Genetica **107:** 171–179, 1999.

Transposable elements as the key to a 21st century view of evolution

James A. Shapiro
University of Chicago, Department of Biochemistry and Molecular Biology, Cummings Life Sciences Center, 920 E. 58th Street, Chicago, IL 60637-4931, USA (Phone: 773-702-1625; Fax: 773-702-0439; E-mail: jsha@midway.uchicago.edu)

Accepted 18 January 2000

Key words: evolutionary feedback, natural genetic engineering, genomic systems, genome-wide transposition, transcriptional regulatory circuits

Abstract

Cells are capable of sophisticated information processing. Cellular signal transduction networks serve to compute data from multiple inputs and make decisions about cellular behavior. Genomes are organized like integrated computer programs as systems of routines and subroutines, not as a collection of independent genetic 'units'. DNA sequences which do not code for protein structure determine the system architecture of the genome. Repetititve DNA elements serve as tags to mark and integrate different protein coding sequences into coordinately functioning groups, to build up systems for genome replication and distribution to daughter cells, and to organize chromatin. Genomes can be reorganized through the action of cellular systems for cutting, splicing and rearranging DNA molecules. Natural genetic engineering systems (including transposable elements) are capable of acting genome-wide and not just one site at a time. Transposable elements are subject to regulation by cellular signal transduction/computing networks. This regulation acts on both the timing and extent of DNA rearrangements and (in a few documented cases so far) on the location of changes in the genomes. By connecting transcriptional regulatory circuits to the action of natural genetic engineering systems, there is a plausible molecular basis for coordinated changes in the genome subject to biologically meaningful feedback.

Introduction

The goal of this presentation is to delve into some conceptual issues in evolutionary theory raised by the existence and action of transposable elements. These elements constitute internal biochemical systems for DNA rearrangement, and they account for a large proportion of genetic changes (e.g. Green, 1987). The existence of transposable elements means that evolutionary variability occurs in the highly regulated realm of cell biology (Alberts et al., 1994). Since there is no reason to suppose that biochemical systems working on DNA are less subject to regulation than any other cellular functions, biological information-processing has the potential to play a major role in genome change during the course of organismal evolution.

The current prevailing view of evolution developed in the first four decades of this century. This perspective combined Darwinian concepts of gradualism and natural selection with random mutation and Mendelian segregation as the mechanisms of evolutionary variability. The early 20th century view of evolution developed its basic outlines before we knew about DNA as the genetic material. As it was reaching its mature formulation, there were a series of landmark discoveries which were to transform our understanding of genome structure, organization and function (Table 1). Among these discoveries were the identification of DNA as the genetic material by Avery et al. (1944) and the deciphering of its double helical structure by Watson and Crick (1953). These set the stage for the future elaboration of molecular genetics. About the same time, McClintock (1950, 1951) discovered that

Table 1. Historical benchmarks

1859	–	Darwin – On the origin of species by means of natural selection
1900	–	Rediscovery of Mendelism
1944	–	Avery, MacLeod and McCarty – DNA as genetic material
1950	–	McClintock – Genomes contain mobile elements restructuring chromosomes and changing patterns of gene expression
1953	–	Watson and Crick – double-helical structure of DNA
1961	–	Jacob and Monod – Operon theory: regulatory proteins and regulatory sites
1968	–	Britten and Kohne – Abundant repetitive DNA in genomes

cells contain internal systems mediating a wide variety of genetic changes, including extensive chromosome rearrangements as well as alterations in the regulated expression of diverse genetic loci.

A decade later, the Operon theory of Jacob and Monod (1961) had several fundamental consequences. It helped us understand the composite, systemic nature of individual genetic loci. It defined a whole new class of genetic elements (protein binding sites). And it made clear how regulatory proteins and their cognate binding sites can form integrated functional systems involving loci at multiple locations in the genome (Monod & Jacob, 1961). Finally, the discovery of abundant repetitive DNA in higher organisms by Britten and Kohn (1968) showed that large fractions of the genome were organizationally (and therefore functionally) different from the unique 'genes' of the pre-DNA era. We now know that repetitive DNA elements are present in all genomes, and that many genomes contain far more repeated sequences than single-copy genetic loci.

A 21st century view of evolution will incorporate a more informational perspective on the structure and operation of genetic systems. One of the main realizations emerging from contemporary cell and developmental biology is that essentially all cellular functions are regulated by interactive 'signal transduction' networks composed of information transfer molecules, such as G proteins, protein kinases, second messengers and transcription factors (Alberts et al., 1994). In effect, these signal transduction networks are now seen to be cellular computation systems allowing cells to evaluate multiple internal and external inputs in order to make appropriate decisions (e.g. which enzymes to synthesize, when to divide, where to move) (Bray, 1990; Gerhart & Kirschner, 1997).

In this informational context, the cellular DNA can be thought of as a storage medium, like a hard disk. It contains coding information for the proteins and RNAs that the cells need to function. This coding information must be dynamically accessible for reading at the right time and in the right amounts as different molecular programs are executed. The coding information for many potential programs are contained within a single genome: for example, housekeeping routines operating in all cells, specialized patterns of protein synthesis leading to distinct cell types, emergency responses to deal with certain repetitive crises, like oxidative damage and starvation, and, in organisms with complex life cycles, programs for making different organisms at each stage, such as caterpillars and butterflies. Coordinately retrieving the information for multiple RNA and protein molecules required to execute each program or set of programs imposes a need for physical organization of the genome and for addressing individual coding regions. This is achieved by using combinations of repeated sequences as address tags on related genetic loci (Britten & Davidson, 1969) and by organizing some loci in physically distinct regions within the genome (so-called chromatin domains; Felsenfeld et al., 1996).

In addition to its coding and physical organization, the genome has other requirements to fulfill as an efficient biological information storage system. It must replicate at the right time once per cell cycle; equal copies of the duplicated genome must be distributed to daughter cells following division; replication errors and physical or chemical damage to sequence information must be detected and repaired; and (most significantly for evolution) there must be a capacity for reprogramming the information content of the genome when necessary. In some organisms, such as man, this kind of reorganization is part of the normal life cycle. Our lymphocytes undergo a tightly regulated sequence of DNA rearrangements to assemble, improve and modify the recognition proteins of our immune sys-

Table 2. Some functions of repetitive DNA elements*

Coordinated expression of unlinked genetic loci
Activator regions
Silencing regions
DNA replication origins
Chromosome end stability (telomeres)
Chromosome distribution during cell division
Centromeres
Chromosome pairing in meiosis
Chromatin organization and timing of
gene expression in development (position effect)

*Specific examples and references in Shapiro, 1999b.

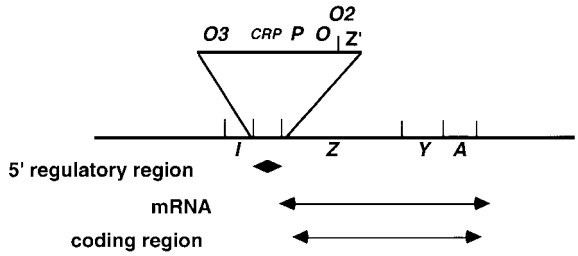

Figure 1. E. coli lac operon (modified from Reznikoff, 1992).

tem (Blackwell & Alt, 1989; Lewis, 1999; Kenter & Wuerfel, 1999). Without the natural genetic engineering that occurs in our B and T cells, we would perish due to severe immunodeficiency. In all organisms that persist in evolution, there is a need for reprogramming in order to survive crises or exploit new ecological opportunities that cannot be handled using the existing genome.

Meeting the genomic requirements listed above involves repetitive DNA sequences (Table 2). As detailed elsewhere in this volume, many of these repeats are also transposable elements.

Transposable elements and genome organization

McClintock called the transposable elements she discovered 'controlling elements' because she observed them popping in and out of individual genetic loci and altering their capacity for expression and their regulation during development (McClintock, 1953, 1956b). Her observations were among the first indications that genetic loci are not indivisible units but rather consist of modular systems built up from different kinds of sequence components. This modular view of each genetic locus as a system is consistent with the results of molecular genetic analyses which have identified a wide variety of components – promoters, enhancers (acting sometimes as activators, sometimes as silencers), introns, exons, splice signals, sequence segments encoding protein domains, transcription terminators, RNA processing signals, etc.

A basic example like the *E. coli lac* operon (Figure 1; Reznikoff, 1992) illustrates the point. The *lac* operon is a composite of coding sequences for three different proteins (Z, Y, A, each containing domains) and a composite 5′ regulatory region containing binding sites for RNA polymerase (the promoter, P), for LacI repressor (the operators, O1, O2 and O3) and for the cAMP receptor protein (CRP). P is only functional as a strong promoter when the Crp-cAMP complex is bound at the CRP site, and cAMP levels are low when glucose is present (catabolite repression; Saier et al., 1996). Transcription from P, even in the presence of Crp-cAMP, is blocked if LacI repressor is not removed from the operators by binding the inducer molecule allolactose (a lactose derivative formed intracellularly by basal levels of beta-galactosidase (LacZ)) activity.

The *lac* operon complex is effectively part of a computational system designed to control expression of the proteins specifically needed for lactose catabolism. The algorithm governing this system can be stated formally – 'If lactose present and if no glucose present, then transcribe ZYA'. It is important to note that this simple computation does not just involve the DNA and transcription factors. Since cytoplasmic allolactose formation requires the presence of low levels of LacY permease and LacZ, and since the activity of the enzyme which synthesizes cAMP is controlled by a glucose-specific protein of the phosphotransferase transport system (Saier et al., 1996), it is clear that computing whether or not to transcribe lacZYA involves components distributed over the whole cell (membrane, cytoplasm and genome). This kind of whole-cell integration is typical of most regulatory computations governing gene expression. More complex 5′ regulatory systems have been described explicitly in computational terms in higher organisms, such as the sea urchin (Yuh et al., 1998).

Not only is each genetic locus itself a rather intricate system, but virtually all cellular and organismal phenotypes are encoded by coordinated networks involving many genetic loci, linked together by common (i.e. repetitive) protein binding sites. The CRP site serves as a repetitive element to integrate the *lac* operon with other loci in the *E. coli* genome whose expression is regulated by glucose availability. McClin-

tock demonstrated the ability of transposable elements to construct such networks. She isolated insertions of related elements in loci on different chromosomes and then showed that the modified loci responded similarly to changes in transposase activity in the same clonal lineages (McClintock, 1956a, 1965). In bacteria, we know that many phenotypes are determined by expression of several operons which share promoter or other regulatory sites. A good example is expression of chemotaxis, motility and flagellar biosynthetic functions encoded in 15 operons at five distinct regions of the *E. coli* genome (MacNab, 1992). These operons share sets of promoter sequences that allow them to respond to a cascade of sigma factors during flagellar biogenesis and assembly of the chemotaxis receptor-signal transduction system. In higher organisms, especially during the cell cycle, cellular differentiation, and multicellular development, the complexity of these coregulated suites of proteins can be far more extensive (Alberts et al., 1994; Gerhart & Kirschner, 1997).

Transposable elements, natural genetic engineering, and the potential for major evolutionary rearrangements

The modular organization of genomes as hierarchical systems requires a capacity for cut-and-splice changes (i.e. natural genetic engineering) that transposable elements can provide to cells (Shapiro, 1992). Without these capacities, functionally significant regulatory signals and repetitive elements could not have been distributed throughout the genome to build up coordinated systems. The accumulation of these integrative repeats, one site at a time, by the gradual addition of random nucleotide substitutions would require an unimaginable length of time and would not be consistent with the punctuated nature of the geologic record. Some events, such as the emergence of flowering plants and many different animal body plans, appear to have occured in relatively short time spans.

The roles that transposable elements may have played in evolution can be deduced from several kinds of information:

- their abundance and distribution in contemporary genomes,
- their biologically useful functions in contemporary genomes,
- database evidence for a past evolutionary role to generate currently functional genomic structures, and
- their capacities, demonstrated in the laboratory, for generating useful genome changes.

On all four counts, it is hard to escape the conclusion that transposable elements have played, and will continue to play, a major role in genome reorganization during episodes of evolutionary change.

Virtually all genomes contain significant numbers of transposable elements. In some bacterial species, as much as 10% of the genome can be composed of IS elements (IS database homepage, http://www-is.biotoul.fr/is.html). Mammalian genomes contain large amounts of repetitive DNA, and the abundances of retrotransposable sequences (principally SINEs and LINEs) is often quoted as over 20% of the human genome (Brosius, 1999b; Lerat et al., 1999; Roy et al., 1999). Over 50% of the maize genome is composed of DNA-based, LINE, and retroviral-like transposable elements, and in some plant species the fraction goes as high as 95%. It is inconceivable that chromsomes could have become so filled with transposable elements without a major role for DNA-based transposition and retrotransposition. It is notable, for example, that each mammalian order has its own set of dispersed SINE elements (Roy et al., 1999). Thus, the process of genome-wide retrotransposition must have occurred many times in mammalian evolution.

Contemporary organisms use transposable elements, or their descendants, for a small number of well-defined functions. A direct example is retrotransposition to regenerate telomeres in *Drosophila* (Pardue, 1999). The mechanistic similarities between immune system rearrangements and the action of many DNA-based transposable elements makes it clear that lymphocyte DNA changes are applications of a modified transposition mechanism (Agrawal et al., 1998; Hiom et al., 1998; Lewis, 1999). A recent example of world-wide evolutionary change has been the emergence over the past five decades of transmissible antibiotic resistance in bacteria. The role of transposable elements and other natural genetic engineering systems, such as conjugative plasmids and the gene casette/integron system for building up antibiotic resistance operons (Recchia & Hall, 1995), is extremely well documented at the molecular level in this major evolutionary event. Whole genome analysis of bacteria is beginning to show a similar story for the evolution of pathogenicity and xeno-

biotic degradation determinants (Mazel et al., 1998; Shapiro, 1999a and references therein). From database analysis, a growing number of cases are being documented in vertebrate genomes where regulatory signals can be traced to vestiges of transposable element insertions (Britten, 1997; Brosius, 1999a). Thus, the accumulating DNA evidence shows that transposable elements have been significant players in past evolutionary change to provide new functional systems.

We know quite a lot about how transposable elements operate from experimental studies (Shapiro, 1983; Berg & Howe, 1989). Indeed, the capacities of transposable elements documented in the laboratory are just those which are needed for many aspects of genome reorganization to create new architectures and functions:

- dispersal of multiple copies of a single sequence element to many genomic locations,
- alteration of regulatory patterns at individual genetic loci, including activation of silent loci (McClintock, 1965; Errede et al., 1981; Green, 1987; Hall, 1999); in this regard it is important to remember that virtually all transposable elements carry transcriptional regulatory signals, such as the promoters, enhancers and terminators in retroviral LTRs,
- mobilization of extended chromosome segments in rearrangements such as inversions, translocations, transpositions, duplications and generation of tandem arrays (see Shapiro, 1982; Pardue, 1999, for some models),
- genetic fusions by DNA-mediated rearrangements (Shapiro & Leach, 1990; Maenhaut-Michel et al., 1997),
- transduction of adjacent 5′ and 3′ sequences by retrotransposition to create novel gene fusions, splice patterns, and exon shuffling (Moran, 1999).

The fact that laboratory experiments with transposable elements produce many of the kinds of genetic changes that are needed to explain evolutionary differences between related but distinct organisms makes it highly likely that these elements provided the biochemical mechanisms for some evolutionarily important rearrangements. It does not make sense for cells to possess molecular agents of genome restructuring and not to use them when restructuring is essential to survival or diversification in evolution.

Test and activation in response to biological feedback

Detailed study of the activities of many transposable elements and other natural genetic engineering systems virtually always indicate that their activation is subject to control by regulatory/signal transduction systems. This was true of the initial discovery of transposable elements, when McClintock found several mobile systems activated in response to repeated chromosome breakage during early maize plant development (McClintock, 1951, 1984). In bacteria, molecular genetic analysis has revealed sites in transposable elements for interaction with cell-cycle (DnaA) and transcriptional control factors (IHF) as well as regulation by Dam methylation, translational frameshifting, transcriptional repressors, and truncated inhibitory forms of transposase (Berg & Howe, 1989).

The phenomenon of adaptive mutation by bacteria illustrates responses to environmental and physiological factors as well. Certain mutations arise more frequently under the stress conditions of selection than they do during normal growth (Foster, 1993; Shapiro, 1995, 1997). In the first adaptive mutation system described, a Mu prophage can join the 5′ end of *araB* and the 3′ end of *lacZ* to generate a hybrid *araB-lacZ* coding sequence, in effect serving as a model for making multidomain proteins through the actions of transposable elements (Shapiro, 1984; Shapiro & Leach, 1990; Maenhaut-Michel et al., 1997). These fusions are completely undectectable during normal growth conditions ($<10^{-10}$), but arise at frequencies as high as 10^{-5} after prolonged aerobic starvation (Maenhaut-Michel & Shapiro, 1994). More detailed studies of regulatory functions involved in the fusion process indicate a complex regulatory network, with the RpoS sigma factor and the Lon and ClpXP proteases involved in Mucts62 repressor inactivation by starvation and the Crp transcription factor required for a subsequent stage of the fusion process (Lamrani et al., 1999). In the widely-studied example of *lac33* frameshift reversion, the key regulatory event appears to be activation of F*lac* plasmid transfer and replication functions (Peters & Benson, 1995; Galitski & Roth, 1995; Radicella et al., 1995; Foster & Rosche, 1999). Hall (1999) describes further examples of adaptive mutation, in particular the activation of IS element insertion into the *ebgR* locus stimulated by selective conditions.

In yeast and higher organisms, there are several controls exerted over retrotransposable elements.

Transcription of Ty retrovirus-like elements is subject to control by the mating type locus (Errede et al., 1991), UV irradiation stimulates Ty transcription and activity (Bradshaw & McEntee, 1989), and Ty3 contains pheromone-response elements in its LTRs so that it is induced to transpose selectively during mating (Kinsey & Sandmeyer, 1995).

Hybrid dysgenesis, as studied in *Drosophila* (Bregliano & Kidwell, 1983; Engels, 1989; Finnegan, 1989; Kidwell & Evgen'ev, 1999) illustrates a particular kind of stress situation – matings between individuals from different populations or even from different species. This kind of stress, related to very small population sizes, may be particularly relevant to evolutionary crises. Hybrid dysgenesis involves both DNA transposons (e.g. P factors) and retrotransposons (e.g. LINE-like I elements). The active elements are stable in their normal host population, but can transpose at rates of over 100% when introduced into an egg cell from a naive population lacking active elements. The consequences are transpositions to multiple sites in the chromosomes of both strains, excisions from established sites, and chromosome rearrangements, like inversions. P factor activity is limited to the germ line by regulated splicing; in the germ line, all four exons encoding the active transposase are spliced together, while in somatic tissues only the first three exons are correctly spliced, leading to production of an inhibitory truncated version of the transposase protein. What is most notable about hybrid dysgenesis is that the multiple changes occur premeiotically in germ line development; thus, after several mitotic divisions, the clonal descendants of a single germinal cell can undergo meiosis to produce a group of gametes. Progeny formed from these gametes will constitute an interbreeding population sharing multiple genetic changes in their chromosomes.

Plants undergo transposable element activation after any one of a number of stresses, including wounding and exposure to fungal extracts (Costa et al., 1999), and some plants subject to chromosome breakage were found to rapidly reorganize the entire genome (McClintock, 1978). Finally, methylation is used not only in bacteria but also in fungi, plants and mammals to regulate the activity of various repeated sequences, including transposable elements (Bestor, 1999; Matzke, 1999).

Clearly, there is accumulating evidence that transposable elements respond to biological inputs via cellular control networks that determine the timing and extent of genetic change they cause. If transposable elements are significant agents of evolutionary reorganization of the genome in response to stress situations, then we should expect to find evidence of major episodic changes at the formation of new taxa. Accordingly, the DNA databases show major changes in the repetitive content of the genome between related taxa (e.g. SINES; Brosius, 1999b; Lerat et al., 1999; Roy et al., 1999).

Transposable elements, non-random genomic changes, and signal transduction

Being able to trigger genetic change in response to stress and other biological inputs in itself presents an important departure from temporal randomness in evolution. If an organism can turn on biochemical systems for genome reorganization when they are most needed, it has gained an important edge in the struggle for survival in a constantly changing biosphere. This advantage probably explains the ubiquity of natural genetic engineering systems in contemporary organisms, all of whose ancestors have undergone multiple episodes of evolutionary variation.

Transposable elements also represent a second kind of non-randomness in their movements through the genome. Even if the target sites lack specificity, it is far from a random event to move a defined segment of DNA hundreds or thousands of base pairs in length that carries transcription signals, coding sequences, splice sites, and binding sites for DNA bending proteins and other determinants. What makes transposable elements such effective and versatile mutagenic agents is their ability to modify and enhance as well as block the activity of genetic loci.

From an informational perspective, however, the most sophisticated use of transposable elements in evolution would occur if they could be guided to particular genomic locations. This would permit them to build up the kinds of integrated networks demonstrated in principle by McClintock (1956a, 1965), and coordinated movements of transposable elements would make it easier to understand how novel, multilocus adaptive systems came into being. One of the major challenges being hurled at evolutionary theory right now is the argument that Darwinian gradualism cannot explain the origin of complex integrated systems needed for adaptation or survival (e.g. Behe, 1996). If a plausible molecular mechanism for accomplishing rapid coordinated genetic changes throughout the genome could be found, then the potential of the

cell for creating functional multilocus systems would be enormously enhanced. Is there evidence that such guidance is possible? The initial answer, based on a number of observations in yeast and *Drosophila*, appears to be positive.

One of the most striking observations about the insertional specificity of the yeast Ty1–Ty4 retroviral-like elements is their strong preference for insertion just upstream of tRNA loci (Voytas & Boeke, 1993). It has been demonstrated in a Ty3 *in vitro* system, that this preference is mediated by direct interactions between RNA polymerase III transcription factors and the retrotransposon integrase (Kirchner et al., 1995). Thus, a direct molecular connection between transcription factors and the integration systems of transposable elements has been demonstrated. The yeast Ty5 element has quite a different insertional specificity; it inserts with a very high preference into transcriptionally silenced regions of the genome, such as telomeres and inactivated mating-type cassettes (Zou et al., 1996). This preference is hypothesized to result from an interaction of the integrase with silencing factors that build up a unique chromatin configuration in silent regions. If correct, this hypothesis says that Ty integration systems can interact specifically with different classes of transcription factor, an assertion that is perfectly reasonable given our current knowledge of protein–protein interactions and the formation of higher-order nucleoprotein complexes in DNA rearrangements (Echols, 1986).

In *Drosophila*, targeting has been observed with P factor constructs (DNA-based elements) used in making transgenic flies. The naturally isolated P factor already has strong specificity for certain genetic loci, such as *singed*. When fragments of several loci are inserted into P factor vectors, other strongly preferred insertion sites appear that are related to the inserted fragment. Including fragments of the *engrailed* (Hama et al., 1990) and *linotte* (Taillebourg & Dura, 1999) loci inside the terminal repeats of a P factor resulted in a very high proportion of insertions into those loci, but not at one particular site. In these two cases, some kind of homology recognition system may be at work (Taillebourg & Dura, 1999). In the case of the *polyhomeotic* (Ph) locus, however, a different result was obtained – a P factor construct containing small fragments of the Ph 5′ regulatory region went preferentially to chromosomal regions containing binding sites for the polyhomeotic and polycomb transcription factors (Fauvarque & Dura, 1993). Like the yeast data, this result also suggested that particular transcription factors could interact with the integration system of a transposable element and guide it to regions containing loci that are part of a regulatory network.

From these early results, which may well be only the tip of the iceberg, we see that transcriptional control proteins and transposable element integration proteins can interact. We have no problem accepting the idea that signal transduction networks involving transcription factors use protein–protein interactions to guide RNA polymerase and all its accessory factors to specific suites of genomic loci in response to a wide variety of biological inputs (Alberts et al., 1994; Gerhart & Kirschner, 1997). There is no reason to assume that protein–protein interactions cannot also occur between signal transduction molecules and transposable element DNA rearrangement proteins. Thus, at a molecular level, there is a plausible mechanism to explain how transposable elements could be targeted to a series of genetic loci whose products already function together. Differently targeted insertions could recruit new proteins into the system. Thus, at moments of extensive genome reorganization, the signal transduction/transposable element interaction can confer on cells a far higher probability of evolving useful new multi-locus systems, at least in their rudimentary forms. Such molecular mechanisms leading to coordinated changes at multiple locations in the genome may help solve the mystery of how complex evolutionary inventions arise in a perfectly natural way.

Summary

The thrust of this presentation has been to point out how the discovery of transposable elements as agents of genome restructuring has brought the question of evolutionary change into the realm of cell biology, where regulation and biological information processing are major factors. We are entering the next century with an increasingly computational view of cells and how they make important decisions. The argument here is that evolutionary change is not exempt from this new perspective. Evidence from a variety of systems indicates that transposable elements can interact in a molecularly plausible way with signal transduction networks, the key information processing entities in the cell. Biological feedback can play a critical role in genomic responses to emergencies (McClintock, 1984). Thus, organisms have a far

more powerful evolutionary potential to generate integrated genomic networks and ensure the survival of their descendants than predicted by current theories of gradualism and random mutation.

Acknowledgements

The ideas expressed here came out of my work on transposable elements and the ways they mediate DNA rearrangements. I am particularly grateful to the late Barbara McClintock for many insights and to my collaborators on phage Mu-mediated rearrangements, Genevieve Maenhaut-Michel, David Leach, and Pat Higgins. This research has been supported by NSF grant MCB-9513728.

References

Agrawal, A., Q.M. Eastman & D.G. Schatz, 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature, 394: 744–751.

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts & J.D. Watson, 1994. Molecular Biology of the Cell, 3rd edn., Garland, New York.

Avery, O.T., C.M. MacLeod & M. McCarty, 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. J. Exp. Med. 79: 137–157.

Behe, M., 1996. Darwin's Black Box: the biochemical challenge to evolution, Free Press, New York.

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

Bestor, T.H., 1999. Sex brings transposons and genomes into conflict. Genetica 107: 289–295.

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

Bradshaw, V.A. & K. McEntee,1989. DNA damage activates transcription and transposition of yeast Ty retrotransposons. Mol. Gen. Genet. 218: 465–474.

Bray, D., 1990. Intracellular signalling as a parallel distributed process. J. Theoret. Biol. 143: 215–231.

Bregliano, J.-C. & M.G. Kidwell, 1983. Hybrid dysgenesis determinants, pp. 363–410 in Mobile Genetic Elements, edited by J.A. Shapiro, Academic Press, New York.

Britten, R.J., 1997. Mobile elements inserted in the distant past have taken on important functions. Gene 205: 177–182.

Britten, R. J. & E.H. Davidson, 1969. Gene regulation for higher cells: a theory, Science 165: 349–357.

Britten, R.J. & D.E. Kohne, 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms, Science 161: 529–540.

Brosius, J., 1999a. RNAs from all categories generate retrosequences that may be accepted as novel genes or regulatory elements. Gene 238: 115–134.

Brosius, J., 1999b. Vertebrate genomes were forged by massive bombardments with retroelements and retrosequences. Genetica 107: 209–238

Costa, A.P.P., K.C. Scortecci, R.Y. Hashimoto, P.G. Araujo, M.-A. Grandbastien & M.-A. Van Sluys, 1999. Retrolycl-1, a member of the Tnt1 retrotransposon super-family in the *Lycopersicon peruvianum* genome. Genetica 107: 65–72.

Echols, H., 1986. Multiple DNA-protein interactions governing high-precision DNA transactions. Science 233: 1050–1056.

Engels, W.R., 1989. P elements in *Drosophila melanogaster*, pp. 437–484 in Mobile DNA, edited by D.E. Berg. and M.M. Howe, American Society for Microbiology, Washington, D.C.

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.

Fauvarque, M.O. & J.M. Dura, 1993. Polyhomeotic regulatory sequences induce developmental regulator-dependent variegation and targeted P-element insertions in *Drosophila*. Genes Dev. 7: 1508–1520.

Felsenfeld, G., J. Boyes, J. Chung, D. Clark & V. Studitsky, 1996. Chromatin structure and gene expression. Proc. Natl. Acad. Sci. USA 93: 9384–9388.

Finnegan, D.J., 1989. Eukaryotic transposable elements and genome evolution. Trends Genet. 5: 103–107.

Foster P.L., 1993. Adaptive mutation: the uses of adversity. Ann. Rev. Microbiol. 47: 467–504.

Foster, P.L. & W.A. Rosche, 1999. Increased episomal replication accounts for the high rate of adaptive mutation in recD mutants of *Escherichia coli*. Genetics 152: 15–30.

Galitski, T. & J.R. Roth, 1995. Evidence that F plasmid transfer replication underlies apparent adaptive mutation. Science 268: 421–423.

Gerhart, J. & M. Kirschner, 1997. Cells, Embryos & Evolution: toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability. Blackwell Science, Malden, Mass.

Green, M.M., 1988. Mobile DNA elements and spontaneous gene mutation, In Eukaryotic Transposable Elements as Mutagenic Agents, Banbury Report 30: 41–50.

Hall, B.G., 1999. Transposable elements as activators of cryptic genes in *E. coli.* Genetica 107: 181–187.

Hama C, Z. Ali & T.B. Kornberg, 1990. Region-specific recombination and expression are directed by portions of the Drosophila engrailed promoter. Genes Dev. 4: 1079–1093.

Hiom, K., M. Melek & M. Gellert, 1998. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell 94: 463–470.

Jacob, F. & J. Monod, 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3: 318–356.

Kassis, J.A., E. Noll, E.P. Vansickle, W.F. Odenwald & N. Perrimon, 1992. Altering the insertional specificity of a *Drosophila* transposable element. Proc. Natl. Acad. Sci. USA 89: 1919–1923.

Kenter, A. & R. Wuerffel, 1999. Immunoglobulin switch recombination may occur by a DNA end-joining mechanism. Annal. N.Y. Acad. Sci. 870: 206–217.

Kidwell, M.G. & M.B. Evgen'ev, 1999. How valuable are model organisms for transposable element studies? Genetica 107: 103–111.

Kinsey P.T. & S.B. Sandmeyer, 1995. Ty3 transposes in mating populations of yeast: a novel transposition assay for Ty3. Genetics 139: 81–94.

Kirchner, J., C.M. Connolly & S.B. Sandmeyer, 1995. Requirement of RNA polymerase III transcription factors for *in vitro* position-specific integration of a retrovirus like element. Science 267: 1488–1491.

Lamrani. S., C. Ranquet, M.-J. Gama, H. Nakai, J.A. Shapiro, A. Toussaint & G. Maenhaut-Michel, 1999. Starvation-induced Mucts62-mediated Coding Sequence Fusion: Roles for ClpXP, Lon, RpoS and Crp. Molec. Microbiol. 32, 327–343.

Lerat, E., F. Brunet, C. Bazin & P. Capy, 1999. Is the evolution of transposable elements modular? Genetica 107: 15–25.

Lewis, S.M., 1999. Evolution of immunoglobulin and T-cell receptor gene assembly. Annal. N.Y. Acad. Sci. 870: 58–67.

MacNab, R., 1992. Genetics and biogenesis of bacterial flagella. Annu. Rev. Genet. 26: 131–158.

Maenhaut-Michel, G. & J.A. Shapiro, 1994. The roles of starvation and selective substrates in the emergence of *araB-lacZ* fusion clones. EMBO J. 13: 5229–5239.

Maenhaut-Michel, G., C.E. Blake, D.R.F. Leach & J.A. Shapiro, 1997. Different structures of selected and unselected araB-lacZ fusions. Molec. Micro. 23: 1133–1146.

Matzke, M.A., M.F. Mette, W. Aufsatz, J. Jakowitsch & A.J.M. Matzke, 1999. Host defenses to parasitic sequences and the evolution of epigenetic control mechanisms. Genetica 107: 271–287.

Mazel, D, B. Dychinco, V.A. Webb & J. Davies, 1998. A distinctive class of integron in the *Vibrio cholerae* genome. Science 280: 605–608.

McClintock, B., 1950. The origin and behavior of mutable loci in maize. Proc. Natl. 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., 1953. Induction of instability at selected loci in maize. *Genetics* 38: 579–599.

McClintock, B., 1956a. Intranuclear systems controlling gene action and mutation. Brookhaven Symp. Biol. 8: 58–74.

McClintock, B., 1956b. 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., 1978. Mechanisms that rapidly reorganize the genome. Stadler Genetics Symp. 10: 25–48.

McClintock, B., 1984. Significance of responses of the genome to challenge. Science 226: 792–801.

Monod, J. & F. Jacob, 1961. Teleonomic mechanisms in cellular metabolism, growth and differentiation. Cold Spr. Harb. Symp. Quant. Biol. 26: 389–401.

Moran, J.V., 1999. Human L1 retrotransposition: insights and peculiarities learned from a cultured cell retrotransposition assay. Genetica 107: 39–51.

Pardue, M.-L. & P.G. DeBaryshe, 1999. *Drosophila* telomeres: two transposable elements with important roles in chromosomes. Genetica 107: 189–196.

Peters, J.E. & S.A. Benson. 1995. Redundant transfer of F' plasmids occurs between *Escherichia coli* cells during nonlethal selection. J. Bacteriol. 177: 847–850.

Radicella, J.P., P.U. Park & M.S. Fox, 1995. Adaptive mutation in *Escherichia coli*: A role for conjugation. Science 268: 418–420.

Recchia, G.D. & R. Hall, 1995. Gene cassettes: a new class of mobile element. Microbiology 141: 3015–3027.

Reznikoff, W.S., 1992. The lactose operon-controlling elements: a complex paradigm. Mol. Microbiol. 6: 2419–2422.

Roy, A.M., M.L. Carroll, D.H. Kass, S.V. Nguyen, A.-H. Salem, M.A. Batzer & P.L. Deininger, 1999. Recently integrated human Alu repeats: finding needles in the haystack. Genetica 107: 149–161.

Saier, M.H., Jr., T.M. Ramseier & J. Reizer, 1996. Regulation of carbon utilization, pp. 1325–1343 in *Escherichia coli* and Salmonella: Cellular and Molecular Biology, 2nd edn., edited by F.C. Neidhardt et al., ASM Press, Washington, D.C.

Shapiro, J.A. 1982, Changes in gene order and gene expression. Natl. Cancer Inst. Monograph 60: 87–110.

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

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

Shapiro, J.A., 1992. Natural genetic engineering in evolution. Genetica 86: 99–111.

Shapiro, J.A., 1995, Adaptive mutation: Who's really in the garden? *Science* 268: 373–374.

Shapiro, J.A., 1997. Genome organization, natural genetic engineering and adaptive mutation. Trends Genet. 13: 98–104.

Shapiro, J.A., 1999a. Natural genetic engineering, adaptive mutation and bacterial evolution, pp. 259–275, in Microbial Ecology and Infectious Disease, edited by E. Rosenberg, ASM Press, Washington.

Shapiro, J.A., 1999b. Genome system architecture and natural genetic engineering in evolution. Annal. N.Y. Acad. Sci. 870: 23–35.

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

Taillebourg, E. & J.M. Dura, 1999. A novel mechanism for P element homing in Drosophila. Proc. Natl. Acad. Sci. USA 96: 6856–6861.

Voytas, D.F. & J.D. Boeke, 1993. Yeast retrotransposons and tRNAs. Trends Genet. 9: 421–427.

Watson, J.D. & F.H.C. Crick, A structure for deoxyribose nucleic acids. Nature 171: 737–738.

Yuh, C.H., H. Bolouri & E.H. Davidson, 1998. Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. Science 279: 1896–1902

Zou, S., N. KE, J.M. Kim & D.F. Voytas, 1996. The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. Genes Dev. 10: 634–645.