

Mobile DNA and Evolution in the 21st Century

James A. Shapiro

Dept. of Biochemistry and Molecular Biology, University of Chicago, Gordon Center
for Integrative Science W123B, 929 E. 57th Street, Chicago, IL 60637
jsha@uchicago.edu

Running title: Evolution in the 21st Century

Keywords: Natural genetic engineering, genome restructuring, horizontal transfer,
symbiosis, whole genome duplication, hybridization.

Abbreviation: ncRNA – non-protein encoding RNA, NHEJ – non-homologous end-
joining

Abstract: Scientific history has a profound effect on theories of evolution. At the beginning of the 21st Century, molecular cell biology has revealed a dense structure of information-processing networks that use the genome as an interactive Read-Write (RW) memory system rather than an organism blueprint. Genome sequencing has documented the importance of mobile DNA activities and major genome restructuring events at key junctures in evolution: exon shuffling, changes in *cis*-regulatory sites, horizontal transfer, cell fusions, and whole genome doublings (WGDs). The natural genetic engineering functions that mediate genome restructuring are activated by multiple stimuli, in particular by events similar to those found in the DNA record: microbial infection and interspecific hybridization leading to the formation of allotetraploids. These molecular genetic discoveries, plus a consideration of how mobile DNA rearrangements increase the efficiency of generating functional genomic novelties, make it possible to formulate a 21st Century view of interactive evolutionary processes. This view integrates contemporary knowledge of the molecular basis of genetic change, major genome events in evolution, and stimuli that activate DNA restructuring with classical cytogenetic understanding about the role of hybridization in species diversification.

Introduction: Summary of the argument

The review assumes that readers of this journal are familiar with the actions of mobile DNA and other genome restructuring functions. It will try to integrate that familiarity into the historical development of evolutionary concepts and incorporate recent discoveries from genome sequencing. Just as our knowledge of mobile DNA has introduced new ways of thinking about hereditary change, the results of sequence analysis have documented several types of genome alterations at key places in evolutionary history which happened within a single generation and affected multiple cellular and organismal characters at the same time: horizontal transfer of large DNA segments, cell fusions and symbioses, and whole genome doublings (WGDs). These rapid multi-character changes are fundamentally different from the slowly accumulating small random variations postulated in Darwinian and neo-Darwinian theory. Cell mergers and WGDs are the kinds of events that activate mobile DNA and genome restructuring. In order to fully integrate the genomic findings with our knowledge of mobile DNA, we have to make use of information about the molecular regulation of mobile DNA activities as well as McClintock's view that cells respond to signs of danger, frequently restructuring their genomes as part of the response [1]. This regulatory/cognitive view of genome restructuring helps us to formulate reasonable hypotheses about two unresolved questions in evolutionary theory: (i) the connections between evolutionary change and ecological disruption and (ii) the origins of complex adaptive novelties at moments of macroevolutionary change.

The historical context for evolutionary ideas.

Since Darwin, three issues have been seen as central to formulating a coherent theory of evolutionary change:

- (i) descent with modification (i.e. the inheritance of novel characters),
- (ii) the origins of hereditary variation, and
- (iii) the operation of natural selection.

All evolutionists accept descent with modification as fundamental to the evolutionary process, but views towards issues (ii) and (iii) have depended on the existing state of biological knowledge in each historical period.

In the 19th Century, Darwin based his thinking on the observations of animal breeders and naturalists. Lacking detailed studies of inheritance, he postulated that change arose randomly as "numerous, successive, slight variations" [2]. Applying the uniformitarian principle he learned from Charles Lyell, his professor of geology [3], Darwin extrapolated that these small changes would accumulate over long periods of time under the guidance of natural selection to produce major adaptive characters, such as the eye, and eventually would lead to the branching off of new species. Thus, classical Darwinism was characterized by its gradualist view of change and ascribed the major role in adaptive innovation to the positive action of natural selection in sequentially favoring ever fitter variants.

In the 20th Century, evolutionists were confronted by an apparent contradiction between Darwinian gradualism and the abrupt changes in individual traits observed to undergo Mendelian segregations in genetic crosses. This contradiction was resolved at mid-century by the neo-Darwinian "Modern Synthesis" that integrated Darwinian gradualism with mathematical population genetics [4] [5]. Like Darwin, his neo-Darwinian followers postulated that the mutational process, which generated allelic variants of individual genes, has to be random in nature. In opposition to Lamarckian ideas, all possibility that organismal history could influence hereditary variation was

excluded. The primary role in determining evolutionary novelty remained with natural selection.

In the 21st Century, we have the legacy of more than five decades of molecular biology. Knowledge of DNA has allowed us to study the mutational process with nucleotide and phosphodiester bond precision [6]. Our DNA-based technology has made it possible to acquire a growing database of genome sequences that permit us to read the history of evolutionary events preserved in the nucleic acid and protein record. Molecular cell biology has uncovered sophisticated networks in all organisms that acquire information about external and internal conditions, transmit and process that information inside the cell, compute the appropriate biochemical or biomechanical response, and activate the molecules needed to execute that response. These information-processing networks are central to the systems biology perspective of the new century. Altogether, we have a radically different conceptual perspective on living organisms than did our predecessors. As a result, we need to ask how this new perspective affects our 21st Century understanding of the evolutionary process. Posing this question and outlining a provisional answer are the goals of this review.

Barbara McClintock: thinking about genome change as a cognitive response to challenge.

In addition to the discoveries of molecular biology, our 21st Century thinking benefits from another major strand of 20th Century research – McClintock’s cytogenetic studies that led her to recognize the internal capabilities cells possess to repair and restructure their genomes. Starting in the 1930s with X-ray-induced chromosome rearrangements, she analyzed how maize cells dealt with broken ends. These studies taught her that maize had the ability to detect broken ends, bring them together, and fuse them to generate novel chromosome structures, including deletions, inversions, translocations, and rings [7] [8] [9] [10] [8-11]. She also found that cells in the embryo, but not in the terminally differentiated endosperm, could “heal” a single broken end by addition of a telomere. In the course of exploiting these repair capabilities to generate deficiencies of maize chromosome IX, she made the discovery of transposable elements, for which she is best known today [12].

Although the general view is that McClintock’s discovery of transposition was most important for revealing a novel mechanism of genomic change, she herself placed the emphasis on another aspect of her work. In conversation, she would often say that she was far more interested in control than she was in transposition. By this, she meant that the ability of her maize plants to regulate expression and restructure their genomes in accordance with their needs was more significant than the mechanics of chromosome rearrangement. She was primarily interested in the sensory and decision-making (*i.e.* cognitive) capacities of cells with damaged genomes. As she expressed it at the conclusion of her Nobel Prize lecture: “In the future, attention undoubtedly will be centered on the genome, with greater appreciation of its significance as a highly sensitive organ of the cell that monitors genomic activities and corrects common errors, senses unusual and unexpected events, and responds to them, often by restructuring the genome” [1]. In the next section, we will see how prescient and compatible with molecular analysis her vision was.

The genome as a RW memory system, not an organism blueprint.

The pioneering molecular biologists expected to provide a firm physical-chemical basis for the traditional 20th Century view that genotype determines phenotype and that genotype changes accidentally during replication [13]. This

expectation of one-way cellular information transfer was articulated most succinctly in Crick's "Central Dogma of Molecular Biology" [14]. Even when Temin and Mizutani discovered reverse transcriptase [15], Crick insisted that the genome was the source of phenotypic information and that nucleic acids as a class were the basic information molecules of the cell [16].

A review of the past five decades of molecular cell biology, including analysis of how mobile DNA operates, leads to a dramatically different picture of cellular informatics [17] [18]. We realize that DNA by itself is inert. It must interact with other molecules for transcription, replication, transmission to daughter cells, and repair. DNA does not change by itself, even when damaged. Point mutations and DNA rearrangements depend upon the action of protein and RNA molecules, and many organisms undergo genome restructurings as a necessary part of the normal life cycle [19] [20] [21] [22]. A great deal of cellular information processing occurs outside the genome – e.g. transcript processing, protein processing and decoration, signal transduction – yet the results of this information processing also feed back onto the genome in the form of alterations in nucleoprotein complexes, chromatin configurations, subnucleoid or subnuclear localization, and sequence or structural changes. In other words, the cell is a multilevel information-processing entity, and the genome is only a part of the entire interactive complex.

We can summarize the change from the simple linear view of the Central Dogma to today's complex systems-based picture of cell informatics by writing out molecular information transfer events as sets of Boolean propositions (adapted from [18]):

- Crick's central dogma of molecular biology:
 1. DNA \implies 2X DNA
 2. DNA \implies RNA \implies Protein \implies Phenotype
- Contemporary picture of molecular information transfers:
 1. DNA + 0 \implies 0
 2. DNA + Protein + ncRNA \implies chromatin
 3. Chromatin + Protein + ncRNA \implies DNA replication, chromatin maintenance/reconstitution
 4. Protein + RNA + lipids + small molecules \implies signal transduction
 5. Chromatin + Protein + signals \implies RNA (primary transcript)
 6. RNA + Protein + ncRNA \implies RNA (processed transcript)
 7. RNA + Protein + ncRNA \implies Protein (primary translation product)
 8. Protein + nucleotides + Ac-CoA + SAM + sugars + lipids \implies Processed and decorated protein
 9. DNA + Protein \implies New DNA sequence (mutator polymerases)
 10. Signals + Chromatin + Protein \implies New DNA structure (DNA rearrangements subject to stimuli)
 11. RNA + Protein + chromatin \implies New DNA structure (retrotransposition, retroduction, retrohoming)
 12. Signals + chromatin + proteins + ncRNA + lipids \implies nuclear/nucleoid localization

SUMMARY: DNA + Protein + ncRNA + signals + other molecules \iff Genome Structure & Phenotype

A helpful analogy for the role of the genome in cellular informatics is as a Read-Write (RW) memory system. This is a fundamentally different idea from the

conventional 20th Century view of the genome as a Read-Only memory (ROM) subject to accidental change. DNA is a multivalent storage medium capable of holding information in nucleotide sequences, chemical modifications, and nucleoprotein complexes. In thinking about how the cell writes information back onto the genome, we can discriminate roughly three different time scales:

1. Within the cell cycle, where the formation and dissolution of transient nucleoprotein complexes predominate;
2. Over several cell cycles, where heritable chromatin configurations can be passed on and then erased or re-imprinted;
3. Over evolutionary time, where sequence variation and genome restructuring play major roles in the emergence of novel characters and adaptive functions.

In this review, our focus is evolution. So the most appropriate question is to ask about the lessons we have learned from genome sequencing. There are many, and we discuss the ones that are most relevant to the action of mobile DNA.

What genomes teach I: protein evolution by exon shuffling.

From the first experiments clarifying the composite structure of lambda repressor in the late 1970s [23], it has become increasingly clear that proteins are largely composed of independently folding and functional domains [24]. At the start of the 21st Century, the *Nature* paper reporting the draft human genome contained two figures illustrating the way transcription factor and chromatin binding proteins have changed by domain accretion and swapping as they evolved from yeast to mammals [25]. The emergence of novel domains and protein evolution by a combinatorial process of domain shuffling are now widely recognized as the major routes to functionally novel molecules. It is of fundamental conceptual significance that the genomic basis of domain-swapping involves the rearrangement of coding segments (exons) rather than the sequential accumulation of random single base/single amino acid changes. Mobile DNA movements rather than replication errors serve as the primary engines of protein evolution.

Both laboratory experiments and genome sequences have documented roles for well-characterized mobile elements in the origin of novel exons and in exon shuffling. By examining the sequences encoding known proteins without applying Repeat-Masker methods, Nekrutenko and Li discovered that over 4% of human mRNAs come from SINE retrotransposons [26]. Since then, additional examples of “exonization” of segments within mobile element and cDNA inserts have accumulated in the literature [27] [28] [29]. Incorporation of sequences from mobile elements into spliced transcripts typically produces exons that encode oligopeptides. Thus, we can recognize well-defined mobile DNA events (transposon or retrotransposon insertions) that are capable of rapidly generating the extended sequences needed to encode novel protein domains. In particular cases, transposase sequences have been exapted to encode DNA binding domains [30]. Since the genomic content of mobile elements is taxonomically specific [31], we may expect to see differences between phylogenetic branches in the new exons they produce.

There are well-documented cases in the DNA record where mobile element systems have served to mobilize, amplify and rearrange exons. The most striking case involves the more than 3000 Pack-MULEs discovered in the rice genome [32]. These composite MULEs (“Mu-like elements”) have inverted terminal repeats flanking

combinations of exons and introns. In many cases, the Pack-MULE at a particular location is flanked by a short target site duplication indicating that it arrived by a transposition mechanism. Some Pack-MULEs contain complete protein coding sequences, a number of which are duplicated in the rice genome. Many Pack-MULEs, however, contain exons lacking translation initiation or termination signals, and there are known rice mRNAs that contain spliced exons from more than one adjacent Pack-MULE. Helitrons played an analogous role in the maize genome [33]. Intriguingly, although helitrons are present in the *Arabidopsis* and rice genomes, they are far less active in exon capture in those two species than in maize [34].

In addition to DNA transposition, there is both genomic and experimental evidence for exon shuffling by LINE element retrotransduction. Retrotransduction occurs when LINE element transcription reads through the 3' polyA signal and produces RNA and cDNA molecules containing downstream sequences from the genome. Such read-through retrotransduction events are found in 15% of all human LINE1 inserts and may account for fully 1% of the human genome [35]. Exon-shuffling by LINE1 retrotransduction occurs in tissue culture cells [36] and has been documented in the evolution of primate genomes [37]. Further mechanisms of exon shuffling may occur when LINE elements introduce double-strand (DS) breaks into a genetic locus [38] or are involved in homologous exchanges between nearby repeats [39].

What genomes teach II: mobile elements and regulatory evolution.

Transcription signals. The appearance of novel coding capacity at a genetic locus frequently results from changes in *cis*-acting regulatory and processing signals without any change in exon content. Mobile DNA has long been known to play a role in this kind of regulatory change. The phenotypes of the first bacterial mutations known to be IS element insertions resulted either from the acquisition of transcriptional stop signals [40] or from the creation of novel transcriptional start sites [41]. In eukaryotes, mutations activating transcription most commonly resulted from insertion of enhancer elements in LTR retroelements [42]. In the case of one apoptosis regulator protein, genome comparison shows that orthologous coding regions in primates and rodents acquired their parallel transcription signals from independent LTR retrotransposon insertions [43]. Sequences of Mu element insertions in maize can alter both the initiation and termination sites for transcription [44]. Examination of the human genome has uncovered over 100 cases where *Alu* elements provided polyA additional signals at the 3' end of expressed sequences [45]. The role of mobile elements in the evolution of transcriptional regulatory sites has been extensively documented from genomic data since the 1990s [46] [47]. Many of these cases display the kind of taxonomic specificity predicted by the phylogenetic distribution of transposons and retrotransposons [48].

Splicing signals. It has been over two decades since Wessler and colleagues discovered the splicing of *Ds* insertions in maize [49]. Not only does *Ds* behave as a mobile intron; it also confers alternative splicing [50]. The same is true of maize retrotransposons [51]. The potential of a single genomic change to encode multiple novel products has been documented in broad beans, where insertion of a CACTA family transposon carries out exon shuffling and provides sites for alternative splicing [52]. Recent studies in the human genome are beginning to clarify the requirements for generating novel splicing patterns by mobile element inserts [53-55].

Chromatin signals. The insertion of a mobile element has a profound effect on local chromatin configuration. Since a major regulatory mechanism for controlling

the activity of mobile elements is incorporation into silenced chromatin [56], individual or clustered elements serve as nucleation sites for heterochromatin domains [57]. Some elements, like *gypsy* in *Drosophila*, carry chromatin insulator determinants that are major contributors to their influence on genome expression [58]. In certain cases, like the FWA and MEDEA loci in *Arabidopsis*, imprinted expression reflects the action of RNAi machinery on sequences derived from a mobile element [59]. Recent studies of imprinted loci in *Arabidopsis* seeds indicates that mobile elements provided many of the recognition sequences for epigenetic control [60].

The connection between mobile elements and chromatin signals is less well-documented in mammals. Nonetheless, there is intriguing evidence that retrotransposons were critical to the origin of an epigenetic control regime necessary for the emergence of mammals in evolution. Knockout experiments in mice show that imprinted loci derived from the Ty3/gypsy retrotransposon family are essential to placental development [61] [62]. These observations suggest that functional exaptation of retrotransposon coding sequences and signals mediating their epigenetic control played a role in the evolution of the placenta, a major developmental invention.

Regulatory RNAs. We are currently learning how much regulation occurs through the action of small RNA molecules. Examination of plant genome sequences has established important links of many small RNAs to DNA transposons (MITES) [63] and led to the suggestion that si- and miRNA regulation evolved from mobile element controls [64]. The rice Pack-MULEs are also associated with small RNA coding sequences [65]. In the human genome, 55 functionally characterized and 85 uncharacterized miRNAs arose from transposons and retrotransposons [66]. Comparison with the mouse genome indicates that miRNAs matching L2 LINE and MIR SINE elements are ancient and conserved, while those matching L1 LINE and DNA elements are primate-specific. As expected from the taxonomic distribution of SINE elements [31], the *Alu*-derived miRNAs are also primate-specific [67]. *Alu* element recombination also appears to have played a role in the expansion of primate miRNA coding arrays [68]. A similar conclusion about the role of mobile elements in the generation of taxonomically-specific miRNAs arose from analysis of marsupial genomes [69].

Regulatory suites encompassing unlinked coding regions. One major aspect of regulatory evolution by mobile elements was illustrated by McClintock in her 1956 Brookhaven Symposium paper on intranuclear systems [70]. This is the ability of related elements to insert at two or more distinct loci and bring them under coordinate regulation. That coregulated loci have arisen in this way during evolution has been documented in mice, where similar retroviral promoters initiate transcription of different loci in oocytes and preimplantation embryos [71]. In the human genome, taxonomically-restricted evolution of the vertebrate REST-controlled transcriptional network has involved LINE element insertions into cis-regulatory sites [72]. It would clearly be of great interest to correlate genome expression data with a survey of loci that share regulatory sequences evolved from related mobile elements.

What genomes teach III: Intercellular horizontal DNA transfer.

Molecular genetics began with the study of intercellular horizontal DNA transfer. The first demonstration of the genetic capacity of DNA molecules involved Pneumococcal transformation [73], and bacterial genetics developed on the basis of cells' capacities to transfer genome segments by transformation, conjugation or viral transduction [74]. Studies of temperate bacteriophages and antibiotic resistance made

us appreciate the multiple molecular mechanisms cells have to incorporate newly acquired DNA independently of extensive sequence homology [75]. From countless experiments, we now have overwhelming evidence for horizontal DNA transfer between species and between the three kingdoms of living cells (Table 1).

Table 1. Modes of intercellular and interkingdom DNA transfer

Horizontal transfer mode	Documented transfers
Uptake of environmental DNA	Bacteria – bacteria [74] [76] [77] Plant – bacteria [78] [79] Plastid transfection [80] Mammalian cell transfection [81] [82] and lipofection [83] [84]
Conjugal transfer	Bacteria – bacteria [74] Bacteria – yeast [85] Bacteria – plant [86] [87]
Viral transduction and GTAs (gene transfer agents)	Bacteria – bacteria [74] [88] Bacteria – plant [89] Animal cell – animal cell [90] Animal cell – virus [91]
Parasitic or endosymbiotic association	Plant – fern [92] Plant – plant [93] [94] Bacteria – invertebrate [95]
Undetermined mechanism	Archaea – bacteria [96]

Horizontal transfer can be a major driver of evolutionary novelty because it permits the acquisition of DNA encoding complex traits in a single event. The genomic data is overwhelming in documenting the fundamental importance of horizontal transfer in the evolution of bacterial and archaeal genomes [96]. Prokaryotic genomes contain plasmids and genomic islands encoding multi-component adaptive characters that range from microbicide resistance [74] [75], virulence [97] [98], and symbiosis [99] to metabolism [100] and magnetotaxis [101]. This has led to a scheme of bacterial and archaeal evolution that has a reticular rather than a branching structure [102]. The possibility that different genome components could display different phylogenies due to horizontal transfer [103] was quite literally inconceivable to Darwin and his mid-20th Century neo-Darwinian successors.

Although we have long been familiar with the prokaryotic story, there is rapidly growing evidence for intercellular and interkingdom horizontal transfer events in the evolutionary history of eukaryotic genomes [104]. The data include phylogenetically dispersed coding sequences [105] and mobile elements [106] [107] [108] as well as incorporation of genomic segments from prokaryotic and eukaryotic endosymbionts [95] and parasites [94]. There is also evidence of host to parasite transfer [93]. In certain microbially diverse ecosystems, such as the rumen, frequent prokaryote to eukaryote transfer occurs [109]. In plants, but not animals, there is extensive horizontal transfer of mitochondrial DNA [110]. Similar transfer is very rarely seen in the plastids [111], and this may be explained because the mitochondria have a DNA uptake system not found in chloroplasts [112]. The functional consequences of horizontal transfer into eukaryotes range from acquisition of single biochemical activities to major restructuring of metabolism [113] to integrating multiple functions needed to occupy new ecologies, as illustrated by fungal pathogens

[114], the anaerobic human parasites *Entamoeba histolytica* and *Trichomonas vaginalis* [115], and plant parasitic nematodes [116].

What genomes teach IV: Cell fusions and intracellular DNA transfer at key junctures in eukaryotic evolution.

One of the early accomplishments of nucleic acid sequencing was to confirm the endosymbiotic origin of mitochondria and plastids [117]. Combined with evidence that the mitochondrion is an ancestral character for all eukaryotes [118], this confirmation places cell fusion events at the root of eukaryotic evolution [119]. For photosynthetic eukaryotes, the original cyanobacterial fusion that generated the ancestral plastid has been followed by a series of secondary symbioses between various eukaryotic lineages and either red or green algae [120]. The most “basal” photosynthetic lineage appears to be the glaucophytes because their plastids retain bacterial peptidoglycans [121]. Through evidence of cell fusions and endosymbiosis, genome sequencing has introduced another major process of rapid and multi-character change into the established evolutionary record. Lacking knowledge of cell biology, such a mechanism of variation was not considered by Darwin and has been largely ignored by his neo-Darwinian followers.

As the following descriptions of various endosymbioses show, DNA mobility between distinct genome compartments was a major feature of adjustment to cell fusion events. Sequence evidence indicates that all the cell fusions in eukaryotic lineages were followed by massive episodes of intracellular horizontal DNA transfer between the organelle and nuclear genomes [119] [122] [123]. That is why the majority of organelle proteins are encoded by the nuclear genome. Moreover, these organelle genomes are remarkably dynamic in their evolution. Mitochondria display a great range of genome size (~6 kb to ~480 kb), and a number of them have strikingly elaborate DNA structures (*e.g.* multiple linear molecules, interlocked circles) and/or modes of expression [124]. There are anaerobic eukaryotes that have lost the oxidative functions of mitochondria, but most of them retain related organelles labelled hydrogenosomes or mitosomes [118].

The history of plastids, descended from cyanobacteria, is somewhat different from that of mitochondria, descended from alpha-protobacteria. In higher plants and photosynthetic algae, the chloroplast genome is relatively stable and falls within a relatively narrow size range of 120 – 160 kb [125]. In heterotrophic or parasitic species that have lost photosynthesis, the plastid genome is reduced but still retained at sizes greater than 34 kb (Table 2) [125] [126]. In the apicomplexan parasites, plastid genomes are known to have undergone extensive structural rearrangements [127]. Non-photosynthetic chloroplast derivatives appear to retain residual functions, such as encoding tRNAs that may be used by mitochondria, activities involved in the biosynthesis of amino acids, fatty acids, isoprenoids, heme, and pigments, and enzymes for detoxifying oxidative radicals [128].

In cells of organisms arising from secondary symbioses with red algae (cryptomonads) or green algae (chlorarachniophytes), there are actually four distinct genome compartments: nucleus, mitochondrion, plastid and nucleomorph (the descendant of the algal nucleus) [129]. The plastid and nucleomorph compartments are surrounded by four rather than two membranes, presumably reflecting their origins by phagocytosis. The two sequenced nucleomorph genomes are 551 kb (*Guillardia theta*, cryptomonad) and 373 kb (*Bigeloviella natans*, chlorarachniophyte), each containing three chromosomes with telomeres. These genomes encode their own

18S eukaryotic ribosomal RNA, other RNAs and proteins (465 and 293, respectively). The nuclear genomes of both species contain coding sequences of red- or green-algal origin, indicating extensive intracellular horizontal transfer [130].

In addition to the remarkable multi-genome cells just described, there are cases of tertiary symbioses in the dinoflagellates, which have fused with green algae, haptophytes, diatoms and cryptomonads [131]. It appears from analyzing the origins of nuclear coding sequences for plastid-targeted proteins that dinoflagellates and other chromalveolates have retained an ability to phagocytose other cells and recruit fragments of their genomes, while that capacity was lost in the photosynthetic lineages leading to green algae, plants and red algae [132].

Table 2. Plastid genome sizes in photosynthetic organisms and their non-photosynthetic relatives [125]

Angiosperms	Plastid genome size (# coding sequences)
<i>Nicotiana tabacum</i>	155.9 kb (146)
<i>Arabidopsis</i>	154 kb
<i>Oryzae</i>	135 kb
<i>Epifagus virginiana</i> (holoparasite)	70.0 kb (53)
Trebouxiophytes (green algae)	
<i>Chlorella vulgaris</i>	150.6 kb (209)
<i>Helicosporidium sp.</i> (invertebrate pathogen)	37.5 kb (54)
<i>Prototheca wickerhamii</i> (saprophytic algal pathogen)	54.1 kb
Euglenids (flagellates)	
<i>Euglena gracilis</i>	143.2 kb (128)
<i>Euglena longa</i> (heterotroph)	73.3 kb (84)
Apicomplexans (containing secondarily acquired red algal plastids)	
<i>Odontella sinensis</i> (photosynthetic diatom)	119.7 kb (175)
<i>Plasmodium falciparum</i> (malarial parasite)	34.7 kb (68)
<i>Toxoplasma gondii</i> (cat parasite)	35.0 kb (65)
<i>Eimeria tenella</i> (poultry pathogen)	34.8 kb (65)
<i>Theileria parva</i> (bovine pathogen)	39.6 kb (71)

What genomes teach V: Whole genome doublings at key places in eukaryotic evolution.

Genome sequencing has made it clear how important the amplification and modification of various genome components has been. Of particular interest has been the formation of families of coding elements for homologous proteins within genomes. Both prokaryote and eukaryote species encode characteristic protein families, which are important guides to the functions those species need to thrive in their particular ecological niches. As complete genome sequences accumulated, it became apparent that not only the genetic loci encoding individual proteins had amplified; large chromosome regions had also undergone duplication processes. These “syntenic” regions carry genetic loci in the same order and orientation. By comparing related taxa, it has been possible to discern phylogenetic branches that have inherited two copies of multiple ancestral segments. These segments are now understood to be the remnants of whole genome doubling (WGD) events at the base of the branch.

Genome doublings have been documented in yeasts [133] [134], ciliated protozoa [135], and plants [136]. There is even evidence of a genome tripling at the base of the angiosperm radiation (Darwin’s “abominable mystery”¹) [138]. In animals, the most important WGD events have been found at the base of the vertebrate lineage, where two successive events gave rise first to all vertebrates and then to jawed vertebrates [139]. This “2R” double WGD event was originally postulated by Ohno in his 1970 book on the essential role of duplications in evolution [140]. Later in vertebrate evolution, there is another WGD event at the origin of teleost fish [139] [141]. Characteristic of transitions marked by WGD events are the rapid formation of a cluster of related species, as in *Paramecium* [135], or the appearance of major innovations, as with the vertebrate skeleton [142] and jaw [139]. WGD is yet another evolutionary process outside the Darwinist perspective that occurs suddenly (i.e. within a single generation) and simultaneously affects multiple phenotypic characters [143]. It is especially significant to note that a genome doubling means that the dispersed coding elements for complex circuits are duplicated, and the two duplicate circuits can then undergo independent modifications as distinct entities [144].

There is an important connection between WGD and synthetic speciation. It is possible to generate new species of plants by interspecific hybridization and genome doubling [136] [145] [146] [147] [148] [149]. Fertile hybrids tend to have tetraploid genomes [146]. Genome doubling helps maintain stability through meiosis because each chromosome in the hybrid has an homologous partner for pairing and crossing over. There is also evidence that genome doubling helps maintain normal transcription patterns [150]. The genome duplication events may occur either during gametogenesis or after fertilization, but in plants the most common process involves diploid gametes [151]. The incidence of spontaneous genome doubling is surprisingly high, reaching 1% of all fertilizations in mice [152].

It is of great theoretical significance that synthetic speciation takes place rapidly after hybridization rather than slowly following repeated selections, as predicted by conventional theory. The evolutionary importance of interspecific hybridization in promoting evolutionary change has been appreciated since a time

¹ Letter to J. D. Hooker, July 22, 1879: Darwin described the rapid rise and early diversification within the angiosperms as “an abominable mystery.” 137. Darwin F, Seward, AC, eds. : *More Letters of Charles Darwin* London: John Murray; 1903.

predating the molecular genetics revolution [153] [154]. Although most synthetic and observational work has been done with plants [155], there are reports of contemporary natural hybridization involving animals [156] [157]. The animal cases include Darwin's finches in the Galapagos Islands [158], long taken as a paradigm of gradualist evolution. The finch case is especially instructive because hybridization leads to abrupt, unpredictable changes in beak shape [159].

Responses of mobile DNA systems to infection, hybridization and genome duplications.

The genomic evidence showing that cell fusions and WGD have occurred at key junctures in eukaryotic evolutionary phylogenies leads to the question of what effect such events (plus the related process of interspecific hybridization) have on mobile DNA and natural genetic engineering functions. The answer is that all these processes are major triggers of genomic instability and restructuring, with microbial infection serving as a proxy for cell fusions [160] [161]. The data on hybridization responses are more extensive in plants (Table 3), but we have enough cases in animals to be confident that the answer there is equally valid (Table 4). Moreover, we know of many cases of hybrid dysgenesis in animals, where activation of mobile elements and widespread genomic changes results from inter-population mating [162] [163] [164] [165]. In at least one intriguing plant case, interspecific mating has triggered genomic instability with formation of a zygote containing only one of the parental genomes [166].

The rapid natural genetic engineering response to genome doubling reflects a tendency to return to the normal diploid state. This poorly understood process of "diploidization" involves chromosome loss, deletions, and chromosome rearrangements [167]. The chief mechanistic basis for activation of natural genetic engineering in response to hybridization and genome doubling appears to be changes in chromatin organization and in epigenetic modifications of the DNA that normally inhibit activity of mobile elements (Tables 3 & 4) [168] [169] [170] [171].

Table 3. Genomic responses to changes in ploidy and interspecific hybridization in plants.

Taxon	Genomic response
Asteraceae (Compositae)	Genome expansion and retrotransposon proliferation in sunflower hybrids [172]
	Chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species [173]
	Rapid chromosome evolution in polyploids [174]
Grasses	Altered methylation patterns and chromosome restructuring in hybrids [175]
Potato	Genome instability in hybrids [176]
	Phenotypic and epigenetic variability in a diploid hybrid [177]
Nicotiana spp. (tobacco)	Elimination of repeated DNA in a synthetic allotetraploid [178]

Rice	Extensive genomic variability induced by introgression from wild rice [179]
	LTR retrotransposon movements in rice lines introgressed by wild rice [180]
	Retrotransposon activation following introgression [181]
Brassica	Rapid genome change in synthetic polyploids [182]
	Large scale chromosome restructuring [183]
Wheat [184]	Sequence loss and cytosine methylation following hybridization and allopolyploidy [185]
	Rapid genome evolution following allopolyploidy [186]
	Changes in methylation patterns of mobile elements in allopolyploids [187]
	Parental repeat elimination in newly synthesized allopolyploids [188]
	Rapid genomic changes in interspecific and intergeneric hybrids and allopolyploids [189]
Arabidopsis	Genomic changes in synthetic polyploids [190]
	Chromosome rearrangements after allotetraploid formation [191]

Table 4. Genomic responses to hybridization in animals.

Taxon	Genomic Response
<i>Drosophila</i> spp.	Increased retrotransposition in interspecific hybrid [192]
<i>Macropus</i> marsupials	Centromere instability in interspecific hybrids [193]
Wallabies	Loss of retroelement methylation and chromosome remodeling in interspecific hybrids [194]
Mouse	Amplification and double minutes in a hybrid [195]
Rice fish (<i>medaka</i>)	Chromosome elimination in an interspecific hybrid [196]

The evolutionary advantages of searching genome space by natural genetic engineering.

One of the traditional objections to Darwinian gradualism has been that it is too slow and indeterminate a process to account for natural adaptations, even allowing for long periods of random mutation and selection. A successful random walk through the virtually infinite dimensions of possible genome configurations simply has too low a probability of success [197]. Is there a more efficient way for cells to search “genome space” and increase their probability of hitting upon useful new DNA structures? There is, and the underlying molecular mechanisms utilize the demonstrated capabilities of mobile DNA and other natural genetic engineering systems [198] [199].

Perhaps the most important aspect of evolutionary change by natural genetic engineering is that it employs a **combinatorial** search process based upon DNA modules that already possess functionality. The evolutionary reuse of functional components has been recognized for many years [200, 201], but it is only with genome sequencing that we have come to appreciate how fundamental and virtually ubiquitous such reuse is. A well-established engineering principle is to build new structures to meet specific requirements by rearranging proven, existing components, as in mechanical structures and electronic circuits. The evolution of proteins by domain accretion and shuffling is one example of an analogous process. Mixing functional domains in new combinations is far more likely to produce a protein with novel activities than is modification one amino acid at a time. Single amino acid changes are more suitable for modulating existing functional properties (*e.g.* ligand binding and allosteric responses) than for generating capabilities that did not previously exist. In addition to the combinatorial search by shuffling of existing exons, greater variability results from the formation of novel exons. We do not yet know a great deal about any biases that may exist in the exonization process. If it is correct to postulate that new functional exons arise by the exaptation of segments of mobile DNA, such as SINE elements, then it will be worthwhile to investigate the coding content of these elements to see if there is any tendency favoring sequences that encode useful folded polypeptide structures.

The second major aspect of evolutionary change by natural genetic engineering is that it generally takes place after an activating event that produces what McClintock called a “genome shock” [202]. Activating events include loss of food [18], infection and interspecific hybridization (Tables 3 and 4) – just the events that we can infer from the geological and genomic records to have happened repeatedly in life history. Episodic activation of natural genetic engineering functions means that alterations to the genome occur in bursts rather than as independent events. Thus, novel adaptations that require changes at multiple locations in the genome can arise within a single generation and can produce progeny expressing all the changes at once. There is no requirement, as in conventional theory, that each individual change be beneficial by itself. The episodic occurrence of natural genetic engineering bursts also makes it very easy to understand the punctuated pattern of the geological record [203]. Moreover, the nature of activating challenges provides a comprehensible link to periodic disruptions in earth history. Geological upheavals that perturb an existing ecology are likely to lead to starvation, alteration of host-parasite relationships, and unusual mating events between individuals from depleted populations.

A particular instance of the potential for stress-activated natural genetic engineering to produce complex novelties is the exaptation of an existing functional network following its duplication by WGD. Domains may be added to various

proteins in the network to allow them to interact with a novel set of input and output molecules. In addition, insertions of connected regulatory signals at the cognate coding regions can generate a new transcriptional control circuit that may allow the modified network to operate under different conditions from its progenitor.

The idea that genomic restructuring events may be integrated functionally to operate coordinately at a number of distinct loci encoding components of a regulatory network may seem extremely unlikely. But the basic requirement for such integration is the ability to target DNA changes to co-regulated regions of the genome. Precisely this kind of targeting has been demonstrated for mobile elements in yeast, where retrotransposon integration activities interact with transcription [204] or chromatin [205] factors, and in *Drosophila*, where P elements can be engineered to “home” in on loci regulated by particular regulatory proteins [206]. In addition, we know that mobile element insertion can be coupled with replication [207] and DNA restructuring with transcription [208]. Naturally, the feasibility of such multi-locus functional integration of genome changes remains to be demonstrated in the laboratory. Fortunately, the experiments are straightforward; we can use appropriately engineered transposons and retrotransposons to search for coordinated multilocus mutations after activation. Clearly, the subject of functionally targeted changes to the genome belongs on the 21st Century mobile DNA research agenda.

Conclusion: A 21st Century view of evolutionary change.

Our ability to think fruitfully about the evolutionary process has greatly expanded thanks to studies of mobile DNA. Laboratory studies of plasmids, transposons, retrotransposons, NHEJ systems, reverse transcription, antigenic variation in prokaryotic and eukaryotic pathogens, lymphocyte rearrangements, and genome reorganization in ciliated protozoa have all made it possible to provide mechanistic explanations for events documented in the historical DNA record [6]. We know that processes similar to those we document in our experiments have been major contributors to genome change in evolution. Using our knowledge of genome restructuring mechanisms, we can generate precise models to account for many duplications, amplifications, dispersals and rearrangements observed at both the genomic and proteomic levels.

The genome DNA record also bears witness to sudden changes that affect multiple characters at once: horizontal transfer of large DNA segments, cell fusions, and WGDs. These data are not readily compatible with earlier gradualist views on the nature of evolutionary variation. However, we are now able to apply the results of findings on the regulation of natural genetic engineering functions in the laboratory and in the field to make sense of the DNA record. Cell fusions and WGDs are events we know to activate DNA restructuring functions (Tables 3 and 4). Thus, it is not surprising that bursts of intracellular horizontal transfer, genome reduction and genome rearrangement follow these initial abrupt changes in the cell’s DNA. How a newly symbiotic cell or one with a newly doubled genome manages the transition to a stable genome structure that replicates and transfers reliably at cell division is another important subject for future research. The lessons we learn about silencing mobile DNA by internal deletion [12] and RNA-directed chromatin modification [209] are likely to prove helpful starting points.

Although there remain many gaps in our knowledge, we are now in a position to outline a distinctively 21st Century scenario for evolutionary change. The scenario includes the following elements:

- (1) Hereditary variation arises from the non-random action of built-in biochemical systems that mobilize DNA and carry out natural genetic engineering.
- (2) Major disruptions of an organism's ecology trigger cell and genome restructuring. The ecological disruptions can act directly, through stress on individuals, or indirectly, through changes in the biota that favor unusual interactions between individuals (cell fusions, interspecific hybridizations). Triggering events continue until a new ecology has emerged that is filled with organisms capable of utilizing the available resources.
- (3) Ecologically-triggered cell and genome restructurings produce organisms which, at some frequency, will possess novel adaptive features that suit the altered environment. Novel adaptive features can be complex from the beginning because they result from processes that operate on pre-existing functional systems, whose components can be amplified and rearranged in new combinations. Competition for resources (purifying selection) serves to eliminate those novel system architectures that are not functional in the new ecology.
- (4) Once ecological stability has been achieved, natural genetic engineering functions are silenced, the tempo of innovation abates, and microevolution can occur to fine-tune recent evolutionary inventions through successions of minor changes.

This 21st Century scenario assumes a major role for the kind of cellular sensitivities and genomic responses emphasized by McClintock in her 1984 Nobel Prize address [1]. Such a cognitive component is absent from conventional evolutionary theory because 19th and 20th Century evolutionists were not sufficiently knowledgeable about cellular response and control networks. This 21st Century view of evolution establishes a reasonable connection between ecological changes, cell and organism responses, widespread genome restructuring, and the rapid emergence of adaptive inventions. It also answers the objections to conventional theory raised by intelligent design advocates because evolution by natural genetic engineering has the capacity to generate complex novelties. In other words, our best defense against anti-science obscurantism comes from the study of mobile DNA because that is the subject that has most significantly transformed evolution from natural history into a vibrant empirical science.

CONFLICTS OF INTEREST: The author has no conflicts of interest.

References

1. McClintock B: **Significance of responses of the genome to challenge.** *Science* 1984, **226**:792-801.
2. Darwin C: *Origin of Species*. London: John Russel; 1859.
3. Lyell C: *Principles of Geology*. Edinburgh: John Murray; 1830.
4. [http://en.wikipedia.org/wiki/Modern_evolutionary_synthesis]
5. Huxley J: *Evolution: the modern synthesis*. London: Allen & Unwin; 1942.

6. Craig NL, Craigie R, Gellert M, Lambowitz AM: *Mobile DNA II*. Washington, DC: American Society of Microbiology Press; 2002.
7. McClintock B: **A Correlation of Ring-Shaped Chromosomes with Variegation in Zea Mays**. *Proc Nat Acad Sci USA* 1932, **18**:677-681.
8. McClintock B: **The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes**. *Genetics* 1938, **23**:315-376.
9. McClintock B: **The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis**. *Proc Nat Acad Sci USA* 1939, **25**:405-416.
10. McClintock B: **The Stability of Broken Ends of Chromosomes in Zea Mays**. *Genetics* 1941, **26**:234-282.
11. McClintock B: **The Fusion of Broken Ends of Chromosomes Following Nuclear Fusion**. *Proc Nat Acad Sci USA* 1942, **28**:458-463.
12. McClintock B: *Discovery And Characterization of Transposable Elements: The Collected Papers of Barbara McClintock*. New York: Garland; 1987.
13. Judson H: *The Eighth Day of Creation: Makers of the Revolution in Biology*. New York: Simon & Schuster; 1979.
14. Crick F: **On protein synthesis**. *Symp Soc Exp Biol* 1958, **12**:138-163.
15. Temin H, S Mizutani: **RNA-dependent DNA polymerase in virions of Rous sarcoma virus**. *Nature* 1970, **226**:1211-1213.
16. Crick F: **Central dogma of molecular biology**. *Nature* 1970, **227**:561-563.
17. Shapiro JA: **Genome informatics: The role of DNA in cellular computations**. *Biological Theory* 2006, **1**:288-301.
18. Shapiro JA: **Revisiting the central dogma in the 21st Century**. *Ann NY Acad Sci* 2009, **in press**.
19. Kunkel B, Losick R, Stragier P: **The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene**. *Genes Dev* 1990, **4**:525-535.
20. Barbour AG, Restrepo BI: **Antigenic variation in vector-borne pathogens**. *Emerg Infect Dis* 2000, **6**:449-457.
21. Gellert M: **V(D)J recombination: RAG proteins, repair factors, and regulation**. *Ann Rev Biochem* 2002, **71**:101-132.
22. Garnier O, Serrano V, Duharcourt S, Meyer E: **RNA-mediated programming of developmental genome rearrangements in Paramecium tetraurelia**. *Mol Cell Biol* 2004, **24**:7370-7379.
23. Pabo C, Sauer, RT, Sturtevant, JM, Ptashne, M: **The lambda repressor contains two domains**. *Proc Nat Acad Sci USA* 1979, **76**:1608-1612.
24. Doolittle RF: **The multiplicity of domains in proteins**. *Annu Rev Biochem* 1995, **64**:287-314.
25. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al: **Initial sequencing and analysis of the human genome**. *Nature* 2001, **409**:860-921.
26. Nekrutenko A, Li WH: **Transposable elements are found in a large number of human protein-coding genes**. *Trends Genet* 2001, **17**:619-621.
27. Sorek R: **The birth of new exons: mechanisms and evolutionary consequences**. *Rna* 2007, **13**:1603-1608.
28. Britten RJ: **Transposable elements have contributed to thousands of human proteins**. *Proc Nat Acad Sci USA* 2006, **103**:1798-1803.

29. Bowen NJ, Jordan IK: **Exaptation of protein coding sequences from transposable elements.** *Genome Dyn* 2007, **3**:147-162.
30. Cordaux R, Udit S, Batzer MA, Feschotte C: **Birth of a chimeric primate gene by capture of the transposase gene from a mobile element.** *Proc Natl Acad Sci U S A* 2006, **103**:8101-8106.
31. Sternberg RV SJ: **How repeated retroelements format genome function.** *Cytogenet Genome Res* 2005, **110**:108-116.
32. Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR: **Pack-MULE transposable elements mediate gene evolution in plants.** *Nature* 2004, **431**:569-573.
33. Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A: **Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize.** *Nat Genet* 2005, **37**:997-1002.
34. Sweredoski M, DeRose-Wilson L, Gaut BS: **A comparative computational analysis of nonautonomous helitron elements between maize and rice.** *BMC Genomics* 2008, **9**:467.
35. Pickeral OK, Makalowski W, Boguski MS, Boeke JD: **Frequent human genomic DNA transduction driven by LINE-1 retrotransposition.** *Genome Res* 2000, **10**:411-415.
36. Moran JV, DeBerardinis RJ, Kazazian HH, Jr.: **Exon shuffling by L1 retrotransposition.** *Science* 1999, **283**:1530-1534.
37. Ejima Y, Yang L: **Trans mobilization of genomic DNA as a mechanism for retrotransposon-mediated exon shuffling.** *Hum Mol Genet* 2003, **12**:1321-1328.
38. Wheelan SJ, Aizawa Y, Han JS, Boeke JD: **Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution.** *Genome Res* 2005, **15**:1073-1078.
39. Han K, Lee J, Meyer TJ, Remedios P, Goodwin L, Batzer MA: **L1 recombination-associated deletions generate human genomic variation.** *Proc Natl Acad Sci U S A* 2008, **105**:19366-19371.
40. De Crombrughe B AS, Gottesman M, Pastan I: **Effect of Rho on transcription of bacterial operons.** *Nature New Biology* 1973, **241**:260-264.
41. Pilacinski W, Mosharrafa E, Edmundson R, Zissler J, Fiandt M, Szybalski W: **Insertion sequence IS2 associated with int-constitutive mutants of bacteriophage lambda.** *Gene* 1977, **2**:61-74.
42. Errede B, Cardillo TS, Wever G, Sherman F, Stiles JJ, Friedman LR, Sherman F: **Studies on transposable elements in yeast. I. ROAM mutations causing increased expression of yeast genes: their activation by signals directed toward conjugation functions and their formation by insertion of Ty1 repetitive elements. II. deletions, duplications, and transpositions of the COR segment that encompasses the structural gene of yeast iso-1-cytochrome c.** *Cold Spring Harb Symp Quant Biol* 1981, **45 Pt 2**:593-607.
43. Romanish MT, Lock WM, van de Lagemaat LN, Dunn CA, Mager DL: **Repeated recruitment of LTR retrotransposons as promoters by the anti-apoptotic locus NAIP during mammalian evolution.** *PLoS Genet* 2007, **3**:e10.
44. Cui X, Hsia AP, Liu F, Ashlock DA, Wise RP, Schnable PS: **Alternative transcription initiation sites and polyadenylation sites are recruited during Mu suppression at the rf2a locus of maize.** *Genetics* 2003, **163**:685-698.

45. Chen C, Ara T, Gautheret D: **Using Alu elements as polyadenylation sites: A case of retroposon exaptation.** *Mol Biol Evol* 2009, **26**:327-334.
46. Britten RJ: **Cases of ancient mobile element DNA insertions that now affect gene regulation.** *Mol Phylogenet Evol* 1996, **5**:13-17.
47. Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, Chew JL, Ruan Y, Wei CL, Ng HH, Liu ET: **Evolution of the mammalian transcription factor binding repertoire via transposable elements.** *Genome Res* 2008, **18**:1752-1762.
48. Marino-Ramirez L, Lewis KC, Landsman D, Jordan IK: **Transposable elements donate lineage-specific regulatory sequences to host genomes.** *Cytogenet Genome Res* 2005, **110**:333-341.
49. Wessler SR, Baran G, Varagona M: **The maize transposable element Ds is spliced from RNA.** *Science* 1987, **237**:916-918.
50. Wessler SR: **The maize transposable Ds1 element is alternatively spliced from exon sequences.** *Mol Cell Biol* 1991, **11**:6192-6196.
51. Varagona MJ, Purugganan M, Wessler SR: **Alternative splicing induced by insertion of retrotransposons into the maize waxy gene.** *Plant Cell* 1992, **4**:811-820.
52. Zabala G, Vodkin L: **Novel exon combinations generated by alternative splicing of gene fragments mobilized by a CACTA transposon in Glycine max.** *BMC Plant Biol* 2007, **7**:38.
53. Lev-Maor G, Ram O, Kim E, Sela N, Goren A, Levanon EY, Ast G: **Intronic Alus influence alternative splicing.** *PLoS Genet* 2008, **4**:e1000204.
54. Gal-Mark N, Schwartz S, Ast G: **Alternative splicing of Alu exons--two arms are better than one.** *Nucleic Acids Res* 2008, **36**:2012-2023.
55. Schwartz S, Gal-Mark N, Kfir N, Oren R, Kim E, Ast G: **Alu exonization events reveal features required for precise recognition of exons by the splicing machinery.** *PLoS Comput Biol* 2009, **5**:e1000300.
56. Girard A, Hannon GJ: **Conserved themes in small-RNA-mediated transposon control.** *Trends Cell Biol* 2008, **18**:136-148.
57. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavigne K, Mittal V, May B, Kasschau KD, et al: **Role of transposable elements in heterochromatin and epigenetic control.** *Nature* 2004, **430**:471-476.
58. Gerasimova TI, Byrd K, Corces VG: **A chromatin insulator determines the nuclear localization of DNA.** *Mol Cell* 2000, **6**:1025-1035.
59. Kinoshita Y, Saze H, Kinoshita T, Miura A, Soppe WJ, Koornneef M, Kakutani T: **Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats.** *Plant J* 2007, **49**:38-45.
60. Gehring M, Bubb KL, Henikoff S: **Extensive demethylation of repetitive elements during seed development underlies gene imprinting.** *Science* 2009, **324**:1447-1451.
61. Ono R, Nakamura K, Inoue K, Naruse M, Usami T, Wakisaka-Saito N, Hino T, Suzuki-Migishima R, Ogonuki N, Miki H, et al: **Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality.** *Nat Genet* 2006, **38**:101-106.
62. Sekita Y, Wagatsuma H, Nakamura K, Ono R, Kagami M, Wakisaka N, Hino T, Suzuki-Migishima R, Kohda T, Ogura A, et al: **Role of retrotransposon-derived imprinted gene, Rtl1, in the fetomaternal interface of mouse placenta.** *Nat Genet* 2008, **40**:243-248.

63. Kuang H, Padmanabhan C, Li F, Kamei A, Bhaskar PB, Ouyang S, Jiang J, Buell CR, Baker B: **Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: new functional implications for MITEs.** *Genome Res* 2009, **19**:42-56.
64. Obbard DJ, Gordon KH, Buck AH, Jiggins FM: **The evolution of RNAi as a defence against viruses and transposable elements.** *Philos Trans R Soc Lond B Biol Sci* 2009, **364**:99-115.
65. Hanada K, Vallejo V, Nobuta K, Slotkin RK, Lisch D, Meyers BC, Shiu SH, Jiang N: **The functional role of pack-MULEs in rice inferred from purifying selection and expression profile.** *Plant Cell* 2009, **21**:25-38.
66. Piriyaopongsa J, Marino-Ramirez L, Jordan IK: **Origin and evolution of human microRNAs from transposable elements.** *Genetics* 2007, **176**:1323-1337.
67. Lehnert S, Van Loo P, Thilakarathne PJ, Marynen P, Verbeke G, Schuit FC: **Evidence for co-evolution between human microRNAs and Alu-repeats.** *PLoS One* 2009, **4**:e4456.
68. Zhang R, Wang YQ, Su B: **Molecular evolution of a primate-specific microRNA family.** *Mol Biol Evol* 2008, **25**:1493-1502.
69. Devor EJ, Peek AS, Lanier W, Samollow PB: **Marsupial-specific microRNAs evolved from marsupial-specific transposable elements.** *Gene* 2009.
70. McClintock B: **Intranuclear systems controlling gene action and mutation.** *Brookhaven Symp Biol* 1956:58-74. .
71. Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB: **Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos.** *Dev Cell* 2004, **7**:597-606.
72. Johnson R, Gamblin RJ, Ooi L, Bruce AW, Donaldson IJ, Westhead DR, Wood IC, Jackson RM, Buckley NJ: **Identification of the REST regulon reveals extensive transposable element-mediated binding site duplication.** *Nucleic Acids Res* 2006, **34**:3862-3877.
73. Avery OT, C.M. MacLeod, M. McCarty: **Studies on the chemical nature of the substance inducing transformation of Pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus Type III.** *J Exp Med* 1944, **79**:137-158.
74. Hayes W: *The Genetics of Bacteria and their Viruses (2nd ed.)*. London: Blackwell; 1968.
75. Bukhari AI, J.A. Shapiro, and S. L. Adhya (Eds.): *DNA insertion elements, plasmids and episomes*. Cold Spring Harbor, New York: Cold Spring Harbor Press; 1977.
76. Averhoff B, Friedrich A: **Type IV pili-related natural transformation systems: DNA transport in mesophilic and thermophilic bacteria.** *Arch Microbiol* 2003, **180**:385-393.
77. Claverys JP, Prudhomme M, Martin B: **Induction of competence regulons as a general response to stress in gram-positive bacteria.** *Annu Rev Microbiol* 2006, **60**:451-475.
78. Kay E, Vogel TM, Bertolla F, Nalin R, Simonet P: **In situ transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria.** *Appl Environ Microbiol* 2002, **68**:3345-3351.

79. de Vries J, Herzfeld T, Wackernagel W: **Transfer of plastid DNA from tobacco to the soil bacterium *Acinetobacter* sp. by natural transformation.** *Mol Microbiol* 2004, **53**:323-334.
80. O'Neill C, Horvath GV, Horvath E, Dix PJ, Medgyesy P: **Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery systems.** *Plant J* 1993, **3**:729-738.
81. Wolff JA, Budker V: **The mechanism of naked DNA uptake and expression.** *Adv Genet* 2005, **54**:3-20.
82. Khalil IA, Kogure K, Akita H, Harashima H: **Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery.** *Pharmacol Rev* 2006, **58**:32-45.
83. Felgner PL, T.R. Gadek, M. Holm, R. Roman, H.W. Chan, M. Wenz, J.P. Northrop, G.M. Ringold and M. Danielsen **Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure.** *Proc Nat Acad Sci USA* 1987, **84**:7413-7417.
84. Zuhorn IS, Kalicharan R, Hoekstra D: **Lipoplex-mediated transfection of mammalian cells occurs through the cholesterol-dependent clathrin-mediated pathway of endocytosis.** *J Biol Chem* 2002, **277**:18021-18028.
85. Heinemann JA, Sprague GF, Jr.: **Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast.** *Nature* 1989, **340**:205-209.
86. Winans SC: **Two-way chemical signaling in *Agrobacterium*-plant interactions.** *Microbiol Rev* 1992, **56**:12-31.
87. Broothaerts W, Mitchell HJ, Weir B, Kaines S, Smith LM, Yang W, Mayer JE, Roa-Rodriguez C, Jefferson RA: **Gene transfer to plants by diverse species of bacteria.** *Nature* 2005, **433**:629-633.
88. Stanton TB: **Prophage-like gene transfer agents-novel mechanisms of gene exchange for *Methanococcus*, *Desulfovibrio*, *Brachyspira*, and *Rhodobacter* species.** *Anaerobe* 2007, **13**:43-49.
89. Chung SM, Vaidya M, Tzfira T: ***Agrobacterium* is not alone: gene transfer to plants by viruses and other bacteria.** *Trends Plant Sci* 2006, **11**:1-4.
90. El-Aneed A: **An overview of current delivery systems in cancer gene therapy.** *J Control Release* 2004, **94**:1-14.
91. Filee J, Pouget N, Chandler M: **Phylogenetic evidence for extensive lateral acquisition of cellular genes by Nucleocytoplasmic large DNA viruses.** *BMC Evol Biol* 2008, **8**:320.
92. Davis CC, Anderson WR, Wurdack KJ: **Gene transfer from a parasitic flowering plant to a fern.** *Proc Biol Sci* 2005, **272**:2237-2242.
93. Davis CC, Wurdack KJ: **Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales.** *Science* 2004, **305**:676-678.
94. Mower JP, Stefanovic S, Young GJ, Palmer JD: **Plant genetics: gene transfer from parasitic to host plants.** *Nature* 2004, **432**:165-166.
95. Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, Torres MC, Giebel JD, Kumar N, Ishmael N, Wang S, et al: **Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes.** *Science* 2007, **317**:1753-1756.
96. Koonin EV, Makarova KS, Aravind L: **Horizontal gene transfer in prokaryotes: quantification and classification.** *Annu Rev Microbiol* 2001, **55**:709-742.

97. Hacker J, Carniel E: **Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes.** *EMBO Rep* 2001, **2**:376-381.
98. Kado CI: **Horizontal gene transfer: sustaining pathogenicity and optimizing host-pathogen interactions.** *Mol Plant Pathol* 2009, **10**:143-150.
99. Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS, et al: **The genome of *Rhizobium leguminosarum* has recognizable core and accessory components.** *Genome Biol* 2006, **7**:R34.
100. Fournier G: **Horizontal gene transfer and the evolution of methanogenic pathways.** *Methods Mol Biol* 2009, **532**:163-179.
101. Ullrich S, Kube M, Schubbe S, Reinhardt R, Schuler D: **A hypervariable 130-kilobase genomic region of *Magnetospirillum gryphiswaldense* comprises a magnetosome island which undergoes frequent rearrangements during stationary growth.** *J Bacteriol* 2005, **187**:7176-7184.
102. Doolittle WF: **Phylogenetic classification and the universal tree.** *Science* 1999, **284**:2124-2129.
103. Beiko RG, Doolittle WF, Charlebois RL: **The impact of reticulate evolution on genome phylogeny.** *Syst Biol* 2008, **57**:844-856.
104. Keeling PJ, Palmer JD: **Horizontal gene transfer in eukaryotic evolution.** *Nat Rev Genet* 2008, **9**:605-618.
105. Andersson JO: **Lateral gene transfer in eukaryotes.** *Cell Mol Life Sci* 2005, **62**:1182-1197.
106. Bartolome C, Bello X, Maside X: **Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes.** *Genome Biol* 2009, **10**:R22.
107. Pace JK, 2nd, Gilbert C, Clark MS, Feschotte C: **Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods.** *Proc Natl Acad Sci U S A* 2008, **105**:17023-17028.
108. Roulin A, Piegu B, Fortune PM, Sabot F, D'Hont A, Manicacci D, Panaud O: **Whole genome surveys of rice, maize and sorghum reveal multiple horizontal transfers of the LTR-retrotransposon Route66 in Poaceae.** *BMC Evol Biol* 2009, **9**:58.
109. Ricard G, McEwan NR, Dutilh BE, Jouany JP, Macheboeuf D, Mitsumori M, McIntosh FM, Michalowski T, Nagamine T, Nelson N, et al: **Horizontal gene transfer from Bacteria to rumen Ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment.** *BMC Genomics* 2006, **7**:22.
110. Bergthorsson U, Adams KL, Thomason B, Palmer JD: **Widespread horizontal transfer of mitochondrial genes in flowering plants.** *Nature* 2003, **424**:197-201.
111. Keeling PJ: **Role of horizontal gene transfer in the evolution of photosynthetic eukaryotes and their plastids.** *Methods Mol Biol* 2009, **532**:501-515.
112. Koulintchenko M, Konstantinov Y, Dietrich A: **Plant mitochondria actively import DNA via the permeability transition pore complex.** *Embo J* 2003, **22**:1245-1254.
113. Rogers M, Keeling PJ: **Lateral transfer and re compartmentalization of Calvin cycle enzymes of plants and algae.** *J Mol Evol* 2004, **58**:367-375.

114. Richards TA, Dacks JB, Jenkinson JM, Thornton CR, Talbot NJ: **Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms.** *Curr Biol* 2006, **16**:1857-1864.
115. Alsmark UC, Sicheritz-Ponten T, Foster PG, Hirt RP, Embley TM: **Horizontal gene transfer in eukaryotic parasites: a case study of *Entamoeba histolytica* and *Trichomonas vaginalis*.** *Methods Mol Biol* 2009, **532**:489-500.
116. Mitreva M, Smant G, Helder J: **Role of horizontal gene transfer in the evolution of plant parasitism among nematodes.** *Methods Mol Biol* 2009, **532**:517-535.
117. Gray MW: **Origin and evolution of organelle genomes.** *Curr Opin Genet Dev* 1993, **3**:884-890.
118. Embley TM: **Multiple secondary origins of the anaerobic lifestyle in eukaryotes.** *Philos Trans R Soc Lond B Biol Sci* 2006, **361**:1055-1067.
119. Embley TM, Martin W: **Eukaryotic evolution, changes and challenges.** *Nature* 2006, **440**:623-630.
120. Lane CE, Archibald JM: **The eukaryotic tree of life: endosymbiosis takes its TOL.** *Trends Ecol Evol* 2008, **23**:268-275.
121. Pfanzagl B, Zenker A, Pittenauer E, Allmaier G, Martinez-Torrecuadrada J, Schmid ER, De Pedro MA, Löffelhardt W: **Primary structure of cyanelle peptidoglycan of *Cyanophora paradoxa*: a prokaryotic cell wall as part of an organelle envelope.** *J Bacteriol* 1996, **178**:332-339.
122. Kleine T, Maier UG, Leister D: **DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis.** *Annu Rev Plant Biol* 2009, **60**:115-138.
123. Martin W: **Gene transfer from organelles to the nucleus: frequent and in big chunks.** *Proc Natl Acad Sci U S A* 2003, **100**:8612-8614.
124. Burger G, Gray MW, Lang BF: **Mitochondrial genomes: anything goes.** *Trends Genet* 2003, **19**:709-716.
125. Krause K: **From chloroplasts to "cryptic" plastids: evolution of plastid genomes in parasitic plants.** *Curr Genet* 2008, **54**:111-121.
126. Lopez-Juez E, Pyke KA: **Plastids unleashed: their development and their integration in plant development.** *Int J Dev Biol* 2005, **49**:557-577.
127. de Koning AP, Keeling PJ: **The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured.** *BMC Biol* 2006, **4**:12.
128. Barbrook AC, Howe CJ, Purton S: **Why are plastid genomes retained in non-photosynthetic organisms?** *Trends Plant Sci* 2006, **11**:101-108.
129. Archibald JM: **Nucleomorph genomes: structure, function, origin and evolution.** *Bioessays* 2007, **29**:392-402.
130. Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG: **Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms.** *Mol Biol Evol* 2006, **23**:2413-2422.
131. Hackett J, Anderson, DM, Erdner, DL, Bhattacharya, D: **Dinoflagellates: A remarkable evolutionary experiment.** *Am J Bot* 2004, **91**:1523-1534.
132. Nosenko T, Lidie KL, Van Dolah FM, Lindquist E, Cheng JF, Bhattacharya D: **Chimeric plastid proteome in the Florida "red tide" dinoflagellate *Karenia brevis*.** *Mol Biol Evol* 2006, **23**:2026-2038.
133. Wolfe KH, Shields DC: **Molecular evidence for an ancient duplication of the entire yeast genome.** *Nature* 1997, **387**:708-713.

134. Scannell DR, Butler G, Wolfe KH: **Yeast genome evolution--the origin of the species.** *Yeast* 2007, **24**:929-942.
135. Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, Segurens B, Daubin V, Anthouard V, Aich N, et al: **Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*.** *Nature* 2006, **444**:171-178.
136. Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, et al: **Widespread genome duplications throughout the history of flowering plants.** *Genome Res* 2006, **16**:738-749.
137. Darwin F, Seward, AC, eds. : *More Letters of Charles Darwin* London: John Murray; 1903.
138. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, et al: **The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla.** *Nature* 2007, **449**:463-467.
139. Nakatani Y, Takeda H, Kohara Y, Morishita S: **Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates.** *Genome Res* 2007, **17**:1254-1265.
140. Ohno S: *Evolution by Gene Duplication* London: George Allen and Unwin; 1970.
141. Semon M, Wolfe KH: **Rearrangement rate following the whole-genome duplication in teleosts.** *Mol Biol Evol* 2007, **24**:860-867.
142. Zhang G, Cohn MJ: **Genome duplication and the origin of the vertebrate skeleton.** *Curr Opin Genet Dev* 2008, **18**:387-393.
143. Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C: **Major ecological transitions in wild sunflowers facilitated by hybridization.** *Science* 2003, **301**:1211-1216.
144. Veron AS, Kaufmann K, Bornberg-Bauer E: **Evidence of interaction network evolution by whole-genome duplications: a case study in MADS-box proteins.** *Mol Biol Evol* 2007, **24**:670-678.
145. Soltis PS, Soltis DE: **The role of hybridization in plant speciation.** *Annu Rev Plant Biol* 2009, **60**:561-588.
146. Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF: **Evolutionary genetics of genome merger and doubling in plants.** *Annu Rev Genet* 2008, **42**:443-461.
147. Wilson A: **Wheat and rye hybrids.** *Edinburgh Botanical Society Transactions* 1876, **12**:286-288.
148. Meister G: **Natural hybridization of wheat and rye in Russia.** *Journal of Heredity* 1921, **12**:467-470.
149. Ma XF, Gustafson JP: **Allopolyploidization-accommodated genomic sequence changes in triticale.** *Ann Bot (Lond)* 2008, **101**:825-832.
150. Hegarty MJ, Barker GL, Brennan AC, Edwards KJ, Abbott RJ, Hiscock SJ: **Changes to gene expression associated with hybrid speciation in plants: further insights from transcriptomic studies in *Senecio*.** *Philos Trans R Soc Lond B Biol Sci* 2008, **363**:3055-3069.
151. Thompson J, Lumaret, R: **The evolutionary dynamics of polyploid plants: origins, establishment and persistence.** *Trends Ecol Evol* 1992, **7**:302-307.

152. Eakin GS, Behringer RR: **Tetraploid development in the mouse.** *Dev Dyn* 2003, **228**:751-766.
153. Anderson E, Stebbins, G.L., Jr. : **Hybridization as an evolutionary stimulus.** *Evolution* 1954, **8**:378–388.
154. Arnold ML: **Transfer and origin of adaptations through natural hybridization: were Anderson and Stebbins right?** *Plant Cell* 2004, **16**:562-570.
155. Otto SP, Whitton J: **Polyploid incidence and evolution.** *Annu Rev Genet* 2000, **34**:401-437.
156. Dowling TS, CL: **The role of hybridization and introgression in the diversification of animals.** *Annu Rev Ecol Syst* 1997, **28**:593–619.
157. Arnold ML: **Natural hybridization and the evolution of domesticated, pest and disease organisms.** *Mol Ecol* 2004, **13**:997-1007.
158. Grant P, Grant, BR: **Hybridization of bird species.** *Science* 1992, **256**:193–197.
159. Grant PR, Grant BR: **Unpredictable evolution in a 30-year study of Darwin's finches.** *Science* 2002, **296**:707-711.
160. Frank O, Jones-Brando L, Leib-Mosch C, Yolken R, Seifarth W: **Altered transcriptional activity of human endogenous retroviruses in neuroepithelial cells after infection with *Toxoplasma gondii*.** *J Infect Dis* 2006, **194**:1447-1449.
161. Kimura Y, Tosa Y, Shimada S, Sogo R, Kusaba M, Sunaga T, Betsuyaku S, Eto Y, Nakayashiki H, Mayama S: **OARE-1, a Ty1-copia retrotransposon in oat activated by abiotic and biotic stresses.** *Plant Cell Physiol* 2001, **42**:1345-1354.
162. Bregliano J, Kidwell, M: **Hybrid dysgenesis.** In *Mobile Genetic Elements*. Edited by Shapiro J. New York: Academic Press,; 1983: 363–410.
163. Bucheton A: **I transposable elements and I-R hybrid dysgenesis in *Drosophila*.** *Trends Genet* 1990, **6**:16-21.
164. Periquet G, Lemeunier F, Bigot Y, Hamelin MH, Bazin C, Ladeveze V, Eeken J, Galindo MI, Pascual L, Boussy I: **The evolutionary genetics of the hobo transposable element in the *Drosophila melanogaster* complex.** *Genetica* 1994, **93**:79-90.
165. Vrana PB, Fossella JA, Matteson P, del Rio T, O'Neill MJ, Tilghman SM: **Genetic and epigenetic incompatibilities underlie hybrid dysgenesis in *Peromyscus*.** *Nat Genet* 2000, **25**:120-124.
166. Wang H, Chai Y, Chu X, Zhao Y, Wu Y, Zhao J, Ngezahayo F, Xu C, Liu B: **Molecular characterization of a rice mutator-phenotype derived from an incompatible cross-pollination reveals transgenerational mobilization of multiple transposable elements and extensive epigenetic instability.** *BMC Plant Biol* 2009, **9**:63.
167. Ma XF, Gustafson JP: **Genome evolution of allopolyploids: a process of cytological and genetic diploidization.** *Cytogenet Genome Res* 2005, **109**:236-249.
168. Vrana PB, Guan XJ, Ingram RS, Tilghman SM: **Genomic imprinting is disrupted in interspecific *Peromyscus* hybrids.** *Nat Genet* 1998, **20**:362-365.
169. Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J, Comai L: **Remodeling of DNA methylation and phenotypic and transcriptional**

- changes in synthetic Arabidopsis allotetraploids. *Plant Physiol* 2002, **129**:733-746.**
170. Josefsson C, Dilkes B, Comai L: **Parent-dependent loss of gene silencing during interspecies hybridization.** *Curr Biol* 2006, **16**:1322-1328.
 171. Brown JD, Golden D, O'Neill RJ: **Methylation perturbations in retroelements within the genome of a Mus interspecific hybrid correlate with double minute chromosome formation.** *Genomics* 2008, **91**:267-273.
 172. Ungerer MC, Strakosh SC, Stimpson KM: **Proliferation of Ty3/gypsy-like retrotransposons in hybrid sunflower taxa inferred from phylogenetic data.** *BMC Biol* 2009, **7**:40.
 173. Lai Z, Nakazato T, Salmaso M, Burke JM, Tang S, Knapp SJ, Rieseberg LH: **Extensive chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species.** *Genetics* 2005, **171**:291-303.
 174. Lim KY, Soltis DE, Soltis PS, Tate J, Matyasek R, Srubarova H, Kovarik A, Pires JC, Xiong Z, Leitch AR: **Rapid chromosome evolution in recently formed polyploids in Tragopogon (Asteraceae).** *PLoS One* 2008, **3**:e3353.
 175. Salmon A, Ainouche ML, Wendel JF: **Genetic and epigenetic consequences of recent hybridization and polyploidy in Spartina (Poaceae).** *Mol Ecol* 2005, **14**:1163-1175.
 176. Marfil CF, Masuelli RW, Davison J, Comai L: **Genomic instability in Solanum tuberosum x Solanum kurtzianum interspecific hybrids.** *Genome* 2006, **49**:104-113.
 177. Marfil CF, Camadro EL, Masuelli RW: **Phenotypic instability and epigenetic variability in a diploid potato of hybrid origin, Solanum ruiz-lealii.** *BMC Plant Biol* 2009, **9**:21.
 178. Skalicka K, Lim KY, Matyasek R, Matzke M, Leitch AR, Kovarik A: **Preferential elimination of repeated DNA sequences from the paternal, Nicotiana tomentosiformis genome donor of a synthetic, allotetraploid tobacco.** *New Phytol* 2005, **166**:291-303.
 179. Wang YM, Dong ZY, Zhang ZJ, Lin XY, Shen Y, Zhou D, Liu B: **Extensive de Novo genomic variation in rice induced by introgression from wild rice (Zizania latifolia Griseb.).** *Genetics* 2005, **170**:1945-1956.
 180. Shen Y LX, Shan X, et al.: **Genomic rearrangement in endogenous long terminal repeat retrotransposons of rice lines introgressed by wild rice (Zizania latifolia Griseb.).** *JOURNAL OF INTEGRATIVE PLANT BIOLOGY* 2005, **47**:998-1008.
 181. Liu B, Wendel JF: **Retrotransposon activation followed by rapid repression in introgressed rice plants.** *Genome* 2000, **43**:874-880.
 182. Song K, Lu P, Tang K, Osborn TC: **Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution.** *Proc Natl Acad Sci U S A* 1995, **92**:7719-7723.
 183. Kantama L, Sharbel TF, Schranz ME, Mitchell-Olds T, de Vries S, de Jong H: **Diploid apomicts of the Boechera holboellii complex display large-scale chromosome substitutions and aberrant chromosomes.** *Proc Natl Acad Sci U S A* 2007, **104**:14026-14031.
 184. Feldman M, Levy AA: **Allopolyploidy--a shaping force in the evolution of wheat genomes.** *Cytogenet Genome Res* 2005, **109**:250-258.
 185. Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA: **Sequence elimination and cytosine methylation are rapid and reproducible**

- responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 2001, **13**:1749-1759.**
186. Ozkan H, Levy AA, Feldman M: **Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group.** *Plant Cell* 2001, **13**:1735-1747.
 187. Dong YZ, Liu ZL, Shan XH, Qiu T, He MY, Liu B: **Allopolyploidy in wheat induces rapid and heritable alterations in DNA methylation patterns of cellular genes and mobile elements.** *Genetika* 2005, **41**:1089-1095.
 188. Han F, Fedak G, Guo W, Liu B: **Rapid and repeatable elimination of a parental genome-specific DNA repeat (pGc1R-1a) in newly synthesized wheat allopolyploids.** *Genetics* 2005, **170**:1239-1245.
 189. Han FP, Fedak G, Ouellet T, Liu B: **Rapid genomic changes in interspecific and intergeneric hybrids and allopolyploids of Triticeae.** *Genome* 2003, **46**:716-723.
 190. Madlung A, Tyagi AP, Watson B, Jiang H, Kagochi T, Doerge RW, Martienssen R, Comai L: **Genomic changes in synthetic *Arabidopsis* polyploids.** *Plant J* 2005, **41**:221-230.
 191. Pontes O, Neves N, Silva M, Lewis MS, Madlung A, Comai L, Viegas W, Pikaard CS: **Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome.** *Proc Natl Acad Sci U S A* 2004, **101**:18240-18245.
 192. Labrador M, Farre M, Utzet F, Fontdevila A: **Interspecific hybridization increases transposition rates of *Osvaldo*.** *Mol Biol Evol* 1999, **16**:931-937.
 193. Metcalfe CJ, Bulazel KV, Ferreri GC, Schroeder-Reiter E, Wanner G, Rens W, Obergfell C, Eldridge MD, O'Neill RJ: **Genomic instability within centromeres of interspecific marsupial hybrids.** *Genetics* 2007, **177**:2507-2517.
 194. O'Neill RJ, O'Neill MJ, Graves JA: **Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid.** *Nature* 1998, **393**:68-72.
 195. Brown JD, Strbuncelj M, Giardina C, O'Neill RJ: **Interspecific hybridization induced amplification of *Mdm2* on double minutes in a *Mus* hybrid.** *Cytogenet Genome Res* 2002, **98**:184-188.
 196. Sakai C, Konno F, Nakano O, Iwai T, Yokota T, Lee J, Nishida-Umehara C, Kuroiwa A, Matsuda Y, Yamashita M: **Chromosome elimination in the interspecific hybrid medaka between *Oryzias latipes* and *O. hubbsi*.** *Chromosome Res* 2007, **15**:697-709.
 197. Moorhead P, MMKaplan, PW Brown (eds.): *Mathematical Challenges to the Neo-Darwinian Interpretation of Evolution.* Philadelphia: Wistar Institute Press 1967.
 198. Shapiro JA: **Genome system architecture and natural genetic engineering in evolution.** *Ann N Y Acad Sci* 1999, **870**:23-35.
 199. Shapiro JA: **A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering.** *Gene* 2005, **345**:91-100.
 200. Jacob F: **Evolution and tinkering.** *Science* 1977, **196**:1161-1166.
 201. Jacob F: **Complexity and tinkering.** *Ann N Y Acad Sci* 2001, **929**:71-73.
 202. Comai L, Madlung A, Josefsson C, Tyagi A: **Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids?** *Philos Trans R Soc Lond B Biol Sci* 2003, **358**:1149-1155.

203. Gould SJ, Eldredge N: **Punctuated equilibrium comes of age.** *Nature* 1993, **366**:223-227.
204. Kirchner J, Connolly CM, Sandmeyer SB: **Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element.** *Science* 1995, **267**:1488-1491.
205. Xie W, Gai X, Zhu Y, Zappulla DC, Sternglanz R, Voytas DF: **Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p.** *Mol Cell Biol* 2001, **21**:6606-6614.
206. Kassis JA: **Pairing-sensitive silencing, polycomb group response elements, and transposon homing in Drosophila.** *Adv Genet* 2002, **46**:421-438.
207. Peters JE, Craig NL: **Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE.** *Genes Dev* 2001, **15**:737-747.
208. Kinoshita K, Honjo T: **Linking class-switch recombination with somatic hypermutation.** *Nat Rev Mol Cell Biol* 2001, **2**:493-503.
209. Lisch D: **Epigenetic regulation of transposable elements in plants.** *Annu Rev Plant Biol* 2009, **60**:43-66.

Tables

