rationale of the large deletions is often unclear. The loss of X chromosomes may serve to accomplish the same dosage compensation achieved by X-chromosome inactivation in other organisms; why the loss of satellite DNA sequences should be advantageous is not obvious (*5*).

The effects of DNA deletions on gene expression vary. Deletion of the entire gene itself obviously stops expression; more subtle deletions may be used, however, to create functional genes (see below).

Making New Genes: DNA Rearrangements in B and T Cells of Mammals

Rearrangement of immunoglobulin genes in mammalian B cells allows a relatively small number of gene segments to be recombined so as to generate a large number of different antigen-binding specificities (*6*). Diversity is created in two ways: first, joining of gene segments is imprecise and nucleotides may be added or removed at the joint. Second, diversity accrues from combinatory joining of any variable (V) to any diversity (D) to any joining (J) segment. Similar mechanisms are used in T cells to create diversity in the α and β subunits of the T-cell receptor (*7*).

Rearrangements of DNA in T and B cells not only create diversity but also affect gene expression. VDJ joining activates the promoter in front of the V segment by bringing it under the influence of the transcriptional enhancer located between J and C segments. The Cμ-enhancer is activated very early in B-cell differentiation, however, and directs the transcription of adjacent sequences, giving rise to short-lived "sterile" Cμ-transcripts prior to VDJ joining (*8*). Even after joining, the main control over the steady-state level of heavy-chain messenger RNA (mRNA) is at the posttranscriptional level (*9*). The transcriptional activation of the correct V-gene promoter is therefore only a side-effect of the DNA rearrangement and not its raison d'être.

Why do T-cell receptor genes only rearrange in T cells? By means of a functional thymidine kinase gene flanked by D and J segments of an unrearranged T-cell receptor gene, Alt and co-workers (*10*) have shown that correct D to J joining can occur after such a construct has been transfected into a pre-B-cell line. Rearrangement of the endogenous T-cell receptor sequences is not observed. All the enzymic machinery required for rearranging T-cell receptor genes must therefore be present in pre-B cells, but the endogenous genes are not accessible to the recombination enzymes, presumably because they are packaged in a more condensed chromatin structure (*11*).

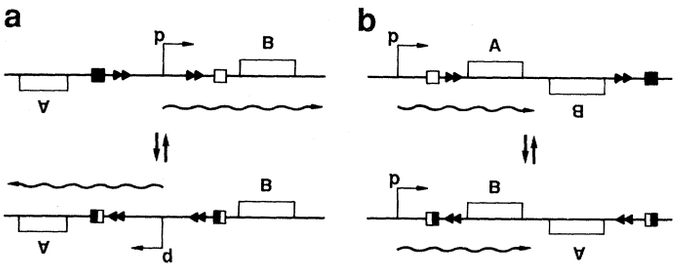

Fig. 1. Inversion control of gene expression. Panel **a** demonstrates the mutually exclusive expression of two structural genes, A and B, by promoter (p→) inversion. Panel **b** illustrates control of expression through gene inversion. Wavy lines indicate mRNA transcripts; white and black squares denote sequences recognized and cleaved by recombinase enzymes.

Although the assembly of genes from gene segments represents a powerful mechanism for creating diversity, the immunoglobulin and T-cell receptor genes are the only proven examples thus far. A possible third example may occur in ciliates. In the massive DNA amplification that occurs during the formation of the macronucleus, segments are deleted from the amplified DNA (*12*). Most of these deletions are invariant (*13*), but some are not, creating genetic heterogeneity within the progeny of a single clone (*14*). The functional significance of this heterogeneity is not yet known, but as there are more than 5000 rearrangement sites in the *Tetrahymena* genome, large variations could potentially be created. Mating-type differentiation in *Tetrahymena* might depend on such differential DNA rearrangements (*15*).

An Overview of DNA Rearrangements Altering Gene Expression

DNA rearrangements altering gene expression can be reversible or irreversible. Two basic mechanisms used in the reversible rearrangements are illustrated in Figs. 1 and 2.

1) Gene expression is controlled by an invertible DNA segment (Fig. 1). There are two versions of this flip-flop control principle as follows. (i) Transcription starts in the invertible segment, and the genes controlled by this promoter are outside the segment. (ii) Transcription starts outside the invertible segment, and the genes controlled by the invertible promoter are inside the segment.

2) Gene expression is activated by the transposition of a copy of the silent gene to an expression site, where it can be transcribed (Fig. 2). In all cases known to us studied thus far, the duplicative transposition occurs by a gene conversion, in which a gene in the expression site is displaced by the incoming gene copy and destroyed. This is also referred to as the cassette model (or principle), the expression site acting as the head of a tape recorder in which gene cassettes can be slotted to be transcribed (*16*). There are also two versions of expression-site control of gene expression. (i) The start of transcription is within the transposed gene (Fig. 2a). In the silent version of the gene, the transcription start is repressed by sequences neighboring the gene. Transposition removes the gene

from its neighbors and leads to derepression. (ii) The transposed gene does not contain a functional transcription start, but transposition places it within an active transcription unit which starts upstream of the transposed gene in the expression site (Fig. 2b).

Inversion control and expression site control differ not only in the number of gene sets that can be controlled (two versus many) but also in the mechanism and enzymic machinery involved. Inversion requires a site-specific recombination, catalyzed by a specific recombinase that is aided by protein factors. Gene conversion may be initiated by a site-specific nuclease, but later steps probably use the enzymic machinery involved in generalized recombination in the cell.

A third mechanism for reversibly altering gene expression has recently been discovered in *Neisseria* bacteria. Addition or subtraction of sequences to a protein-coding region can break or restore the reading frame, resulting in control of gene expression at the translational level. Whether this mechanism is used outside *Neisseria* is not known.

The irreversible DNA rearrangements altering gene expression are diverse. We have already mentioned the rearrangements giving rise to functional immunoglobulin and T-cell receptor genes and the deletions introduced in some ciliate genes during macronucleus formation. A third example is the formation of a functional nitrogen fixation operon in the cyanobacterium *Anabaena* by excision of a DNA segment (*17*). Finally, precisely regulated gene rearrangements result from the site-specific insertion of episomes and proviruses into chromosomal DNA and from the genetic colonization of plants by bacteria, such as *Agrobacterium* (*18*). In none of these examples is the evidence strong that alteration of gene expression is the main function of the DNA rearrangement.

Inversion Control of Gene Expression

Inversion control of gene expression is only found in bacteria and bacterial viruses, and even in these microorganisms it affects only a few genes (*19*). The classical example of the inversion of a promoter-containing segment (mechanism a in Fig. 1) is provided by the phase variation (antigenic variation) of *Salmonella* (*19, 20*). Individual *Salmonella* bacteria can alternate between production of two types of flagellar protein, called H1 and H2. In the (+) phase, the gene for H1 is transcribed together with an adjacent gene that codes for a repressor of the gene for H2, located elsewhere in the genome. In the (−) state the promoter is inverted, the H1 gene and the repressor for the H2 gene are not transcribed, and hence H2 is produced. H1 and H2 differ in antigenic properties, and, since the flagellar protein is the dominant antigen of *Salmonella*, switching improves the chance of survival in a host with an effective immunological defense. Switching occurs at low rate, 10^{-5} to 10^{-3} per bacterial division, and requires a specific recombinase, called Hin.

A genetic switch in which the genes (or part of them) are inverted and the promoter is outside the inverted segment (mechanism b in Fig. 1) is found in bacteriophage mu and related phages (*21*). In this case the phage population can alternate the production of two different sets of tail fibers, which allows the phage to extend its host range (*21*).

DNA sequence comparisons have shown that the two forms of inversion control are closely related (*19, 22*). The enzymic machinery used in inversion has probably been derived from mobile genetic elements, as the site-specific recombinases catalyzing inversion are about 40% homologous to tnpR, an enzyme required for resolution of the cointegrate intermediate in transposon movement (*19, 22*). Inversion is brought about by a homologous recombination between two short identical but inverted sequences that flank the invertible segment and is stimulated by the presence of an additional

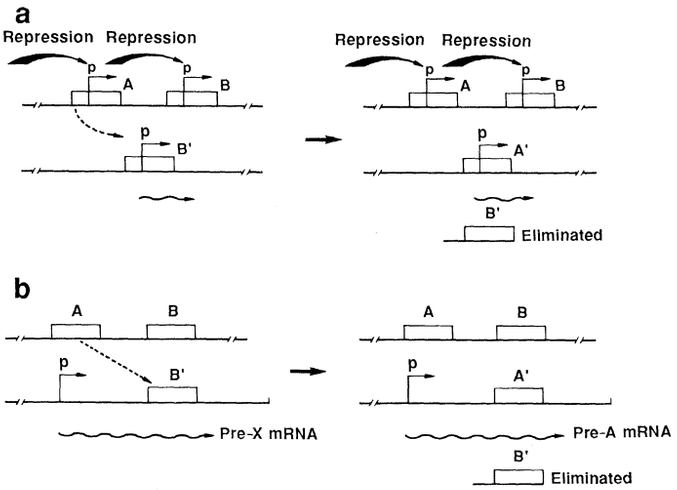

Fig. 2. Expression site control of gene expression. Panal **a** shows mutually exclusive expression of structural genes A and B by the cassette mechanism. Genes A and B have a functional promoter but are not transcribed because of *cis*-mediated repression (represented by curved arrows). Transposition of gene A mediated by gene conversion (dotted arrow) replaces the expressed gene copy B′ with a copy of the silent gene A to give the expressed gene A′. Panel **b** shows control of expression through promoter addition. Duplicative gene conversion generates a second copy of gene A downstream of an active promoter and displaces the previously active gene.

DNA sequence (*20, 21*). The inversion reaction has been reproduced in vitro and requires only the recombinase, a protein factor, and Mg^{2+} (*19–21*).

Invertible DNA segments are not restricted to prokaryotes. They have also have been observed in a circular yeast plasmid (*23*) and in the linear duplex DNA of herpesviruses. The yeast plasmid encodes a site-specific recombinase (*24*) that shows weak homology to a bacterial family of recombinases, exemplified by the integrase that helps to insert lambda prophage into the bacterial host genome (*25*). Inversion does not affect gene expression, but allows plasmid amplification (*26*). The function of the invertible DNA segments of herpes viral DNA is not known (*27*), and the enzyme responsible for inversion has not been identified.

Switching the Mating Type of Yeast

Homothallic strains of the budding yeast *Saccharomyces cerevisiae* grow as haploid cells of either the *a* or α mating type. Cells of opposite mating type can fuse to form *a*/α diploids, which in response to starvation undergo meiosis to produce four haploid spores. The many physiological differences between *a* and α cells and between haploid and diploid yeast cells are determined by the DNA sequences present at a single genetic locus, the mating-type locus *MAT*. Haploid but not diploid cells undergo frequent interconversion of mating type during growth. Switching of mating type can occur in precisely determined, pedigree-dependent fashion as often as once every cell division (*16, 28, 29*).

Figure 3 summarizes the organization of mating-type genes on chromosome III of an α cell. All the information required to encode the α phenotype is present within a 2.5-kb region of the *MAT*α locus, which gives rise to two divergently transcribed mRNAs, α1 and α2. An identical copy of the 2.5 kb of α-specific information is present on chromosome III at *HML*α, but this is not expressed because of *cis*-mediated repression by sequences flanking *HML* (*30*). Sequences encoding *a*1 and *a*2, the gene products encoding the *a* phenotype, reside at the opposite end of chromosome III at the

HMRa locus. The *a*1 and *a*2 promoters at *HMRa* are also transcriptionally inactive as a result of the transcriptional "silencer" sequences flanking *HMR* (*31*).

Switching of yeast mating type is the result of a duplicative gene conversion event (compare Fig. 2a), which replaces the α sequences at *MAT*α with *a* sequences from *HMRa*. The *a*1 and *a*2 promoters thereby escape the *cis*-mediated repression at *HMR* and their expression at *MATa* converts the α yeast cell into an *a* cell. A key observation in understanding the regulation of mating-type interconversion was the detection of a double-strand DNA break within the *MAT* locus in DNA prepared from haploid cells undergoing frequent switching (*32*). This DNA break maps to the YZ junction within the *MAT* locus and is made by a specific endonuclease that recognizes and cleaves a >16-bp sequence near the YZ junction (*33*). Cloned DNAs of *HMR, HML*, and *MAT* are cleaved with equal efficiency in vitro, but in vivo duplex DNA breaks are seen only at *MAT*. This preferential cleavage at *MAT* is the result of a more open chromatin configuration (*34*). Inaccessibility of *HML* and *HMR* chromatin is known to depend on the full expression of four different *SIR* genes (*35*), but the mechanism of *SIR* repression has not yet been determined. Thus, the unidirectional transfer of information from *HML* or *HMR* to *MAT* is the result of the more active chromatin configuration of the expressed cassette that makes *MAT*, but not *HML* or *HMR*, a substrate for the YZ endonuclease in vivo. The double-strand break initiates the elimination of sequences at *MAT* and their replacement by *HML* or *HMR* sequences. This occurs by double-strand gap repair, probably with the enzymes involved in generalized recombination (*36*).

The evolutionary advantage of mating-type interconversion is that it allows a single cell to give rise to progeny that can undergo mating, keeping the haploid phase short. Why yeast cells should have chosen to use the cassette mechanism to regulate the mutually exclusive expression of two sets of genes is not clear. Constant gene rearrangements at *MAT* are bought at a price, aberrant recombinations of mating-type sequences being generated at high frequency (*28, 29, 37*), even though the endonuclease used in mating-type switching is very specific.

Antigenic Variation in Trypanosomes

The African trypanosomes, exemplified by *Trypanosoma brucei*, are unicellular protozoa that can multiply for long periods in the bloodstream of mammalian hosts. The parasite evades the host immune defense by repeatedly changing its surface coat, the only parasite structure exposed to host antibodies (*38*). The switching of coat composition has the hallmarks of a gene rearrangement: it occurs at low rate (10^{-7} to 10^{-6} per trypanosome division) in laboratory strains, and switching is not detectably influenced by external stimuli, such as antibodies (*39*).

The surface coat consists mainly of a single protein species, the variant-specific surface glycoprotein or VSG (*40*). A major route for VSG gene activation involves the duplicative transposition of a silent gene to a telomeric expression site (*41*), where it is transcribed, as schematically indicated in Fig. 4a. There are some 10^3 chromosome-internal silent VSG genes detected by hybridization (*42*), which accounts for the enormous coat repertoire of the trypanosome (*43*). As the switching rate is so low, no intermediates in switching have been identified. Indirect evidence strongly suggests, however, that transposition is the result of a gene conversion that involves short stretches of homology that vary in length (*44*). Upstream of the VSG gene, these consist of a few imperfect 70-bp repeats. Longer stretches of these repeats are present in front of most telomeric VSG genes (*45*). A different homology stretch is present downstream of VSG genes. This homology starts in the region that encodes the COOH-terminal region of the mature protein and extends beyond the mRNA end. The homology between different VSG genes in this area is patchy, one of the longest stretches being a 14mer positioned just before the polyadenylated tail (*44, 46*).

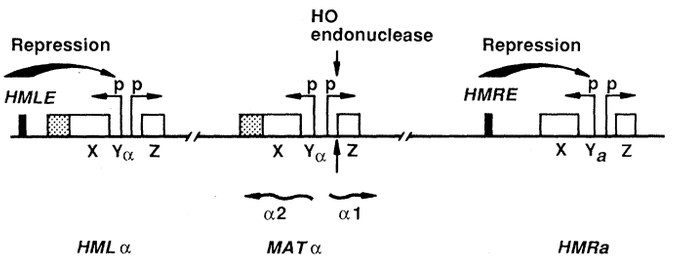

Fig. 3. Schematic representation of the organization of mating-type genes on chromosome III of a *S. cerevisiae* α cell (adapted from *31*). Different blocks of sequence homology between *HML*, *MAT*, and *HMR* are denoted by shaded or open (X and Z) boxes. Yα and Y*a* are nonhomologous sequence blocks containing the α1, α2 and *a*1, *a*2 genes. The position of the double-strand DNA break found at *MAT*α is marked by the vertical arrows. Sequences at *HMRE* are essential for the *cis*-mediated repression of the *a*1 and *a*2 promoters at Y*a*, *HMLE* sequences are essential for *cis*-mediated repression of the α1 and α2 promoters at Yα. The wavy lines indicate mRNAs; as pictured, these are exaggerated in length to make them visible.

These short recombination areas may be involved in a gene conversion process (*47*) as is suggested by several arguments (*46*). First, the resident gene in the expression site is invariably displaced by the incoming gene and lost. Second, the length of the transposed segment varies in different transpositions of the same gene. Upstream of the gene, the crossover may occur at different positions in different 70-bp repeats (*48*); downstream of the gene, the crossover may occur anywhere in the recombination region (*49*). Third, the

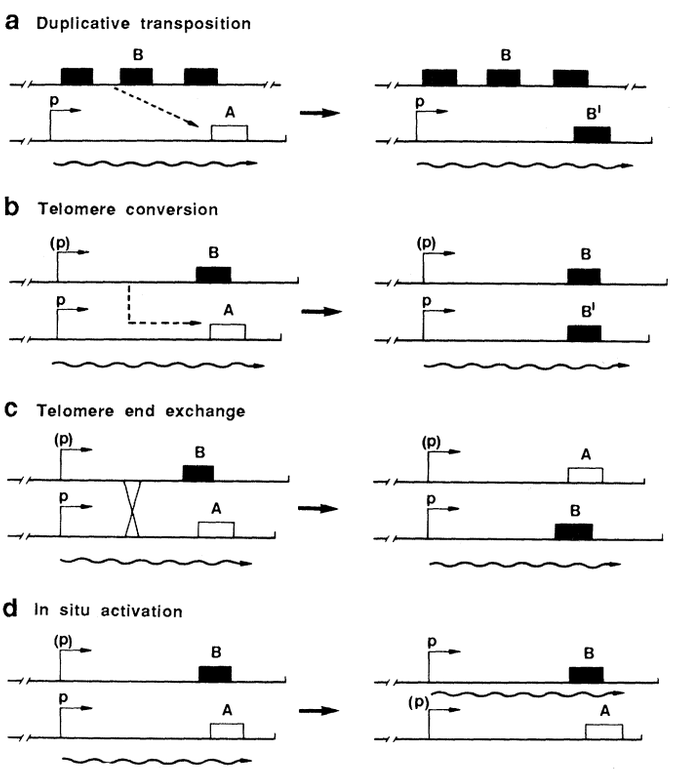

Fig. 4. Four ways in which an inactive VSG gene B can be activated with concomitant inactivation of the previously active gene A (*46*). Inactivated promoter sequences are enclosed by parentheses (p).

transposed segments containing VSG genes are not flanked by (inverted) repeats. Hence, VSG genes seem to transpose by gene conversion of other VSG genes and not as mobile elements that can insert in new locations. The absence of one of the recombination areas in some VSG genes (*42, 50, 51*) may explain why these genes are activated infrequently (*51*).

It is unlikely that transposition involves an RNA intermediate, because silent VSG genes seem to lack a promoter and are not detectably transcribed (*46*). It is also possible that no specific endonuclease is involved, but that VSG gene transposition depends on accidental DNA breaks and is catalyzed by the general recombination machinery of the cell. Whether such gene rearrangements can still be called programmed rearrangements is an interesting question that we shall return to below.

In summary, a major route for the duplicative transposition of VSG genes resembles in mechanism the mating-type switch in yeast, but fundamental differences also exist, as indicated in Table 1. Trypanosomes contain 10^3 silent genes, rather than two; transposition activates the transposed gene by promoter addition (Fig. 2b) rather than by derepression (Fig. 2a); gene conversion is mediated by short, imperfectly homologous segments rather than by large blocks of perfect homology; the switching rate is 10^{-6} to 10^{-7} rather than once per cell division; and switching may not involve a specific endonuclease. The more detailed analysis of VSG gene-switching presented in the next section reveals additional complications further accentuating the difference with the mating-type switch.

The Nonduplicative Activation of VSG Genes

Silent VSG genes that reside at telomeres can be activated in several ways, as indicated in Fig. 4. (i) Activation by a duplicative transposition (Fig. 4b): This transposition differs from the usual transposition of chromosome-internal VSG genes by the size of the segment transposed. This segment may start as far as 50 kb upstream of the gene and end far downstream of the gene and possibly include the whole telomere (*44, 52*). Hence, the name telomere conversion. In occasional transpositions, the DNA segments transferred are much smaller than a complete gene (*44*). Such segmental gene conversions further increase the number of different coats that one trypanosome and its progeny can produce. (ii) Activation without detectable duplication (*53*): This can either occur by a reciprocal recombination, as indicated in Fig. 4c, or by activation in situ (Fig. 4d). Only two examples of reciprocal recombination have been found (*54*), and it is probably rare. The frequent nonduplicative activation of VSG genes early in infection is therefore usually due to activation of a VSG gene in situ.

How the in situ activation (and inactivation) of VSG genes occurs has thus far defied analysis. The problem is that the start of VSG gene transcription lies very far upstream of the gene and has only recently been reached by chromosome walking (*55*). What happens at the promoter when transcription is activated or inactivated is still unknown. A single mobile promoter unit that hops from telomere to telomere has been ruled out by the demonstration that two different telomeric sites of VSG gene expression may be simultaneously active (*56, 57*). It is therefore likely that expression sites are activated or inactivated independently of each other; some form of gene rearrangement might still be involved, for example, insertion of active or inactive promoter cassettes by gene conversion. This would account for the stochastic, low-frequency character of switching. This character can also be explained without invoking gene rearrangements, however (*58, 59*).

In summary, expression site control of surface antigens is complicated because the trypanosome contains multiple telomeric expression sites for VSG genes, and these sites can be activated or inactivated independently (*60*). How this occurs and how many sites exist remain to be determined. Various investigators have reviewed the ins and outs of trypanosome antigenic variation (*44, 46, 61*).

Antigenic Variation in *Borrelia*

In many respects the eubacterial genus *Borrelia* resembles the African trypanosomes in its behavior. *Borrelia hermsii* causes relapsing fever in humans; it multiplies in the bloodstream; it escapes from host antibodies by resorting to antigenic variation; and its antigenic variation is controlled by gene rearrangements (*62*). In *Borrelia* the immunodominant surface protein is called variable major protein (VMP), and the progeny of a single organism has been shown to give rise to at least 26 different VMPs. The recent analysis of two VMP genes demonstrated that these genes are activated by their duplicative transposition to an expression site. How duplication activates the transposed gene is not yet known, but this is probably

Table 1. Gene activation by programmed gene rearrangements; modes and mechanisms.

System	Gene rearrangement				Gene activation by
	Nature	Mechanism	Size of homologous sequences involved	Enzymes involved	
Phase variation in *Salmonella*	Inversion	Homologous recombination	Short	Specific recombinase DNA-binding proteins	Promoter addition
Mating-type switch in yeast	Duplicative transposition	Gene conversion	Long	Specific endonuclease standard recombination machinery	Derepression
Trypanosome antigenic variation	Duplicative transposition	Gene conversion?	Short	?	Promoter addition
	Reciprocal recombination	Homologous recombination?	Short?	?	Promoter addition
Immunoglobulin gene formation in mammals	Deletion/ inversion	Homologous but imprecise recombination	Short	Recombinase +?	Enhancer addition
Opacity protein variation in *Neisseria*	Deletion/ inversion	?	Short	?	Restoring reading frame

by promoter addition rather than derepression, as the expressed gene copies are also transcribed after cloning in *Escherichia coli*, whereas the silent copies are not. The mechanism of the transposition remains to be clarified. Plasterk *et al.* (*63*) favor a reciprocal recombination between silent VMP gene and expression site with subsequent loss of the displaced old expressed copy. In our opinion the published evidence does not exclude gene conversion (*47*), however, and this would seem the simpler mechanism.

An interesting feature of VMP genes is their location on relatively small (30-kb) linear plasmids (*62, 63*), which is analogous to the location of VSG genes in trypanosomes on abundant minichromosomes (*64, 65*). Whether the similarity is spurious or has a functional basis, for example, ease of transposition, remains to be seen.

Antigenic Variation in *Neisseria*

Neisseria gonorrhoeae, the causative agent of gonorrhea, provides the third example of antigenic variation controlled by transposition of genes to an expression site. The immunodominant antigen in this case is pilin, the protein subunit of the long hairlike protein appendages (pili) on the surface of *Neisseria* (*66*). The bacterium can vary the antigenic composition of these pili, but it can also switch off pilin synthesis altogether, resulting in a pilus minus (P^-) phenotype (*67, 68*). Switching between P^+ and P^- occurs frequently and is known as phase variation.

Several important molecular details of antigenic and phase variation remain to be filled in. To simplify the discussion, we present some of the major facts in a speculative scheme in Fig. 5. Comparison of the pilins produced by different antigenic variants of *Neisseria* has established (*69, 70*) that the protein can be divided in three segments, each taking up about one-third of the protein, a constant NH_2-terminal segment, a semivariable segment, and a hypervariable COOH-terminal segment. There are seven clusters of silent pilin genes in *N. gonorrhoeae*, strain MS11 (*67*). Sequence analysis of cluster *pil* S1 has shown that it contains six different genes and that these genes are incomplete because they lack part of the NH_2-terminal sequences of pilin (*70*). All genes but one lose homology with the gene in the expression site at the COOH-terminus (see Fig. 5). It is likely that the silent genes present in the other clusters are also truncated at their 5′ end, because a probe for the DNA sequence that codes for the signal peptide of the pilin precursor only hybridizes to restriction fragments containing the expression site (*67*). These data explain why the NH_2-terminal part of all pilin variants is identical: this sequence is present in its entirety only in the expression site.

How silent genes move into the expression site is not yet known in detail. In two cases, movement has been shown to occur by a duplicative transposition of a silent gene to the expression site (*67, 68*), and gene conversion seems the most plausible mechanism for this. The necessary homology between the silent genes and the gene in the expression site may be provided by the DNA sequence that codes for the constant part of the protein, part of which is present in some silent genes, and by conserved segments of DNA sequence scattered through the semivariable and hypervariable regions (*69, 70*). Other conserved repeats are present between some silent genes and downstream of silent genes and of the gene in the expression site (*70*).

Recent results suggest that the phase variation from P^+ to P^- can be due to transfer of a nonfunctional silent pilin gene into the expression site (*68, 69*); whether this can account for all cases of phase variation, some of which involve deletions in the expression site (*67–69*), is still unclear.

The pili are not the only surface components of *Neisseria* bacteria

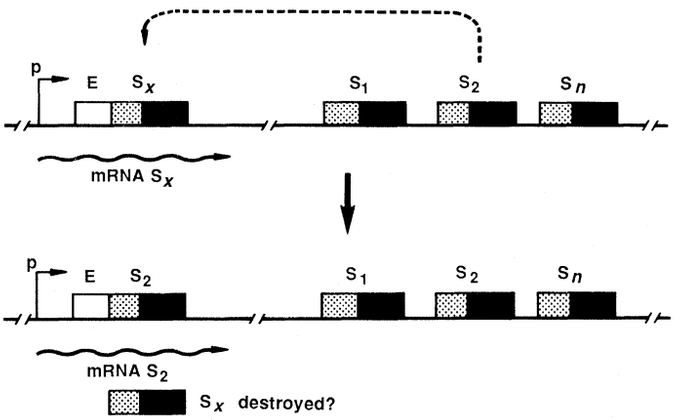

Fig. 5. Switching of pilin gene expression in *Neisseria*. This simplified and speculative scheme incorporates data from (*66–69*). $E(S_x)$ denotes the pilin gene expression site; S_1 to S_n represent silent copies of pilin genes that can be transposed to the expression site (dotted arrow). The white, hatched, and black blocks represent the constant, variable, and hypervariable segments of pilin genes, respectively. A complication omitted from this figure is that the *N. gonorrhoeae* strain most intensively analyzed at the DNA level contains two identical expression sites, whereas other strains have only one.

that can be varied. The opacity protein (protein II) is a second major surface protein that undergoes antigenic and phase variation. *Neisseria* contains many genes for the opacity protein, but in contrast to the pilin genes these *opa* genes are not truncated, and the expression of these genes is controlled at the translational level and not at transcriptional initiation (*71*). A mechanism for this translational control has recently been suggested by Stern *et al.* (*71*). They have found that the leader peptide sequence of the opacity protein mRNAs contains a series of adjacent CTCTT pentamers, varying between 7 and 28 units in different genes. Addition or removal of one or two pentamers will throw translation out of frame and prevent synthesis of the corresponding protein. Variation in pentamer number should be readily accomplished by gene conversion, nonhomologous recombination, or errors in replication (*71*). Although it seems wasteful to control such a large gene family at the translational level, it is certainly a novel variation on the theme of programmed gene rearrangements.

Although the antigenic changes in pilin and opacity protein help the gonococcus to escape host immune defense, these changes also affect other complex parasite-host interactions, such as gonococcal colonization of mucosal surfaces and invasion of epithelial cells (*72*). In this respect antigenic variation in *Neisseria* differs from that in trypanosomes and *Borrelia*.

Segmental Gene Conversion in Chicken Immunoglobulin Genes

Gene transpositions need not involve complete genes. For antigenic variation to be effective, only part of the surface antigen that is exposed to antibody need be replaced. Indeed, DNA inversion in phage mu only affects a segment of the tail fiber genes, gene replacement is only partial in *Neisseria* pilin variation, and segmental gene conversion has also been observed in trypanosome antigen genes. The most exuberant application of this principle, however, is found in the generation of the somatic genes for chicken imumnoglobulins.

Recent work by Reynaud and co-workers (*73*) has demonstrated that chickens follow a route to immunological diversity that differs radically from that found in mammals. There is only a single V, J, and C segment for the light chain of the immunoglobulin in

chickens, and these segments are joined in a precise way. Diversity in these primordial light chain genes is created solely by somatic mutation brought about by multiple rounds of segmental gene conversion in which the segments are derived from a large family of pseudo-V genes. Diversity of the genes for heavy chains is probably created in an analogous fashion.

It is clear from these recent results that immunological diversity in chicken B cells and antigenic diversity in some of the parasites discussed are created by the same mechanism. In both cases a gene residing in an expression site is altered by gene conversion. The functional consequences differ, of course. In B cells, the result is maximal population diversity of terminally differentiated cells; in the parasites, population diversity is kept to a minimum and each cell remains in principle capable of expressing the entire repertoire of silent genes. Irreversible segmental gene conversions between silent genes may, however, contribute to repertoire diversification in the parasites as well (*44, 70*).

Gene Activation by Programmed Rearrangement: A Comparison of Modes and Mechanisms

Table 1 illustrates the remarkable diversity of the main systems analyzed thus far. Genes may be activated by inversion, deletion, reciprocal recombination, or duplicative transposition. Activation is usually the result of the linkage of the rearranged gene to an element that activates transcription, a promoter or enhancer. The exception is the yeast mating-type switch, in which the transposed gene is activated by the removal of its promoter from repressing neighbor sequences. The mating-type switch is also exceptional in the switching frequency: once per cell cycle, against 10^{-3} and lower for all other systems studied.

In several cases the DNA sequences and the enzymic machinery involved in the gene rearrangement have been characterized. The inversion of bacterial DNA segments requires a short sequence flanking the invertible segment. This is recognized by a sequence-specific recombinase, aided by additional DNA-binding proteins. A sequence-specific recombinase is also responsible for the irreversible gene rearrangements that create functional genes for immunoglobulins and T-cell receptors. Probably the same enzyme system rearranges both sets of genes. The differentiation state of the cell determines which gene set is accessible to the enzyme or enzymes. Chromatin configuration is probably an essential determinant of accessibility; transcription uncovers the DNA for enzymic cleavage. DNA accessibility also plays an important role in the mating-type switch in yeast. This is initiated by a duplex cut in the *MAT* locus, which acts as the expression site for mating-type information, but the DNA sequence recognized by the specific endonuclease is also present at the silent loci that are not cut. Chromatin structure is therefore an important element in the decision about which gene will be rearranged, and it may codetermine the direction of gene transposition.

The five examples presented in Table 1 cover the whole range of known mechanisms in gene activation by gene rearrangement. All examples of inversion control resemble phase variation in *Salmonella* in mechanism. Antigenic variation in *Borrelia* and pilus variation in *Neisseria* resemble the duplicative transposition mode of antigenic variation in trypanosomes. The rearrangement of T-cell receptor genes is identical to that of immunoglobulin genes. A novel mechanism was recently discovered for the control of the opacity genes of *Neisseria*: insertions or deletions controlling the translational reading frame. Whether this mechanism is used elsewhere in nature remains to be seen.

Why Control Gene Expression by DNA Rearrangement?

DNA rearrangements are not necessary for the regulation of gene expression; the vast majority of the prokaryotic and eukaryotic genes analyzed are tightly regulated without moving a base pair in DNA. Programmed DNA rearrangements also have obvious disadvantages. No DNA recombination system is error-free, and the errors may destroy essential cellular functions or activate oncogenes (*74*). So why control gene expression by DNA rearrangements?

The answer to this question must lie in the common features shared by most of these programmed DNA rearrangements: they are uninduced; they are often reversible; they may occur at very low rates; and they allow a cell to express only one of a set of genes at a time (intergenic exclusion). The first three points seem most important, because it must be difficult for a unicellular organism to program infrequent, uninduced, reversible switches merely by repressor or activator mechanisms. In contrast, gene rearrangements are well suited for that purpose. They are catalyzed by constitutive enzymes and therefore do not require induction; they can be made reversible by site-specific recombination (inversion control) or by the generation of extra, disposable, gene copies (expression site control); and the levels of the enzymes involved can easily be set for infrequent events, as exemplified by transposon movement (*75*).

Other features of gene activation by DNA rearrangements seem less universal. Intergenic exclusion is obviously important in the synthesis of surface antigens, where tight control of individual genes is essential to avoid a patchwork surface coat. However, several systems discussed in this article do not use this feature at optimal advantage. Trypanosomes have at least two and probably several expression sites for surface antigen genes, and these sites can be active at the same time. In the yeast mating-type system the silent alleles are kept silent by repression, and this repression can be lost by mutations in any of four unlinked genes. Even the simple phase variation in *Salmonella* makes no optimal use of gene rearrangements for intergenic exclusion, because one of the two genes for flagellar proteins is controlled not directly by gene rearrangements but indirectly via a repressor. Obviously repressor control is tight enough to avoid patchwork flagellae.

In many cases then, activity of genes seems to be controlled by gene rearrangement, whenever it is advantageous to preadapt a small fraction of the population to sudden changes or to opportunities in the environment that come too fast for adequate reaction to them. This explains why most known examples of gene control by gene rearrangements regulate the interaction of an organism with its environment (*18, 19*): surface antigens, host range, and adhesion. It is conceivable that other properties, such as preadaptation to major changes in nutrient supply, might be controlled in this way. Starvation is not immediately lethal, however, and new food is usually not a temporary opportunity. It might therefore be more fruitful for the organism to elaborate an adaptive rather than a preadaptive response.

The Borderline Between Incidental and Programmed Gene Rearrangements

At first sight there is fundamental difference between incidental mutations and the programmed gene rearrangements discussed in this article. Mutational rearrangements are rare and usually deleterious. They often involve short stretches of accidental sequence homology (*76*) or no homology at all (*77*). They are the consequence of DNA damage or by-products of normal reactions in DNA recombination, repair, winding, or unwinding. In contrast, pro-

grammed gene rearrangements are an essential part of normal development, always involve sequence homology, and often require specific recombinases or nucleases. They may be as frequent as once per cell cycle. Although these are clear and profound differences, the borderline between programmed and incidental rearrangements is not always so obvious. We discuss three examples here, trypanosomes, major histocompatibility complex (MHC), and maize.

The rate at which African trypanosomes change coat is only 10^{-6} to 10^{-7} per cell division (*39*). As we discussed in a previous section, the coat change usually involves the transfer of a silent coat gene to a telomeric expression site on another chromosome by gene conversion (*47*), and we have questioned whether such low rates of gene conversion could not result from accidental recombination events. We know of no data for accidental gene conversion of the average *T. brucei* gene, but in the yeast *Saccharomyces*, a simple eukaryote like *T. brucei* (*78*), spontaneous mitotic gene conversion occurs at rates varying between 2.5×10^{-8} and 10^{-4} per cell division (*36, 79*). It is also instructive that the mating type switches at a rate of 10^{-6} in yeast strains (*29*) that are unable to make the specific endonuclease that cuts the *MAT* locus (see Fig. 3). One could therefore argue that trypanosomal gene rearrangements that occur at a rate of 10^{-7} are not programmed in the sense that they require specific enzymes. They could result from the background noise of accidental gene conversion. The program element might be restricted to the availability of a sufficient repertoire of silent surface coat genes and a few expression sites in which these genes can be expressed.

Viewed in this admittedly speculative light, the programmed switching of trypanosome coats has more elements in common with the diversity of the mammalian MHC than might appear at first sight. The MHC codes for a set of surface molecules that show marked polymorphism (*80*). Biologically significant differences between polymorphic alleles may arise by replacement of small gene segments by gene conversion (*81*). The mutation rate of functional MHC genes is 4×10^{-5} per locus per gamete. Even if only one-tenth of these mutations would result from gene conversions, the rate of replacements would be equivalent to that in trypanosomes. As with trypanosomes, selection at the population level probably drives the diversification of the MHC gene repertoire; host antibodies select in the case of trypanosomes; and selection in mammals is driven by pathogens. The programmed aspect of the gene conversions in the MHC would be the presence of a large repertoire of MHC genes and, possibly, hot spots for the initiation of gene conversion (*82*).

It should be obvious, however, that it would not be useful to include MHC polymorphisms among the programmed gene rearrangements. Whereas one trypanosome can switch from among a large number of individual coat genes, the diversity of the MHC is based on allelism, and any mouse is stuck with the alleles it has. In our definition, programmed gene rearrangements benefit an organism and its direct descendants and the intraspecies heterogeneity of the MHC does not fall under this definition.

Programming for Change: Mobile Elements in Maize

The third example discussed here to illustrate the borderline between programmed and mutational gene rearrangements comes from work with higher plants. The transposable elements discovered by McClintock in plants can affect adjacent genes in a developmentally controlled fashion, but, in addition, their movement can be under developmental control (*83*). Since one complete element can influence the behavior of many incomplete elements in the same cell, the overall effect of these mobile elements on plant development can be impressive. It is therefore not surprising that McClintock initially thought that normal plant development might be controlled by these elements even though she had considered the possibility that these elements were "foreign genetic systems, such as viruses or episomes of some type, that had become incorporated into the maize genome" (*84*, p. 163). The interpretation that McClintock's plant-controlling elements would be an essential component of normal plant development has never been popular, and it has been disproven by molecular studies (*85*) showing that these elements have properties similar to those of mobile elements in bacteria or fungi (*85*). Nevertheless, these elements have three striking properties that make them suitable for playing an important role in restructuring the plant genome under conditions of environmental stress (*86, 87*). First, transposition is induced by stress, such as chromosome breaks (*88*) or virus infection (*89*). The mechanism of this induction is unknown, but the phenomenon is reminiscent of the excision of phage DNA from the genome of their host as part of the reaction to the bacterial distress call in response to stress. Second, movement of the element leaves a footprint, because excision is not precise (*85, 90*). Third, the genetic activity of the element can be suppressed within a few plant generations by DNA modification (*91*). It is difficult to believe that these properties are accidental idiosyncrasies of parasitic elements. Rather it seems likely that evolutionary selection of these elements has allowed rapid genomic change to create a diversity of new genotypes that have permitted the species to survive in changing external conditions (*86, 88*).

In a similar vein, it has been argued that most mobile genetic elements are well-adapted symbionts of the genome in which they reside, the increased evolutionary flexibility that these mutator genes provide to their host offsetting the deleterious mutations incurred by their infrequent movements. Even though strict control of movement (*75, 92*) and target sequence specificity (*93*) may limit the damage inflicted by mobile elements, the effects of transposition on the host and its progeny are too uncertain to include such movements among the programmed gene rearrangements.

Prospects

For a long period, the interest in programmed gene rearrangements controlling gene expression was fueled by speculations that early development in animals and plants might be steered by rearrangements. The totipotency of somatic plant cells and at least some animal nuclei has dampened the enthusiasm for such ideas (*94*), although they are by no means forgotten (*95*). Indeed, expression site control is reversible, and its use in development would not necessarily violate totipotency. From what we know now, however, it seems likely that the major applications for programmed gene rearrangements in multicellular organisms have been found and that less hazardous methods are used to control early development. This does not detract from the importance of further delineating the role of gene rearrangements in the complex life of simple organisms and in the construction of antibodies and T-cell receptors.

Another motive for the study of programmed gene rearrangement stems from the use of this device by pathogens to vary surface antigens and evade the immune defense of humans. Precise information of the switching process might reveal a serious weakness that might be exploited in killing African trypanosomes or gonococci. It is still uncertain how many pathogens use gene rearrangements to vary their surface antigens, but the few cases known now are not likely to be the only ones. Potential candidates are the gene rearrangements of unknown function observed in the protozoan

parasite *Babesia bovis* (*96*) and the remarkable phenotypic variability in the pathogenic yeast *Candida albicans*, which can switch among at least seven phenotypes with a combined frequency of 10^{-4} per division (*97*). The study of programmed gene rearrangements will continue to alert biologists to the inherent flexibility of genomes and to the importance of incidental, mutational rearrangement for disrupting and reshaping genomes. Such alterations can lead to inherited disease (*98*), cancer (*99*), or resistance to carcinostatic drugs (*100*).

Finally, the area between programmed and incidental gene rearrangements remains a major challenge. Few experiments have been done to study the advantages and disadvantages of mobile elements in a genome (*101*) and the interesting speculations about the possible advantages of mobile promoters (*102*) deserve critical testing. Although we may not find "entirely new mechanisms for cellular differentiation in both development and evolution" (*103*, p. xvi), there is certainly still a lot to do.

REFERENCES AND NOTES

1. M. Simon and I. Herskowitz, in *Genome Rearrangement*, M. Simon and I. Herskowitz, Eds., vol. 20 of *UCLA Symposium on Molecular and Cellular Biology* (Liss, New York, 1980), pp. xv–xix; P. Borst, *Biochem. Int.* **12**, 567 (1986).
2. R. T. Schimke, *Cell* **37**, 705 (1984); J. L. Hamlin, J. D. Milbrandt, N. H. Heintz, J. C. Azizkhan, *Int. Rev. Cytol.* **90**, 31 (1984); G. R. Stark and G. M. Wahl, *Annu. Rev. Biochem.* **53**, 447 (1984); K. Alitalo, *Trends Biochem. Sci.* **10**, 194 (1985).
3. F. C. Kafatos, W. Orr, C. Delidakis, *Trends Genet.* **1**, 301 (1985).
4. S. Beermann, *Chromosoma* **60**, 297 (1977).
5. H. Tobler, in *Germline—Soma Differentiation*, W. Hennig, Ed., vol. 13 of *Results and Problems in Cell Differentiation* (Springer-Verlag, New York, 1986), pp. 1–69.
6. P. Leder, *Sci. Am.* **246**, 102 (May 1982); S. Tonegawa, *Nature (London)* **302**, 575 (1983); G. D. Yancopoulos and F. W. Alt, *Annu. Rev. Immunol.* **4**, 339 (1986).
7. L. Hood, M. Kronenberg, T. Hunkapiller, *Cell* **40**, 225 (1985); M. Kronenberg, G. Siu, L. Hood, N. Shastri, *Annu. Rev. Immunol.* **4**, 529 (1986); M. Malissen *et al.*, *Nature (London)* **319**, 28 (1986); D. Baltimore, *ibid.*, p. 12.
8. D. J. Kemp, A. W. Harris, S. Cory, J. M. Adams, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2876 (1980); G. D. Yancopoulos and F. W. Alt, *Cell* **40**, 271 (1985).
9. X. Wang and K. Calame, *Cell* **43**, 659 (1985); T. Gerster, D. Picard, W. Schaffner, *ibid.* **45**, 45 (1986).
10. G. D. Yancopoulos, T. K. Blackwell, H. Suh, L. Hood, F. W. Alt, *ibid.* **44**, 251 (1986).
11. Several laboratories have reported endonuclease activities in B- and non-B-cell nuclear extracts capable of cleaving sequences in or around nonamer/heptamer blocks. [S. Desidero and D. Baltimore, *Nature (London)* **308**, 860 (1984); T. Kataoka, S. Kondo, M. Nishi, M. Kodaira, T. Honjo, *Nucleic Acids Res.* **12**, 5995 (1984); T. J. Hope, R. J. Aguilera, M. E. Minie, H. Sakano, *Science* **231**, 1141 (1986).]
12. H. Kraut, H. J. Lipps, D. M. Prescott, *Int. Rev. Cytol.* **99**, 1 (1986); C. F. Brunk, *ibid.*, p. 49.
13. D. Dawson, B. Buckley, S. Cartinhour, R. Myers, G. Herrick, *Chromosoma* **90**, 289 (1984); M.-C. Yao, J. Choi, S. Yokoyama, C. F. Austerberry, C.-H. Yao, *Cell* **36**, 433 (1984); R. C. Callahan, G. Shalke, M. A. Gorovsky, *ibid.*, p. 441; J. M. Cherry and E. H. Blackburn, *ibid.* **43**, 747 (1985); G. Herrick *et al.*, *ibid.*, p. 759; L. A. Klobutcher, D. L. Jahn, D. M. Prescott, *ibid.* **36**, 1045 (1984).
14. C. F. Austerberry, C. D. Allis, M.-C. Yao, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7383 (1984); E. A. Howard and E. H. Blackburn, *Mol. Cell. Biol.* **5**, 2035 (1985).
15. E. Orias, *Dev. Genet.* **2**, 185 (1981).
16. J. B. Hicks, J. N. Strathern, J. Herskowitz, in *DNA Insertion Elements, Plasmids and Episomes*, A. J. Bukhari, J. A. Shapiro, S. L. Adhya, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1977), pp. 457–462.
17. When *Anabaena* is starved for nitrogen, it initiates a developmental program in which about one in ten cells terminally differentiate into heterocysts with a concomitant induction of nitrogen fixation enzymes. Induction triggers the excision of an 11-kb sequence from the *nif* operon, resulting in a functional operon transcribed from a single promoter. Excision is catalyzed by a site-specific recombinase, which cuts a short repeat that flanks the 11-kb excision and which is unrelated to other known site-specific recombinases. The functional consequences of this gene rearrangement are not yet known. Circularization of the excised DNA does not create a new gene; it is also unclear why it should be advantageous to interrupt the *nif* operon by an extra chunk of DNA in vegetative cells. Nevertheless, the operation may represent a form of gene control by excision and a method for irreversible terminal differentiation. It is basically different from most other systems discussed in this review in that it is a gene rearrangement induced by an environmental cue [J. W. Golden, S. J. Robinson, R. Haselkorn, *Nature (London)* **314**, 419 (1985); P. J. Lammers, J. W. Golden, R. Haselkorn, *Cell* **44**, 905 (1986); R. Haselkorn, J. W. Golden, P. J. Lammers, M. E. Mulligan, *Trends Genet.* **2**, 255 (1986)].
18. J. A. Shapiro, Ed., *Mobile Genetic Elements* (Academic Press, New York, 1983).
19. M. Silverman and M. Simon, in (*18*), pp. 537–557; R. H. A. Plasterk and P. van de Putte, *Biochim. Biophys. Acta* **782**, 111 (1984); N. L. Craig, *Cell* **41**, 649 (1985).
20. R. C. Johnson and M. I. Simon, *Cell* **41**, 781 (1985).
21. R. Kahmann, R. Rudt, C. Koch, G. Mertens, *ibid.*, p. 771; H. E. Huber, S. Jida, W. Arber, T. E. Bickle, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3776 (1985); S. Jida and R. Hiestand-Nauer, *Cell* **45**, 71 (1986); R. Kanaar, P. van de Putte, N. R. Cozzarelli, *Biochim. Biophys. Acta* **866**, 170 (1986).
22. R. H. A. Plasterk, A. Brinkman, P. van de Putte, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5355 (1983).
23. J. L. Hartley and J. E. Donelson, *Nature (London)* **286**, 860 (1980).
24. J. F. Senecoff, R. C. Bruckner, M. M. Cox, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7270 (1985); B. J. Andrews, G. A. Proteau, L. G. Beatty, P. D. Sadowski, *Cell* **40**, 795 (1985).
25. P. Argos *et al.*, *EMBO J.* **5**, 433 (1986).
26. F. C. Volkert and J. R. Broach, *Cell* **45**, 541 (1986).
27. K. L. Poffenbergen, E. Tabares, B. Roizman, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2690 (1983); J. Chou and B. Roizman, *Cell* **41**, 803 (1985).
28. I. Herskowitz and Y. Oshima, in *Molecular Biology of the Yeast Saccharomyces*, J. N. Strathern, E. W. Jones, J. R. Broach, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1981), pp. 181–209; K. A. Nasmyth, *Annu. Rev. Genet.* **16**, 439 (1982); G. F. Sprague, L. C. Blair, J. Thorner, *Annu. Rev. Microbiol.* **37**, 623 (1983).
29. J. E. Haber, in (*18*), pp. 559–619.
30. J. B. Feldman, J. B. Hicks, J. R. Broach, *J. Mol. Biol.* **178**, 815 (1984).
31. J. Abraham, K. A. Nasmyth, J. N. Strathern, A. J. S. Klar, J. B. Hicks, *ibid.* **176**, 307 (1984); A. H. Brand, L. Breeden, J. Abraham, R. Sternglanz, K. A. Nasmyth, *Cell* **41**, 41 (1985).
32. J. N. Strathern *et al.*, *Cell* **31**, 183 (1982); A. J. S. Klar, J. N. Strathern, J. A. Abraham, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 77 (1984).
33. R. Kostriken, J. N. Strathern, A. J. S. Klar, J. B. Hicks, F. Heffron, *Cell* **35**, 167 (1983); R. Kostriken and F. Heffron, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 89 (1984).
34. K. A. Nasmyth, *Cell* **30**, 567 (1982).
35. A. J. S. Klar, J. B. Hicks, J. N. Strathern, *ibid.* **28**, 551 (1982).
36. J. W. Szostack, T. L. Orr-Weaver, R. J. Rothstein, F. W. Stahl, *ibid.* **33**, 25 (1983).
37. D. H. Beach and A. J. S. Klar, *EMBO J.* **3**, 603 (1984).
38. K. Vickerman, *Nature (London)* **273**, 613 (1978).
39. G. S. Lamont, R. S. Tucker, G. A. M. Cross, *Parasitology*, **92**, 355 (1986); J. J. Doyle, H. Hirumi, K. Hirumi, E. N. Lupton, G. A. M. Cross, *ibid.* **80**, 359 (1980).
40. G. A. M. Cross, *ibid.* **71**, 393 (1975).
41. J. H. J. Hoeijmakers, A. C. C. Frasch, A. Bernards, P. Borst, G. A. M. Cross, *Nature (London)* **284**, 78 (1980).
42. L. H. T. van der Ploeg *et al.*, *Nucleic Acids Res.* **10**, 5905 (1982).
43. A. Capbern, C. Giroud, T. Baltz, P. Mattern, *Exp. Parasitol.* **42**, 6 (1977).
44. E. Pays, *Prog. Nucleic Acid Res. Mol. Biol.* **32**, 1 (1986); ——— *et al.*, *Cell* **42**, 821 (1985); E. Pays, *Trends Genet.* **2**, 21 (1986).
45. A. Y. C. Liu, L. H. T. van der Ploeg, F. A. M. Rijsewijk, P. Borst, *J. Mol. Biol.* **167**, 57 (1983); D. A. Campbell, M. P. van Bree, J. C. Boothroyd, *Nucleic Acids Res.* **12**, 59 (1984).
46. P. Borst, *Annu. Rev. Biochem.* **55**, 701 (1986).
47. The term "gene conversion" is used here for any process in which gene A appears to have donated part of its sequence to partly homologous gene B, without being altered itself. Formally, however, gene conversion denotes a nonreciprocal transfer of genetic information, and this can only be distinguished from reciprocal recombination, followed by loss of one of the recombinant alleles, if the parents and progeny involved in the genetic exchange can all be analyzed. This can be done in spore-forming fungi, for which the term "gene conversion" was originally coined, but not in most other organisms discussed here.
48. T. de Lange, J. M. Kooter, J. Luirink, P. Borst, *EMBO J.* **4**, 3299 (1985).
49. P. A. M. Michels *et al.*, *J. Mol. Biol.* **166**, 537 (1983).
50. R. Aline, Jr., *et al.*, *Nucleic Acids Res.* **13**, 3161 (1985).
51. S. Longacre and M. Eisen, *EMBO J.* **5**, 1057 (1986); C. W. Roth, S. Longacre, A. Raibaud, T. Baltz, H. Eisen, *ibid.*, p. 1065.
52. T. de Lange, J. M. Kooter, P. A. M. Michels, P. Borst, *Nucleic Acids Res.* **11**, 8149 (1983).
53. R. O. Williams, J. R. Young, P. A. O. Majiwa, *Nature (London)* **282**, 847 (1979).
54. E. Pays, M. Guyaux, D. Aerts, N. van Meirvenne, M. Steinert, *ibid.* **316**, 562 (1985); C. Shea, D. J. Glass, S. Parangi, L. H. T. van der Ploeg, *J. Biol. Chem.* **261**, 6056 (1986).
55. J. M. Kooter and P. Borst, *Nucleic Acids Res.* **12**, 9457 (1984); unpublished results.
56. A. W. C. A. Cornelissen, P. J. Johnson, J. M. Kooter, L. H. T. van der Ploeg, P. Borst, *Cell* **41**, 825 (1985).
57. T. Baltz *et al.*, *Nature (London)* **319**, 602 (1986).
58. J. R. Young, J. E. Donelson, P. A. O. Majiwa, S. Z. Shapiro, R. O. Williams, *Nucleic Acids Res.* **10**, 803 (1982).
59. A. Bernards, N. van Harten-Loosbroek, P. Borst, *ibid.* **12**, 4153 (1984).
60. Antigenic variation as a stratagem to elude host defense can only work if the population heterogeneity of the parasite remains limited. This is indeed seen in trypanosomiasis (*41*), but how it is accomplished is still not completely clear, although several factors are known to contribute to the loose temporal order in which antigenic variants appear in a chronic infection as follows. Some variants outgrow others [E. N. Miller and M. J. Turner, *Parasitology* **82**, 63 (1981); J. R. Seed, R. Edwards, J. Sechelski, *J. Protozool.* **31**, 48 (1984); P. Myler, A. L. Allen, N. Agabian, K. Stuart, *Infect. Immun.* **47**, 684 (1985)] and therefore tend to appear early. VSG genes activated very early are invariably located near chromosome ends [J. R. Young, J. S. Shah, G. Matthyssens, R. O. Williams, *Cell* **32**, 1149 (1983); A. Y. C. Liu, P. A. M. Michels, A. Bernards, P. Borst, *J. Mol. Biol.* **182**, 383 (1985); R. F. Aline, J. K. Scholler, R. G. Nelson, N. Agabian, K. Stuart, *Mol. Biochem. Parasitol.* **17**, 311 (1985)]. Also, VSG genes that lack some of the recombination areas tend to be activated late, suggesting that the degree of homology between silent gene and expression site may codetermine its ease of activation (*49*). Clearly, temporal control of surface antigen gene expression is an essential component of the expression site control of a large repertoire of silent

genes. Space does not permit a more extensive discussion of this issue, and we refer the reader to (*44*) and (*49*).
61. P. Borst, in (*18*), pp. 621–659; R. G. Parsons, R. G. Nelson, N. Agabian, *Immunol. Today* **5**, 43 (1984); J. E. Donelson and A. C. Rice-Ficht, *Microbiol. Rev.* **49**, 107 (1985); J. C. Boothroyd, *Annu. Rev. Microbiol.* **39**, 475 (1985).
62. P. A. Barstad, J. E. Coligan, J. E. Raum, A. G. Barbour, *J. Exp. Med.* **164**, 1302 (1985); J. T. Meier, M. J. Simon, A. G. Barbour, *Cell* **41**, 403 (1985).
63. R. H. A. Plasterk, M. J. Simon, A. G. Barbour, *Nature (London)* **318**, 257 (1985); T. H. Birkbeck and C. W. Penn, Eds., *Symposium Series of the Society for General Microbiology*, (IRL Press, McLean, VA, 1986), vol. 19, pp. 127–146.
64. L. H. T. van der Ploeg, D. C. Schwartz, C. R. Cantor, P. Borst, *Cell* **37**, 77 (1984).
65. R. O. Williams, J. R. Young, P. A. O. Majiwa, *Nature (London)* **299**, 417 (1982).
66. T. F. Meyer, N. Mlawer, M. So, *Cell* **30**, 45 (1982); G. K. Schoolnik, R. Fernandez, J. Y. Tal, J. Rothbard, E. C. Gotschlich, *J. Exp. Med.* **159**, 1351 (1984); T. F. Meyer, E. Billyard, R. Haas, S. Storzbach, M. So, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6110 (1984).
67. E. Segal, E. Billyard, M. So, S. Storzbach, T. F. Meyer, *Cell* **40**, 293 (1985); E. Segal, P. Hagblom, H. S. Seifert, M. So, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2177 (1986).
68. J. Swanson, S. Bergström, O. Barrara, K. Robbins, D. Corwin, *J. Exp. Med.* **162**, 729 (1985); S. Bergström, K. Robbins, J. M. Koomey, J. Swanson, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3890 (1986); J. Swanson *et al.*, *Cell* **47**, 267 (1986).
69. P. Hagblom, E. Segal, E. Billyard, M. So, *Nature (London)* **315**, 156 (1985).
70. R. Haas and T. F. Meyer, *Cell* **44**, 107 (1986).
71. A. Stern, P. Nickel, T. F. Meyer, M. So, *ibid.* **37**, 447 (1984); A. Stern, P. Nickel, T. F. Meyer, *ibid.* **47**, 61 (1986).
72. J. E. Heckels, in *Antigenic Variation in Infectious Diseases*, T. H. Birkbeck and C. W. Penn, Eds. (Society for General Microbiology, IRL Press, Spec. Publ. no. 19, McLean, VA, 1986), p. 77.
73. C. A. Reynaud, V. Anquez, A. Dahan, J. C. Weill, *Cell* **40**, 283 (1985); C. A. Reynaud, personal communication.
74. Y. Tsujimoto, J. Gorham, J. Cossman, E. Jaffe, C. M. Croce, *Science* **229**, 1390 (1985).
75. E. A. Raleigh and N. Kleckner, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1787 (1986).
76. A. M. Albertini, M. Hofer, M. P. Calos, J. H. Miller, *Cell* **29**, 319 (1982).
77. K. Ikeda and M. Shiozaki, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 401 (1984).
78. P. Borst *et al.*, *Mol. Biochem. Parasitol.* **6**, 13 (1982).
79. S. Scherer and R. W. Davis, *Science* **209**, 1380 (1980); R. Egel, *Nature (London)* **290**, 191 (1981); T. Petes and G. R. Fink, *ibid.* **300**, 216 (1982); M. D. Mikus and T. D. Petes, *Genetics* **101**, 369 (1982); G. R. Fink and T. D. Petes, *Nature (London)* **310**, 728 (1984). The rate increases as genes move closer together, the lowest rates being observed for genes on different nonhomologous chromosomes. The rate of 2.5×10^{-8} per cell division holds for chromosome-internal single-copy genes; the rate for a telomeric member of a multigene family might be higher.
80. L. Hood, M. Steinmetz, B. Malissen, *Annu. Rev. Immunol.* **1**, 529 (1983).
81. P. Kourilsky, *Trends Genet.* **2**, 60 (1986). (This is by no means undisputed. See J. Klein, *CRC Crit. Rev. Immunol.*, in press.)
82. F. Brégégère, *Biochemie* **65**, 229 (1983).
83. N. V. Fedoroff, in (*18*), pp. 1–63.
84. B. McClintock, *Brookhaven Symp. Biol.* **18**, 162 (1985).
85. H. P. Döring and P. Starlinger, *Cell* **39**, 253 (1984).
86. B. McClintock, *Science* **226**, 792 (1984).
87. K. Courage *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 329 (1984).
88. B. McClintock, in *Mobilization and Reassembly of Genetic Information*, W. A. Scott *et al.*, Eds. (Academic Press, New York, 1980), pp. 11–19.
89. S. L. Dellaporta *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **49**, 321 (1984).
90. H. Saedler and P. Nevers, *EMBO J.* **4**, 585 (1985).
91. V. L. Chandler and V. Walbot, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1767 (1986).
92. D. Roberts, B. C. Hoopes, W. R. McClure, N. Kleckner, *Cell* **43**, 117 (1985); F. A. Laski, D. C. Rio, G. M. Rubin, *ibid.* **44**, 7 (1986); D. C. Rio, F. A. Laski, G. M. Rubin, *ibid.*, p. 21.
93. G. S. Roeder and G. R. Fink, in (*18*), pp. 299–329; S. Jida, J. Meyer, W. Arber, *ibid.*, pp. 159–221; G. M. Rubin, *ibid.*, pp. 329–361; D. J. Finnegan, *Int. Rev. Cytol.* **93**, 281 (1985).
94. D. D. Brown, *Science* **211**, 667 (1981); M. A. DiBerardino, N. J. Hoffner, L. D. Etkin, *ibid.* **224**, 946 (1984).
95. J. E. Till, *J. Cell. Physiol. Suppl.* **1**, 3 (1982); I. Herskowitz, *Curr. Top. Dev. Biol.* **18**, 1 (1984).
96. A. F. Cowan, O. Bernard, N. Stewart, D. J. Kemp, *Cell* **37**, 653 (1984).
97. B. Slutsky, J. Buffo, D. R. Soll, *Science* **230**, 666 (1985); B. Slutsky, M. Staebell, M. Pfaller, D. R. Soll, unpublished data.
98. N. A. Jenkins *et al. Nature (London)* **293**, 370 (1981).
99. G. Klein and E. Klein, *ibid.* **315**, 190 (1985); J. M. Adams, *ibid.*, p. 542.
100. R. T. Schimke, *Cancer Res.* **44**, 1735 (1984); J. R. Riordan and V. Ling, *Pharmacol. Ther.* **28**, 51 (1985).
101. L. Chao, C. Vargas, B. B. Spear, E. C. Cox, *Nature (London)* **303**, 633 (1983); D. L. Hart, D. E. Dijkhuizen, R. D. Miller, L. Green, J. de Framond, *Cell* **35**, 503 (1983).
102. A. E. Reynolds, J. Felton, A. Wright, *Nature (London)* **293**, 625 (1981); A. E. Reynolds, S. Mahadevan, S. F. J. LeGrice, A. Wright, *J. Mol. Biol.* **191**, 85 (1986); D. Charlier, J. Piette, N. Glansdorff, *Nucleic Acids Res.* **10**, 5935 (1982).
103. J. A. Shapiro, in (*18*), pp. xi–xvi.
104. We are indebted to our colleages in The Netherlands Cancer Institute and two reviewers for suggestions and critical comments and to E. Spronk and B. van Houten for typing many versions of the manuscript. The research of P.B. is supported by grants from The Netherlands Foundation for Chemical Research (SON), The Netherlands Organization for the Advancement of Pure Research (ZWO), Queen Wilhelmina Fund, and World Health Organization. D.R.G. is supported by a postdoctoral fellowship from the European Programme of the Royal Society.