



## Gene duplication and other evolutionary strategies: from the RNA world to the future

Jürgen Brosius

*Institute of Experimental Pathology, Center for Molecular Biology of Inflammation, University of Münster,  
Von-Esmarch-Str. 56, D-48149 Münster, Germany  
E-mail: RNA.world@uni-muenster.de; fax: +49 251 835 8512*

Received 16.05.2002; Accepted in final form 29.08.2002

*Key words:* barriers of genetic exchange, Darwinian principles, evolutionary transitions, exaptation, group selection, Lamarckism, memes, nuon duplication, nuons, primordial sex, retroposition, RNA world, RNP world

### Abstract

Beginning with a hypothetical RNA world, it is apparent that many evolutionary transitions led to the complexity of extant species. The duplication of genetic material is rooted in the RNA world. One of two major routes of gene amplification, retroposition, originated from mechanisms that facilitated the transition to DNA as hereditary material. Even in modern genomes the process of retroposition leads to genetic novelties including the duplication of protein and RNA coding genes, as well as regulatory elements and their juxtaposition. We examine whether and to what extent known evolutionary principles can be applied to an RNA-based world. We conclude that the major basic Neo-Darwinian principles that include amplification, variation and selection already governed evolution in the RNA and RNP worlds. In this hypothetical RNA world there were few restrictions on the exchange of genetic material and principles that acted as borders at later stages, such as Weismann's Barrier, the Central Dogma of Molecular Biology, or the Darwinian Threshold were absent or rudimentary. RNA was more than a gene: it had a dual role harboring, genotypic and phenotypic capabilities, often in the same molecule. Nuons, any discrete nucleic acid sequences, were selected on an individual basis as well as in groups. The performance and success of an individual nuon was markedly dependent on the type of other nuons in a given cell. In the RNA world the transition may already have begun towards the linkage of nuons to yield a composite linear RNA genome, an arrangement necessitating the origin of RNA processing. A concatenated genome may have curbed unlimited exchange of genetic material; concomitantly, selfish nuons were more difficult to purge. A linked genome may also have constituted the beginning of the phenotype/genotype separation. This division of tasks was expanded when templated protein biosynthesis led to the RNP world, and more so when DNA took over as genetic material. The aforementioned barriers and thresholds increased and the significance and extent of horizontal gene transfer fluctuated over major evolutionary transitions. At the dawn of the most recent transformation, a fast evolutionary transition that we will be witnessing in our life times, a form of Lamarckism is raising its head.

### Introduction

A very reasonable, albeit unproven scenario, maintains that early cellular life went through a stage in which RNA macromolecules constituted the major biopolymers (Woese, 1967; Crick, 1968; Orgel, 1968). Their functions in such primordial cells were not only the storage and replication of genetic infor-

mation, but they were also involved in structural and enzymatic tasks as well. Earlier forms of life based on inorganic materials or other polymers related or unrelated to extant polynucleic acids (Orgel, 1998) are likely to have preceded RNA-based forms (Cairns-Smith and Davis, 1977; Cairns-Smith, 1982; Wächtershauser, 1992). At this point, it is still difficult to explain the origin of nucleotides (Ferris, 1994;

Orgel, 1998) and there is a vast void concerning initial steps towards the polymerization of nucleotides into oligo- and polynucleotides (Kiedrowski, 1986; Li and Nicolaou, 1994; James and Ellington, 1995; Szathmary, 1997) and their self-organization into living entities (Jantsch, 1979; Kauffman, 1993). For more detailed discussions of these issues, the reader is referred to excellent reviews and books (multiple authors, 1987; Gesteland *et al.*, 1999; Orgel, 1998; Fry, 2000; Wills and Bada, 2000; Joyce, 2002).

Yet, in this seemingly despairing situation, one should remember that only two decades ago, an answer to the question whether protein biopolymers preceded nucleic acid biopolymers or *vice versa* seemed equally elusive. The discovery of catalytic RNA resolved this dilemma virtually overnight (Kruger *et al.*, 1982; Guerrier-Takada *et al.*, 1983; Gilbert, 1986; Westheimer, 1986; Gesteland *et al.*, 1999). Despite some lingering uncertainties we can consider the RNA world as a reasonable theoretical outpost from which we can make rather effortless associations to extant cells as they still contain RNAs that function in various compartments and diverse biochemical tasks. By hypothesizing the nature of this RNA world, we may also eventually be able to venture back to examining earlier, even simpler stages of life. Thus, we can begin to integrate a hypothetical RNA world into an attempt to understand biological and evolutionary principles spanning 3–4 billion years of macromolecular, cellular and organismal transitions. Importantly, the RNA world is not yet behind us; particularly in eukaryotic cells, RNA still plays a pervasive role in numerous biochemical and regulatory pathways (Brosius, 1999d; Herbert and Rich, 1999; Filipowicz, 2000; Eddy, 2001; Erdmann *et al.*, 2001a, 2001b; Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001; Lee and Ambros, 2001; Mattick, 2001; Huttenhofer and Brosius, 2002; Storz, 2002).

In this chapter, we will first discuss the impact of RNA in extant cells with an emphasis on its continuous conversion to DNA. This section will remain brief as much of it has been dealt with in earlier articles (Brosius, 1991, 1999a, 1999b, 1999c, 1999d; Brosius and Gould, 1992; Brosius and Tiedge, 1995b). The rest of the chapter will consider what effect envisioning an RNA world built on Neo-Darwinian principles has on accepted evolutionary views – always with the caveat of our limited understanding of such a distant era.

## Duplication of genes and regulatory elements by retroposition

One of the major transitions in life occurred when the RNA world gradually evolved into the RNP world (Maynard Smith and Szathmary, 1995; Szathmary and Smith, 1995). At that stage an extremely versatile biopolymer, protein, joined RNA in structural and catalytic tasks. Simple, untemplated peptides were probably already present in the RNA world. What set the RNP world apart was the advent of translation; the templated synthesis of polypeptides at proto-ribosomes, that became increasingly sophisticated over time (Woese, 1980, 2001; Maizels and Weiner, 1987; Szathmary, 1999b; Brosius, 2001). The advent of versatile protein biopolymers was an important prerequisite for the next major transition, when DNA gradually replaced RNA as genetic material (Brosius, 1999d). The RNA to DNA transition probably occurred via a process of retroposition with the aid of an early form of the enzyme reverse transcriptase and other enzymes presumably as well. Hence, retroposition led to the establishment and growth of a DNA genome at the onset of the DNA world. When such retroposition events occurred in the ‘germ line’ some became fixed in the genomes of populations. The process of retroposition apparently did not vanish when all genomic RNA was converted into DNA but continues in many lineages even to this day (Jurka, 1998; Brosius, 1999a, 1999d). Despite the fact that retroposition is ancient it is still frequent today. We observe that at least 42% of the human genome is still discernible as having been contributed by retrotransposons (Lander *et al.*, 2001; Venter *et al.*, 2001). Even this is likely an underestimate, as retropositions that occurred more than 150–200 million years ago, are no longer identifiable due to the continuous onslaught of mutations. In fact, one might assume that all founding DNA sequences were generated by retroposition. Once the DNA genome became the template of enzymatic replication, additional duplication events based mainly on DNA replication errors expanded the genome. Such mechanisms, also occurring in modern cells, included simple stuttering at expanding homopolymers as well as oligonucleotide repeats involving a few base-pairs at a time or segmental duplications involving megabases.

While only a little more than a decade ago it was not trivial to convey the message that retroposition has the potential to generate functional genes (Brosius, 1991) and not just junk in the form of ret-

ropseudogenes, it is now apparent that many intronless genes were duplicated by retroposition via processed mRNA intermediates. One of the difficulties was the fact that retroposition had an apparent disadvantage compared to segmental duplication or whole genome duplication; mechanisms that were once thought to be solely responsible for generating novel genes (Bridges, 1936b; Lewis, 1951; Stephens, 1951; Nei, 1969; Ohno, 1970; and other chapters in this volume). Usually the mRNA-derived gene copy transposes without its associated promoters and other regulatory elements and hence, is most likely to be inactive. On the other hand, a retrogene that chances to integrate near resident promoter elements might lead to a selective advantage by immediately exhibiting a different temporal and or spatial expression pattern in comparison to the source gene. In contrast, a segmentally duplicated gene is likely to be expressed in the same way as the source gene. There is now a steadily growing list of genes generated by retroposition that often exhibit differential expression patterns in comparison to their source genes and were recruited or exapted (Gould and Vrba, 1982) into variant or novel functions (see for example <http://exppc01.uni-muenster.de/expath/alltables.htm#table3>). In fact, the majority of genes that lack introns may have been generated by retroposition (Brosius, 1999a). According to J.J. Emerson and M. Long, ~15–25% of all genes in the human genome lack introns in their coding regions (personal communication).

Retroposition also generates novel genes encoding non-messenger RNAs. A decade ago, the idea that cells might rely on novel RNA molecules rather than proteins for potential functional innovations was peculiar to scientists. However, a novel, neuron-specific, small non-messenger RNA (BC1 RNA, age ~60–110 MY) has been observed in rodents and another analogous neuron-specific RNA (BC200 RNA, age ~35–55 MY) has been discovered in anthropoid primates (Sutcliffe *et al.*, 1982; DeChiara and Brosius, 1987; Watson and Sutcliffe, 1987; Martignetti and Brosius, 1993a, 1993b; Brosius and Tiedge, 1995a, 2001; Tiedge *et al.*, 1991, 1993). When we are finished analyzing eukaryotic genomes, we believe we will not only have discovered a plethora of ancient RNA genes but also a number of RNA genes that arose well past the stages of the RNA/RNP worlds (Hüttenhofer *et al.*, 2001; Hüttenhofer and Brosius, 2002); as recently as a few million years ago (Wang *et al.*, 2002).

The continued use of RNAs in diverse cellular tasks and the recruitment of novel RNAs into biological systems is further testimony to the fact that the RNP world— at least in Eukarya – is not yet history. Despite their antiquity, RNA molecules are still superior to proteins in certain aspects. For example, RNAs are able to interact with other nucleic acids at higher levels of specificity and RNA molecules are capable of bringing together protein molecules from different cellular compartments (e.g., via RNA binding sites) that otherwise would never interact with each other.

Demonstrably, retroposition has the potential to generate novel genes that may encounter different regulatory elements at the respective loci of integration. The process also facilitates the opposite scenario: retronuons can insert at more or less random positions in the genome and potentially modulate the expression of targeted genes. Retronuons of the SINE, LINE and LTR subfamilies frequently contribute promoter elements, enhancers, silencers, polyadenylation signals, and splice sites that act on the targeted genes. LINE retronuons have been implicated in the inactivation of X-chromosomes (Lyon, 1998, 2000; Bailey *et al.*, 2000). For further details the reader is referred to a number of reviews on the subject (Brosius and Gould, 1992; Shapiro, 1992; McDonald, 1993, 1995; Britten, 1996, 1997; Kidwell and Lisch, 1997; Brosius, 1999a, 1999b, 1999c, 2003b; Makalowski, 2000; Sverdlov, 2000; Mattick, 2001) or to our www page (<http://exppc01.uni-muenster.de/expath/alltables.htm#table1> and <http://exppc01.uni-muenster.de/expath/alltables.htm#table2>).

In summary, retropositional duplication of protein encoding genes can lead to their admixture with resident regulatory regions at the locus of retronuon integration and conversely, novel regulatory regions can combine with resident genes through retronuon dispersal. Hence, retroposition is a process that constantly and pervasively modulates genomes.

### **What can the RNA world reveal about the validity of evolutionary principles?**

Like a rock in the surf, the RNA world can be viewed as a vantage point in a sea of uncertainty. From there, we can almost reach on dry feet (and better understand) the shores of modern cells. We may also have a chance to venture out into the stormy and uncharted waters of the pre-RNA world life.

In addition to imagining molecular events at the beginning of cellular life, the RNA world may, in the not so distant future, be reconstructed in an experimental setting offering the chance to physically reverse engineer primitive cells based on RNA as enzymes, genetic material and structural components. There are efforts under way to generate selected functional RNA nuons in the test tube via selection/amplification schemes (Famulok and Jenne, 1999; Wilson and Szostak, 1999) that may eventually contribute to man-made life forms capable of reenacting life in a primitive cell reminiscent of the RNA world (Bartel and Unrau, 1999; Szostak *et al.*, 2001). It is obvious that such a reconstruction would not accurately retrace the path of evolution, nevertheless, the resulting RNA-based cells could serve as models for life in the RNA world. One can expect this feat to have been achieved within a few decades.

By extrapolating from such hypothetical but reasonable scenarios, aspects of which may be testable in the future, we may be able to address important questions in evolutionary biology at their roots. Also it should be remembered that modern cells have still not completely left the RNA world. It is apparent that extant cells harbor many more RNA species than anticipated (Hüttenhofer *et al.*, 2001). These RNAs still function in a wide spectrum of cellular regulation (Eddy, 1999, 2001; Hüttenhofer and Brosius, 2002; Storz, 2002). Even RNAs that arose quite recently (i.e., after the mammalian radiation) have been exapted into novel functions (DeChiara and Brosius, 1987; Brosius and Gould, 1992; Martignetti and Brosius, 1993a, 1993b; Brosius, 1999c; Cavaillé *et al.*, 2000). The vigorous and diverse 'RNA' life in contemporary cells is an important link to primordial RNA/RNP worlds (Jeffares *et al.*, 1998; Poole *et al.*, 1998; Brosius, 1999d; Herbert and Rich, 1999; Mat-tick, 2001).

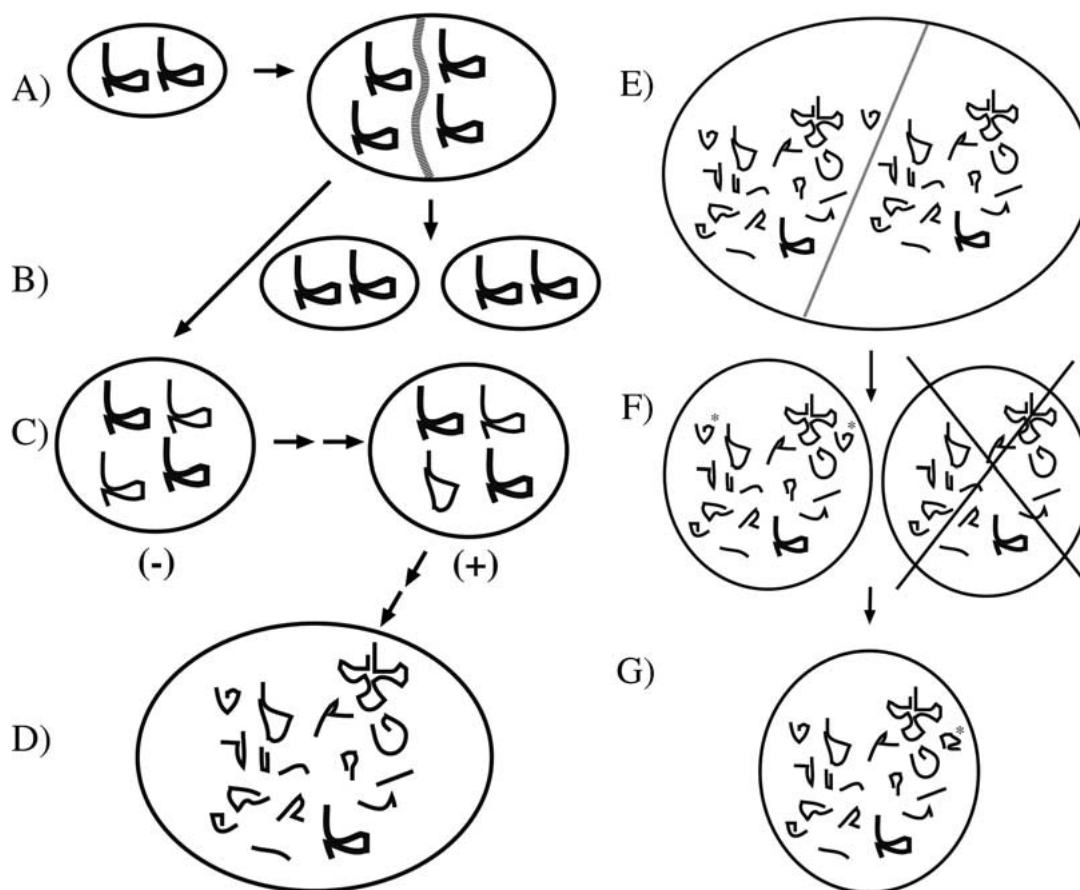
If evolutionary theories are correct, they should apply to all stages of evolution, including primal ones. As we will see, the core predications including the evolutionary importance of unit duplications (genes, nuons) and the Darwinian principles can remain valid in the RNA or RNP worlds, while the roles that other processes, fundamental to evolution, might have fluctuated quantitatively or require modification at different stages of cellular or organismal organization and complexity.

## Scenes from an imaginary RNA world

In the following we will demonstrate that, in the RNA world gene (nuon) duplication must have played a decisive role in evolution, selection acted upon individual RNA nuons as well as on groups of nuons, and primordial sex (exchange of RNA nuons) was promiscuous. Due to the latter the evolution of distinct species in the RNA world was virtually impossible. As RNA nuons determined genotype *and* phenotype of a cell, we can conceive of a world with Neo-Lamarckian elements defined as the inheritance of acquired characteristics barring directed modification of genetic material. Importantly, the significance of amplification of genetic material, its random variation (including recombination) and selection, remains a pillar of evolutionary theory even in the RNA/RNP worlds. Several other principles did not apply. For example, Weismann's Barrier and the Central Dogma of Molecular Biology were virtually absent or rudimentary.

### *Modest beginnings of an RNA-based world*

As stated above, it is not yet clear, how ribonucleotide phosphates, the building blocks of RNA, were generated from inorganic matter and what drove their polymerization. In the beginning, a self-replicating biopolymer could have arisen by chance. However, if more than one replicator was necessary their containment was vital (Szostak *et al.*, 2001). Thus, unless these replicators were present at high concentrations in a given environment, co-optation of a semi-permeable (lipid) membrane should have occurred at an early stage (Maynard Smith and Szathmáry, 1995; Szostak *et al.*, 2001). Cycles of replication and cellular growth (Figure 1A) with spontaneous 'cell' division could have constituted an early form of RNA-based life (Figure 1B). If a larger undivided 'cell' (Figure 1A) had a selective advantage over a 'cell' with less nuons, we can perhaps equate such a scenario with the first 'genome' ('nuonome') duplication in the RNA world. Some of the replication-incompetent nuon variants (mutants, shown by thinner lines in Figure 1C -) were possibly the first 'genomic parasites'. Through gradual alteration of sequence/structure and subsequent exaptation of the variant biopolymers into novel functions, some of these may have conveyed a selective advantage to the cell (Figure 1C +) by facilitating recruitment of or by enzymatically aiding the synthesis of precursor mole-



*Figure 1.* An RNA world can be imagined that already encompassed ‘gene’ duplication, generating new nuons by amplification/mutation. Two RNA biopolymers capable of replication enclosed by a primordial membrane multiplied and the primordial cell grew in size (A). The normal route was eventual division by physical forces resulting in two daughter cells containing similar sets of RNA nuons (B). Alternatively, some of the RNA nuons lost their ability to replicate (thin lines) but were propagated by the replicators (bold lines). Carrying a functionless nuon as ballast must have translated into a high cost and thus, cells containing ‘parasitic’ nuons were negatively selected (C, -). Like all nuons however, the superfluous nuon may have mutated and on rare occasions, altered forms may have accidentally added a function for the benefit of the cell. Such a cell was positively selected (C, +). Continued recruitment of duplicated and altered nuons established more complex cells with different functional nuons and a replicator (D). Duplication (E) and unequal distribution of nuons to daughter cells (F) resulted in one that possessed an extra copy of a given nuon type (asterisk) and another (on the right) with none. This led to a reduced or loss of viability of the latter (crossed out). The daughter cell with duplicate nuons might have had a further advantage over daughter cells resulting from an equal distribution of nuons if the duplicated and subsequently altered (asterisk) nuon was exapted into a variant or novel function (G).

cules. If the depicted scenario or a similar one unfolded, early symbiosis and group selection emerged. In the same vein, an ‘altruistic’ replicator had a selective advantage over a ‘selfish’ replicator as it facilitated provision of ‘raw material’ for its own success: the altered nuon, itself replication incompetent, might have provided a different auxiliary metabolic function and thus would have been among the first exaptations (Gould and Vrba, 1982; Brosius and Gould, 1992).

Various RNA nuons could have been generated in

a similar fashion (Figure 1D) leading to more complicated cells with enhanced metabolic and reproductive capabilities. Thus, the principle of nuon (gene) duplication (Bridges 1936a; Lewis 1951; Muller *et al.* 1936; Nei 1969; Ohno 1970; Stephens 1951; Sturtevant 1925), as well as other principles of Neo-Darwinian evolution, namely random variation and selection, would have played major roles in an evolving RNA world.

### Evolution by nuon duplication and exchange

As described above on a minimal scale, replication of RNA molecules would have led to their amplification and, after growth of cell content and incorporation of external or internally synthesized lipids into the membrane, there could have been spontaneous or triggered division (Figure 1E,F). For simplification only a duplication of each nuon type ( $n=2$ ) rather than multiplication ( $n>2$ ) is shown; of course, multiple copies of each nuon ( $n>2$ ) were most likely generated. In the absence of sophisticated mechanisms that would have enabled equal distribution of each nuon type to the daughter cells, multiple copies would be vital. According to the formula  $P = (1 - 1/2^n)^N$  where  $P$  is the probability,  $N$ , the number of different nuons per cell and  $n$ , the number of copies of each component, a genome of 10, 100, or 1000 different nuons would require a copy number of 10, 14, or 17, respectively, to have a 99% chance that each nuon type be found in both daughter cells. It is not clear how a daughter cell with a low copy number of a given nuon could later re-adjust its copy number. Therefore, above a certain genome size, linkage of different types of RNA nuons into heteromeric concatemers ('chromosomes') might have been advantageous.

The unequal distribution of individual nuons would have resulted in an early form of nuon (gene) duplication. While the absence or underrepresentation of one nuon type conceivably could have resulted in a selective disadvantage or non-viability of the affected daughter cell (Figure 1F, right), the presence of an extra nuon copy or overrepresentation of one nuon type (Figure 1F, left), especially at a high error rate of RNA replication, might have altered, modulated and/or changed the function of such nuons. Such a variant nuon may have provided a selective advantage (Figure 1G).

The exchange of genetic material (primordial sex) was probably rampant among RNA-based cells. Just as cells could have spontaneously divided to yield two daughter cells, two cells with different sets of RNA nuons (genetic material) may have fused. As a result, nuons would have been shuffled and re-distributed to daughter cells yielding sets of nuons ('nuonomes') that were different from those of the initial cells. Depending on the distribution, this may have been advantageous (e.g., in a different environment) or disadvantageous (if important nuons were

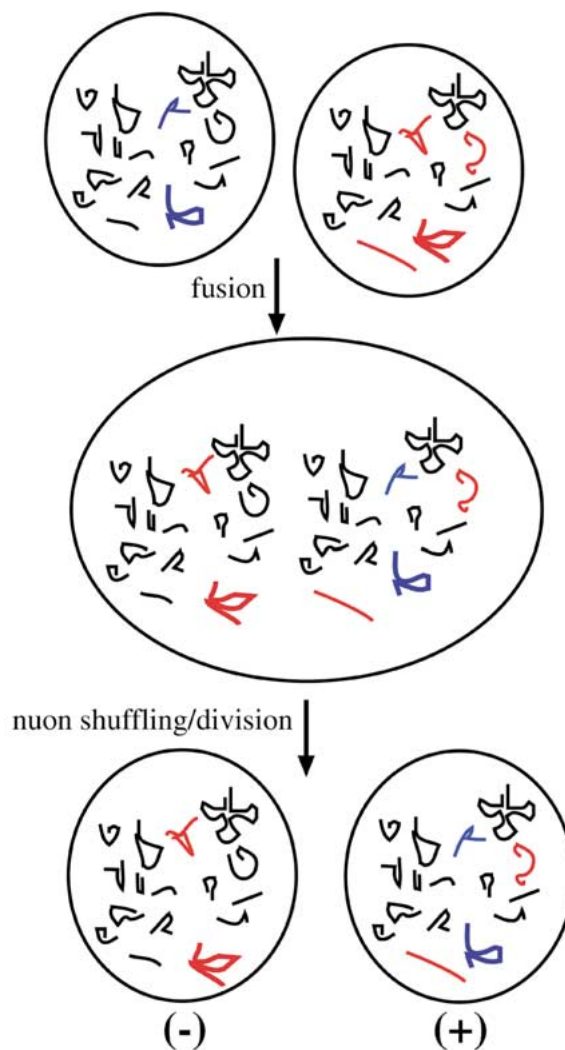


Figure 2. Sex in an imaginary RNA world. Two cells with both different (blue and red) and identical (black) sets of nuon types (top) fused and after diffusion and amplification (not shown) of some of the nuons (center), divided again (bottom). Some nuon combinations in a cell had a selective advantage (+) over others (-). Replicators are in bold. The bottom left cell may have possessed a better replicator but the replicator was better only in association with different accessory nuons ('genetic' background).

underperforming in the different 'nuonomic' environment) to one or both cells (Figure 2).

### Phenotype was genotype

The following example is included to demonstrate that in the RNA world phenotype and genotype were inseparable in a given nuon. The scenario is highly

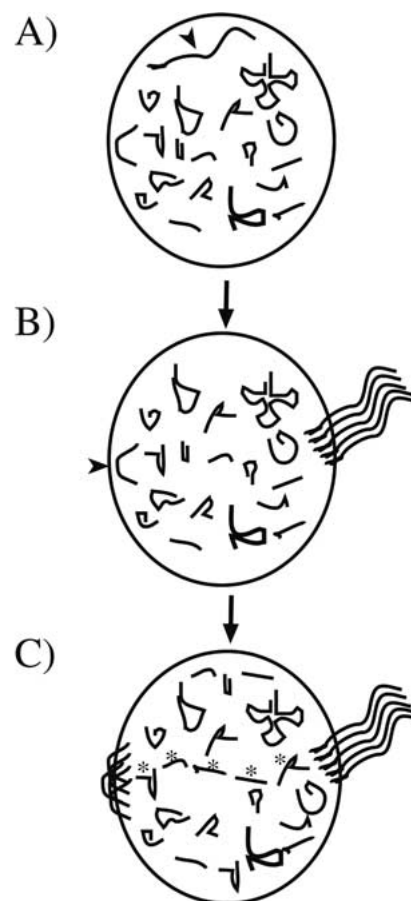
speculative as it implies that homo- or hetero-multimeric RNA biopolymers associated to form functional units for such things as motility (flagelloids), for (chemo)sensation or signal transduction. Nevertheless, it illustrates a point that is also valid for less complicated and not so speculative functions of RNA biopolymers.

Even though there is no evidence whether motility evolved in RNA-based cells, movements surely would have been advantageous to a part of a population that could then move away from nutrient depleted areas or otherwise unfavorable conditions. The evolution of sensory systems and signaling pathways involving RNA biopolymers and other small molecules would have provided a significant selective advantage in the search for favorable and avoidance of unfavorable conditions (Figure 3). Despite its speculative nature, the example illustrates how in the RNA world genotype and phenotype were a single entity. A fortuitous mutation in one of the RNA components of the sensory cluster may have improved its performance. From there it is not hard to imagine that this RNA nuon spread in cells and populations (through primordial sex) over time. Such a mechanism is also conceivable for less hypothetical RNA nuons if a newly arisen mutation was under positive selection.

#### *Chaining nuons – an early evolutionary compromise*

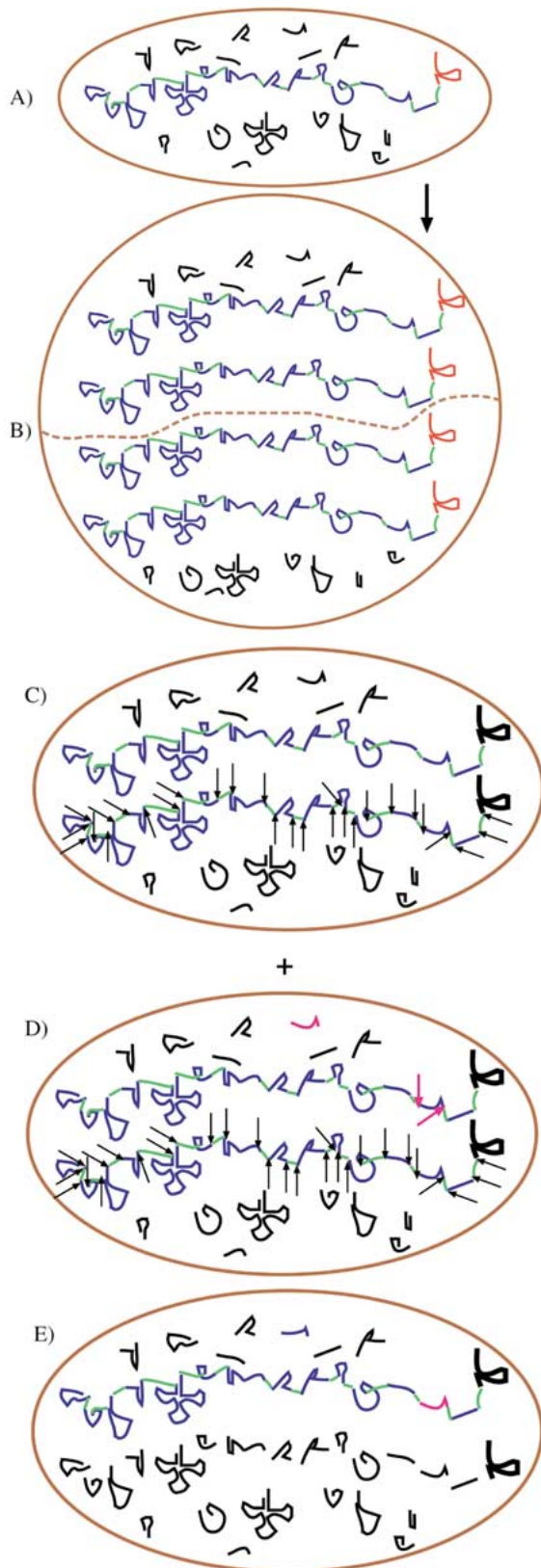
For reasons discussed above, linkage of several nuons into an RNA genome might have been advantageous. In addition to the joined genomic form (the ‘germ line’) there would still have been free RNA nuons (somatic nuons) in the ‘cells’. Occasional replication of these individual nuons could still have taken place (Figure 4A,B). After replication, chiefly of the joined genetic form, the cell divided (Figure 4B,C) and a portion of the joined forms was processed into individual functional nuons (Figure 4C). A major problem of this scenario is posed by the necessity of maintaining some of the genetic chains intact while producing enough of the functional free nuons. However, RNA processing is still an essential mechanism during RNA biogenesis in all extant cells (Herbert and Rich, 1999). Despite the absence of any direct proof, it is attractive to contemplate that many forms of RNA processing had their roots in the RNA world (Brosius, 1999d).

An interesting aspect is that the appearance of chiefly genetic (versus somatogenetic) RNA already



*Figure 3.* In an imaginary RNA world phenotype equaled genotype. In a highly speculative scenario, an RNA nuon (arrowhead) generated by duplication and modification (perhaps in association with duplicates as a homopolymer) acquired a function in cell motility (B). This was advantageous. Likewise, other nuons (arrowhead, B) arose by amplification/modification and generated a complex of molecules that was able to sense cues from the environment (photons, pH, temperature, nutrients) (C). A hypothetical chain of signal transducing molecules (asterisks) may have linked the two systems (C).

began to establish a barrier. It was not impenetrable, as a randomly altered ‘somatic’ nuon could occasionally be added to the linked ‘germ line’ RNA chain, or an altered nuon could replace the corresponding nuon in the ‘genome’ by some sort of recombination (Figure 4D,E). While establishing a barrier in one sense, recombination of nuons with the genome generated a ‘safer’ environment for the survival of selfish nuons: interconnected with proto-chromosomes they would have been more refractory to purging. This situation is reminiscent of modern genomes, in which disadvantageous nuons can hitch a ride by link-



age to positive traits (Brosius and Tiedge, 1995b; Brosius, 1999a).

## Discussion and Conclusions

Despite, or perhaps because of, our gaps in knowledge about early life forms, it was not difficult to construct a hypothetical RNA world that is solidly anchored within the framework of Neo-Darwinian principles. Akin to established *in vitro* amplification/selection experiments (Szathmary, 1990; Ellington and Szostak, 1990; Robertson and Joyce, 1990; Tuerk and Gold, 1990; Wilson and Szostak, 1999; Szostak *et al.*, 2001) evolutionary change is always mediated by amplification of nuons and the selection of random alterations. A minor difference between this type of *in vitro* evolution and natural selection is that, *in vitro*, usually only the best performers are selected for subsequent rounds of amplification while the latter usually removes only individuals of lower fitness from the population. This ensures maintenance of a large genetic diversity in populations (Mayr, 2001). In natural settings, there are additional principles at work such as the conflict between integration of and insulation from foreign genetic material (see below).

### Group selection, group sex and chains of nuons

As outlined above and previously noted by Szathmary (1999a), in the RNA/RNP worlds the selection of RNA nuons enclosed by a membrane can only be interpreted as group selection. In an RNA world with limited numbers of nuons, a selfish parasitic RNA nuon could pose a serious disadvantage to the cell. This situation could only have a positive outcome if the cell lost the selfish nuon or by fortuitous alteration

*Figure 4.* Linkage of nuons and the origin of RNA processing in an imaginary RNA world. In addition to different individual nuons (black) a more advanced RNA-based cell could also have featured the same nuons in a linked form (blue) giving rise to one or several proto-chromosomes that might have included a terminal replicator as a tag shown in red (Maizels and Weiner, 1987). The joined nuons (A) were interspersed by short linkers (green). RNA chromosomes may have facilitated replication (B) and equal nuon distribution to daughter cells (C, D). The ‘intervening’ linker sequences may have been necessary for efficient processing (arrows) of individual nuons from the joined proto-chromosome (C, D, E). A free variant nuon (magenta) may have recombined with the linked homologue by the appropriate cuts (magenta arrows) in the proto-chromosome (D) followed by replacement and re-ligation (E).

and adaptation/exaptation of the formerly selfish nuon into a functional role. Already at the level of the RNA world, the types and variants of other nuons in the cell, its 'genetic background', could have played a decisive role in selecting for or against a given nuon. Clearly, a nuon's performance greatly depended on its 'co-players' and a given environment.

The reshuffling of nuons creating ever – changing genetic backgrounds was favored by the unlimited exchange of genetic material. Free sex was gradually curbed by linkage of different nuons. This would have facilitated equal distribution of genetic material during cell division and begun to establish more specialized genomes; gradually separating genotype and phenotype. This eventually led to the complex DNA-based genomes found in extant organisms. Hence, evolutionary novelties would have been selected more directly in the RNA world and more indirectly in increasingly complex systems at the molecular (chromosomes, DNA), subcellular (nuclear compartmentalization) and organismal (compartmentalization into the germ line versus somatic cells) levels.

### Evolutionary transitions

It is dangerous to categorize life into lower or higher, primitive or advanced forms (Dawkins, 1992; Gould, 1996). All too often, progress is viewed from an

anthropocentric pedestal. One forgets that even reduction in complexity, such as reduction in genome size, or reduction of gene numbers can be evolutionary progress, depending on the 'demands' of changing environments. Nevertheless, when considering extinct and extant forms of life, common sense dictates that there must have been an increase of complexity over time. Such an increase is apparent when comparing the limited number of different nuons in a fledging RNA world to the thousands or tens of thousands of genes in extant organisms (see also Adami *et al.*, 2000). Over time, one observes the evolution and increasing use of novel biopolymers. Furthermore, mergers of cells occurred, in which one partner became an organelle (Margulis, 1970) or, on more subtle terms, in which individual cells associated to form multicellular and differentiated organisms. Transitions to levels of increasing complexity have been characterized by, to name a few, the origin of translation, chromosomes, eukaryotes, sex, multicellular organisms and social groups (Maynard Smith and Szathmary, 1995). Numerous other transitions such as the rise of epigenetic phenomena, (e.g., imprinting of

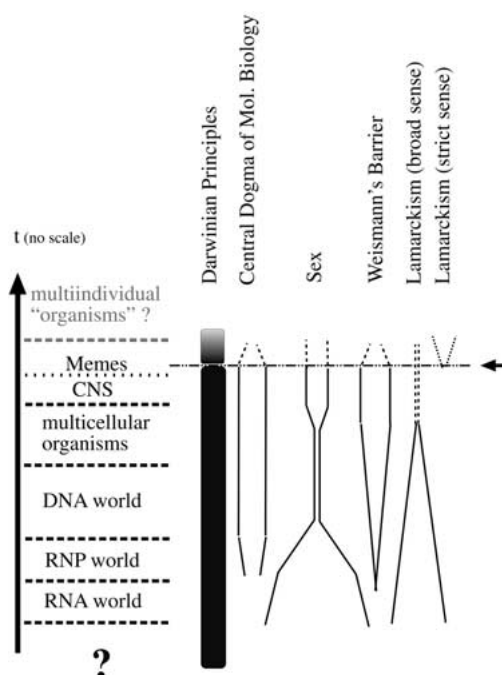


Figure 5. Stages of evolutionary transitions versus variance of evolutionary dogmas. The left panel shows selected major transitions in evolution resulting in different stages over time (a time scale was not attempted). The nature of such stages and transitions prior to the RNA world are not yet perceived (question mark). Major principles (amplification, variation, and selection of units) remained constant over the various transitions, others shifted in presence or scope. Broken lines indicate enhanced uncertainty. The stippled horizontal line indicated by the arrow on the right denotes present time, beyond which we can only speculate.

genes that arose in placental mammals) are not considered here.

In Figure 5, Maynard Smith's and Szathmary's concept of major transitions has been modified slightly (Maynard Smith and Szathmary, 1995). Only a selection from numerous transitional stages is included. Starting from a likely cellular progenitor in the RNA world (see above), the first major transition considered was the advent of templated protein biosynthesis, transferring the catalytic and structural tasks of RNA macromolecules, to a totally different class of macromolecules. The next major transition considered was the transfer of genetic information from RNA to DNA. A prerequisite of which was a major transition itself, namely the concatenation of RNA nuons into 'RNA chromosomes', perhaps allowing improved replication and/or more equal distribution of RNA nuons to daughter cells. The transition from RNA to DNA as genetic material was a gradual

process that obviously continues in many modern genomes in the form of retroposition (Brosius, 1999a). The next evolutionary stage was reached when cells associated and gradually formed multicellular and eventually increasingly differentiated organisms some of which subsequently sequestered germ line cells from somatic cells.

The evolution of a nervous system heralded the beginning of another major transition. It provided a notable selective advantage as information about the environment could be processed rapidly and transformed the organism into a better 'survival machine' for its genes (Dawkins, 1976). There are precedents of non-genetic transfer of information in social insects, such as the locations of food sources in bees, or transmission of 'culture' in mammals, as exemplified by songs of whales or tool utilisation of chimpanzees (Bonner, 1980). In one lineage, however, a highly developed central nervous system (CNS) was exapted into an instrument that permitted the direct, non-genetic transfer of information at an unprecedented magnitude (Gould and Vrba, 1982). This transfer could not only flow horizontally, but also from one generation to the next with equal efficiency. The vehicle of transfer is spoken, written or digitized information, memes in a broad sense (Dawkins, 1976; Blackmore, 1999).

At this point, it should be stressed that at least the upper levels of biological complexity depicted in Figure 5 are biased, to all intents and purposes, towards the evolutionary fate of our own species. In a correct and all-encompassing representation this scheme would include all lineages of extant and extinct life (Gould, 1996, 2002). A large proportion of species (presumably all Prokaryotes) have apparently not made major transitions beyond the one leading to the use of DNA as genetic material. The scheme would even encompass organisms, such as viruses, endosymbionts, or parasites, which made the transition to a new 'stage' via simplification of complexity. It is noteworthy that the evolutionary stages depicted in Figure 5 seem to imply, although there was no attempt to introduce a time scale, that the more recent transitions occurred in shorter successions, as if evolutionary complexity accelerated when higher levels of complexity were reached.

We are possibly at the threshold of a novel stage representing analogues of multicellular organisms, namely super-organisms whose units are not cells but individual organisms, even social entities. Again, this is not the first time in evolution that such a transition

occurred, there are already notable examples in insects (Giraud *et al.*, 2002).

*The 'softer' evolutionary principles are neither present nor constant over the entire transitional spectrum*

The aforementioned evolution of a linear 'genome' also illustrates how, unlike the constant Neo-Darwinian tenets, other evolutionary principles were subject to modification during major evolutionary transitions. Weismann's Barrier could be considered a continuation of the demarcation of the initial 'germ line'. While, in a strict sense, Weismann's barrier applies only to organisms with a germ line that is separate from somatic cells (Pollard, 1984), one may postulate that barriers existed to varying degrees at several evolutionary stages. Initially, in the RNA world, there was no barrier whatsoever, apart from a presumed membrane. Primordial sex which experimented with new combinations of different nuons, was virtually unlimited (Figure 3). When nuons merged into protochromosomes (Figure 4), a distinct novel cellular RNA nuon, generated either by amplification and subsequent mutation of a free endogenous nuon or by exchange with another cell harbouring different nuons, would encounter a barrier. Such a barrier was evident simply by the reduced ease at which the novel nuon could be integrated into the 'germ line'. This contrasts with the simple uptake in an unlinked genome of separate nuons, a frequent event that would have been sufficient to ensure propagation into daughter cells. Conceivably, integration into the proto-chromosome was a means to overcome the barrier (as depicted in Figure 4D,E). As different levels of complexity were reached, the barrier became temporarily less penetrable when, for example, DNA became the genetic material. Nevertheless, penetration of the barrier may have varied over evolutionary transitions, for example, by cells evolving various recombinatorial mechanisms. With the establishment of additional barriers against the indiscriminate exchange of genetic material, but not necessarily concomitant with these events as additional parameters might also have varied, unlimited sex (horizontal gene transfer) was curbed.

One of the consequences of this primordial puritanism was the gradual separation of phenotype and genotype. It might have constituted one of the first evolutionary barriers. The Darwinian Threshold (Woese, 2002) triggered by templated translation

leading to more complex and interconnected cell designs is yet another important barrier; according to Woese (2002) the one that ‘truly represents the origin of species’. I myself tend to consider the origin of species as a more prolonged process that might have included several critical transition points requiring several barriers, such as (i) the aforementioned linkage of individually replicated RNA nuons into a concatenated RNA genome, (ii) the Darwinian Threshold (Woese, 2002) and, in particular, (iii) the action of reverse transcriptase, the enzyme responsible for converting RNA into DNA genomes (Brosius, 2003a).

#### *Evolution is conflict – conflict is evolution*

Perhaps a further evolutionary force, in addition to amplification, variation, and selection, is conflict – an erratic, yet everpresent dynamism. Over evolutionary time, there must have been a continuous oscillation between the optimal value of accuracy in replication of genetic information and the availability of sufficient variation for Darwinian natural selection. Similarly, we observe the expansion and contraction of ‘superfluous’ genetic information in intergenic sequences and introns as a constant fluctuation. To a significant extent, retroposition, itself a product of an ancient but still ongoing conflict of reverse transcriptase encoding retrons with their host genomes, is responsible for major genome expansions. Sex poses yet another riddle: An organism devotes a large proportion of its energy to protecting and promoting every single nucleotide in its genome just to squander 50% of them in a single sweep in the act of propagation. In this context, Ghiselin’s suggestion that a species should be considered an individual deserves additional scrutiny (Ghiselin 1974; Ghiselin 1989; Ghiselin 1997). The need to co-opt genetic material from other cells, organisms and even species is in apparent conflict with the simultaneous establishment of barriers against invasion by foreign nucleic acids. Similar conflicts involving insulation against and integration of evolutionary units not only apply to nuons, genes and genomes but also extend to organellar symbionts, cells, individuals and societies. Often, one gets the impression of a delicate tug of war on a ridge – if one team is too strong and wins, it pulls both into the abyss, what we would call extinction or an evolutionary dead-end street. Perhaps, co-operational strategies evolved, in part, to alleviate evolutionary conflicts.

#### *Lamarckism raises its head*

At the dawn of a major transition in our own lineage we witness major shifts in the evolutionary impact of Weismann’s barrier, sex, and the Darwinian Threshold (Woese, 2002) without abandoning the basic Neo-Darwinian principles. In the RNA/RNP worlds (Figure 5) there was little separation between molecules acting as functional units and those acting as genetic material. Nevertheless, variations in RNA nuons and their immediate selection cannot be considered Lamarckism in the strict sense as the mutations were random and not geared towards a specific nuon in ‘need’ of modification. Only at the stage of memes are we able to observe Neo-Lamarckism: a learned behavior can be transmitted from individual to individual not only in a vertical but also in a horizontal transfer (Jablonka and Lamb, 1995; Jablonka *et al.*, 1998). Of course, this knowledge or meme could not direct the alteration of genetically based behavior – up to now.

There is no doubt that somatically rearranged genes could make their way to the germ line by viral transfer (Steele *et al.*, 1998). Such events, however, are few and far between and they are certainly, like retroposition, random events with no directionality. Almost timidly however, the same authors allude to future possibilities: ‘we may very well discover and harness and oversee our genetic destiny’ (Steele *et al.*, 1998). It cannot be denied that a novel form of Lamarckian evolution is on the rise and it might remain in effect for some time. The same species that obviously uses memes in a most efficient manner is not only able to conceptualize the genetic mechanisms of evolution, but has acquired the possibility to free itself from the dictatorship of the genes by engineering the germ line of virtually all living species including its own. It can, for example, strive to correct genetic disease, introduce desired traits, design genes from scratch, and introduce additional chromosomes (Silver, 1997). Weismann’s Barrier and the Darwinian Threshold will recede, sex may disappear or experience transformation akin to the one from primordial sex to meiotic sex and Lamarckian mechanisms will persist and expand. Traits that have been acquired and back translated into the genetic code during a lifetime will be passed genetically to the next generation. The Central Dogma of Molecular Biology that sought to extend Weismann’s Barrier to the molecular level will receive yet another, perhaps its most serious, blow. Is this the first time in 3.5 billion years that something

akin to Lamarckian evolution arose or has it been tried before and proven to be inferior (e.g., too costly) compared to Darwinian selection (Hayes, 1999)?

### *Peaceful coexistence of Darwinism and Lamarckism?*

The Darwinian principles of amplification, variation and selection appear to have permeated all major transitions in all forms of life like a monolith (Figure 5). Despite the rise of Lamarckism they will remain the cornerstones of evolution in the distant future, if our particular lineage has one.

### Acknowledgements

Thanks to Michael Famulok, Michael Ghiselin, Manyuan Long, Ernst Mayr, Günter Theißen, Leigh Van Valen and Carl Woese for references or discussions. I would like to dedicate this paper to Wally Gilbert on occasion of his seventieth birthday. The author is supported by the German Human Genome Project through the BMBF (#01KW9966).

### References

- Adami, C., Ofria, C. and Collier, T.C. (2000) Evolution of biological complexity. *Proc. Natl. Acad. Sci. USA*, **97**, 4463–4468.
- Bailey, J.A., Carrel, L., Chakravarti, A. and Eichler, E.E. (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc. Natl. Acad. Sci. USA*, **97**, 6634–6639.
- Baltimore, D. (1970) Viral RNA-dependent DNA polymerase. *Nature*, **226**, 1209–1211.
- Bartel, D.P. and Unrau, P.J. (1999) Constructing an RNA world. *Trends Cell Biol.*, **9**, M9–M13.
- Blackmore, S. (1999) *The Meme Machine*. Oxford University Press, Oxford, UK.
- Bonner, J.T. (1980) *The Evolution of Culture in Animals*. Princeton University Press, Princeton, NJ.
- Bowler, P.J. (1992) Lamarckism. In *Keywords in Evolutionary Biology* (Keller, E.F. and Lloyd, E.A.), Harvard University Press, Cambridge, MA, pp. 188–193.
- Bridges, C. (1936a) The Bar 'gene': a duplication. *Science*, **83**, 210–211.
- Bridges, C.B. (1936b) Genes and chromosomes. *Teaching Biol.*, 17–23.
- Britten, R.J. (1996) DNA sequence insertion and evolutionary variation in gene regulation. *Proc. Natl. Acad. Sci. USA*, **93**, 9374–9377.
- Britten, R.J. (1997) Mobile elements inserted in the distant past have taken on important functions. *Gene*, **205**, 177–182.
- Brosius, J. (1991) Retroposons--seeds of evolution. *Science*, **251**, 753.
- Brosius, J. (1999a) Genomes were forged by massive bombardments with retroelements and retrosequences. *Genetica*, **107**, 209–238.
- Brosius, J. (1999b) Many G-protein-coupled receptors are encoded by retrogenes. *Trends Genet.*, **15**, 304–305.
- Brosius, J. (1999c) RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene*, **238**, 115–134.
- Brosius, J. (1999d) Transmutation of tRNA over time. *Nat. Genet.*, **22**, 8–9.
- Brosius, J. (2001) tRNAs in the spotlight during protein biosynthesis. *Trends Biochem. Sci.*, **26**, 653–656.
- Brosius, J. (2003a) The contribution of RNAs and retroposition to evolutionary novelties. *Genetica*, in press.
- Brosius, J. (2003b) Echoes from the past – are we still in an RNP world? In preparation.
- Brosius, J. and Gould, S.J. (1992) On 'nomenclature': a comprehensive (and respectful) taxonomy for pseudogenes and other 'junk DNA'. *Proc. Natl. Acad. Sci. USA*, **89**, 10706–10710.
- Brosius, J. and Gould, S.J. (1993) Molecular constructivity. *Nature*, **365**, 102.
- Brosius, J. and Tiedge, H. (1995a) Neural BC1 RNA: Dendritic localization and transport. In Lipshitz, H.D. (ed.) *Localized RNAs*. R.G. Landes, Austin, TX, pp. 289–330.
- Brosius, J. and Tiedge, H. (1995b) Reverse transcriptase: mediator of genomic plasticity. *Virus Genes*, **11**, 163–179.
- Brosius, J. and Tiedge, H. (2001) Dendritic BC1 RNA: Intracellular transport and activity-dependent modulation. In Richter, D. (ed.) *Cell polarity and subcellular RNA localization*. Springer Verlag, Berlin, pp. 129–138.
- Cairns-Smith, A.G. (1982) *Genetic Takeover and the Mineral Origins of Life*. Cambridge University Press, Cambridge, UK.
- Cairns-Smith, A.G. and Davis, C.J. (1977) In Duncan, R. and Weston-Smith, M. (eds.), *Encyclopaedia of Ignorance*. Pergamon Press, p. 397–403.
- Cavaill e, J., Buiting, K., Kiefmann, M., Lalonde, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.P., Brosius, J. and Huttenhofer, A. (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl. Acad. Sci. USA*, **97**, 14311–14316.
- Crick, F.H.C. (1958) On protein synthesis. *Symp. Soc. Exp. Biol.*, **12**, 138–183.
- Crick, F.H.C. (1968) The origin of the genetic code. *J Mol Biol*, **38**, 367–379.
- Darwin, C. (1872) *The origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. John Murray, London.
- Dawkins, R. (1976) *The selfish gene*. Oxford University Press, Oxford, UK.
- Dawkins, R. (1982) *The extended phenotype*. Freeman, San Francisco.
- Dawkins, R. (1992) Progress. In Keller, E.F. and Lloyd, E.A. (eds.), *Keywords in evolutionary biology*. Harvard University Press, Cambridge, MA, pp. 263–272.
- DeChiara, T.M. and Brosius, J. (1987) Neural BC1 RNA: cDNA clones reveal nonrepetitive sequence content. *Proc. Natl. Acad. Sci. USA*, **84**, 2624–2628.
- Eddy, S.R. (1999) Noncoding RNA genes. *Curr. Opin. Genet. Dev.*, **9**, 695–699.

- Eddy, S.R. (2001) Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.*, **2**, 919–929.
- Ellington, A.D. and Szostak, J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818–822.
- Erdmann, V.A., Barciszewska, M.Z., Hochberg, A., de Groot, N. and Barciszewski, J. (2001a) Regulatory RNAs. *Cell. Mol. Life Sci.*, **58**, 960–977.
- Erdmann, V.A., Barciszewska, M.Z., Szymanski, M., Hochberg, A., de Groot, N. and Barciszewski, J. (2001b) The non-coding RNAs as riboregulators. *Nucleic Acids Res.*, **29**, 189–193.
- Famulok, M. and Jenne, A. (1999) Catalysis based on nucleic acid structures. *Top. Curr. Chem.*, **202**, 101–131.
- Ferris, J. (1994) Origins of life - chemical replication. *Nature*, **369**, 184–185.
- Filipowicz, W. (2000) Imprinted expression of small nucleolar RNAs in brain: time for RNomics. *Proc. Natl. Acad. Sci. USA*, **97**, 14035–14037.
- Fry, I. (2000) *The Emergence of Life on Earth: A Historical Overview*. Rutgers University Press, New Brunswick, NJ.
- Gesteland, R.F., Cech, T.R. and Atkins, J.F. (1999) *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Ghiselin, M.T. (1974) A radical solution to the species problem. *Systematic Zoology*, **23**, 536–544.
- Ghiselin, M.T. (1989) *Intellectual Compromise. The Bottom Line*. Paragon House, New York.
- Ghiselin, M.T. (1994) The imaginary Lamarck: a look at bogus 'history' in schoolbooks. *The Textbook Letter*. Sept-Oct.
- Ghiselin, M.T. (1997) *Metaphysics and the Origin of Species*. State University of New York Press, Albany, NY. [ISBN 0-7914-3467-2].
- Gilbert, W. (1986) The RNA world. *Nature*, **319**, 618.
- Giraud, T., Pedersen, J.S. and Keller, L. (2002) Evolution of super-colonies: The Argentine ants of southern Europe. *Proc. Natl. Acad. Sci. USA*, **99**, 6075–6079.
- Gould, S.J. and Vrba, E.S. (1982) Exaptation – a missing term in the science of form. *Paleobiology*, **8**, 4–15.
- Gould, S.J. (1996) *Full House. The Spread of Excellence from Plato to Darwin*. Harmony Books, New York.
- Gould, S.J. (2002) *The Structure of Evolutionary Theory*. The Belknap Press of Harvard University Press, Cambridge, MA.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.
- Haruna, I. and Spiegelman, S. (1965) Autocatalytic synthesis of a viral RNA in vitro. *Science*, **150**, 884–886.
- Hayes, B. (1999) Experimental Lamarckism. *Am. Sci.*, **87**, 494–498.
- Herbert, A. and Rich, A. (1999) RNA processing and the evolution of eukaryotes. *Nat. Genet.*, **21**, 265–269.
- Hüttenhofer, A. and Brosius, J. (2002) Experimental RNomics. In *Functional Genomics* (Eds. Galperin, M. and Koonin, E.V.). Horizon Scientific Press, New York.
- Hüttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachelier, J.-P. and Brosius, J. (2001) RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.*, **20**, 2943–2953.
- Jablonka, E. and Lamb, M.J. (1995) *Epigenetic Inheritance and Evolution in the Lamarckian Dimension*. Oxford University Press, Oxford, UK.
- Jablonka, E., Lamb, M.J. and Avital, E. (1998) Lamarckian mechanisms in Darwinian evolution. *Trends Ecol. Evol.*, **13**, 206–210.
- James, K.D. and Ellington, A.D. (1995) A search for missing links between self-replicating nucleic acids and the RNA world. *Origins Life Evol Biosphere*, **25**, 515–530.
- Jantsch, E. (1979) *Die Selbstorganisation des Universums: Vom Urknall zum Menschlichen Geist*. Carl Hanser Verlag, München, Germany.
- Jeffares, D.C., Poole, A.M. and Penny, D. (1998) Relics from the RNA world. *J. Mol. Evol.*, **46**, 18–36.
- Joyce, G.F. (2002) The antiquity of RNA-based evolution. *Nature* **418**, 241–221.
- Jurka, J. (1998) Repeats in genomic DNA: mining and meaning. *Curr. Opin. Struct. Biol.*, **8**, 333–337.
- Kauffman, S.A. (1993) *The Origins of Order. Self-Organization and Selection in Evolution*. Oxford University Press, Oxford, UK.
- Kidwell, M.G. and Lisch, D. (1997) Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA*, **94**, 7704–7711.
- Kiedrowski, G.v. (1986) A self-replicating hexadeoxynucleotide. *Angew. Chem. Int. Ed. Engl.*, **25**, 932–935.
- Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, **31**, 147–157.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science*, **294**, 853–858.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczy, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
- Lau, N.C., Lim, L.P., Weinstein, E.G. and Bartel, D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*, **294**, 858–862.
- Lee, R.C. and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*, **294**, 862–864.
- Lewis, E.B. (1951) Pseudoallelism and gene evolution. *Cold Spring Harbor Symp. Quant. Biol.*, **16**, 159–174.
- Li, T. and Nicolaou, K.C. (1994) Chemical self-replication of palindromic duplex DNA. *Nature*, **369**, 218–221.

- Lyon, M.F. (1998) X-chromosome inactivation: a repeat hypothesis. *Cytogenet. Cell Genet.*, **80**, 133–137.
- Lyon, M.F. (2000) LINE-1 elements and X chromosome inactivation: a function for 'junk' DNA? *Proc. Natl. Acad. Sci. USA*, **97**, 6248–6249.
- Maizels, N. and Weiner, A.M. (1987) Peptide-specific ribosomes, genomic tags, and the origin of the genetic code. *Cold Spring Harbor Symp. Quant. Biol.*, **52**, 743–749.
- Makalowski, W. (2000) Genomic scrap yard: how genomes utilize all that junk. *Gene*, **259**, 61–67.
- Margulis, L. (1970) *Origin of Eukaryotic Cells*. Yale University Press, New Haven, CT.
- Martignetti, J.A. and Brosius, J. (1993a) BC200 RNA: a neural RNA polymerase III product encoded by a monomeric Alu element. *Proc. Natl. Acad. Sci. USA*, **90**, 11563–11567.
- Martignetti, J.A. and Brosius, J. (1993b) Neural BC1 RNA as an evolutionary marker: guinea pig remains a rodent. *Proc. Natl. Acad. Sci. USA*, **90**, 9698–9702.
- Mattick, J.S. (2001) Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep.*, **2**, 986–991.
- Maynard Smith, J. and Szathmáry, E. (1995) *The Major Transitions in Evolution*. Oxford University Press, Oxford, UK.
- Mayr, E. (1960) The emergence of evolutionary novelties. In Tax, S. (ed.) *Evolution after Darwin*, Vol. 1. The University of Chicago, Chicago, IL, pp. 349–380.
- Mayr, E. (2001) *What Evolution Is*. Basic Books, New York.
- McDonald, J.F. (1993) Evolution and consequences of transposable elements. *Curr. Opin. Genet. Dev.*, **3**, 855–864.
- McDonald, J.F. (1995) Transposable elements: possible catalysts of organismic evolution. *Trends Ecol. Evol.*, **10**, 123–126.
- Mendel, G. (1866) Versuche über Pflanzenhybriden. *Verhandl. naturforsch. Ver. Brünn*, **4**, 3–47.
- Mendel, G. (1870) Über einige aus künstlicher Befruchtung gewonnene Hieraciumbastarde. *Verhandl. naturforsch. Ver. Brünn*, **8**, 26–31.
- Muller, H.J., Prokofyeva-Belgovskaya, A.A. and Kossikov, K.V. (1936) Unequal crossing-over in the Bar mutant as a result of duplication of a minute chromosome section. *C.R. (Doklady) Acad. Sci. URSS*, **1**, 87–88.
- multiple authors. (1987) *Evolution of Catalytic Function*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Nei, M. (1969) Gene duplication and nucleotide substitution in evolution. *Nature*, **221**, 40–42.
- Ohno, S. (1970) *Evolution by Gene Duplication*. Springer, New York, NY.
- Orgel, L.E. (1968) Evolution of the genetic apparatus. *J. Mol. Biol.*, **38**, 381–393.
- Orgel, L.E. (1998) The origin of life – a review of facts and speculations. *Trends Biochem. Sci.*, **23**, 491–495.
- Pollard, J.W. (1984) Is Weissmann's barrier absolute?, pp. 291–314, In *Beyond Neo-Darwinism. An introduction to the new evolutionary paradigm*, edited by M.-W. Ho and P.T. Saunders. Academic Press, London.
- Poole, A.M., Jeffares, D.C. and Penny, D. (1998) The path from the RNA world. *J. Mol. Evol.*, **46**, 1–17.
- Robertson, D.L. and Joyce, G.F. (1990) Selection *in vitro* of an RNA enzyme that specifically cleaves singlestranded DNA. *Nature* **344**, 467–468.
- Shapiro, J.A. (1992) Natural genetic engineering in evolution. *Genetica* **86**: 99–111.
- Silver, L.M. (1997) *Remaking Eden. Cloning and Beyond in a Brave New World*. Avon, New York, NY.
- Steele, E.J., Lindley, R.A. and Blanden, R.V. (1998) *Lamarck's Signature: How retrogenes are Changing Darwin's Natural Selection Paradigm*. Perseus, Cambridge, MA.
- Stephens, S.G. (1951) Possible significance of duplication in evolution. *Adv. Genet.*, **4**, 247–265.
- Storz, U. (2002) Counting all genes: how many other RNAs exist and what do they do? *Science*, **296**, 1260–1263.
- Sturtevant, A.H. (1925) The effects of unequal crossing over at the Bar locus in *Drosophila*. *Genetics*, **10**, 117–147.
- Sutcliffe, J.G., Milner, R.J., Bloom, F.E. and Lerner, R.A. (1982) Common 82-nucleotide sequence unique to brain RNA. *Proc. Natl. Acad. Sci. USA*, **79**, 4942–4946.
- Sverdlov, E.D. (2000) Retroviruses and primate evolution. *BioEssays*, **22**, 161–171.
- Szathmáry, E. (1990) Towards the evolution of ribozymes. *Nature*, **344**, 115.
- Szathmáry, E. (1997) Origins of life. The first two billion years. *Nature*, **387**, 662–663.
- Szathmáry, E. (1999a) The first replicators. In *Levels of Selection* (Ed. Keller, L.). Princeton University Press, Princeton, NJ, pp. 31–52.
- Szathmáry, E. (1999b) The origin of the genetic code: amino acids as cofactors in an RNA world. *Trends Genet.*, **15**, 223–229.
- Szathmáry, E. and Smith, J.M. (1995) The major evolutionary transitions. *Nature*, **374**, 227–232.
- Szostak, J.W., Bartel, D.P. and Luisi, P.L. (2001) Synthesizing life. *Nature*, **409**, 387–390.
- Temin, H.M. (1970) Viral RNA-dependent DNA polymerase. *Nature*, **226**, 1211–1213.
- Tiedge, H., Chen, W. and Brosius, J. (1993) Primary structure, neural-specific expression, and dendritic location of human BC200 RNA. *J. Neurosci.*, **13**, 2382–2390.
- Tiedge, H., Freneau, R.T., Weinstock, P.H., Arancio, O. and Brosius, J. (1991) Dendritic location of neural BC1 RNA. *Proc. Natl. Acad. Sci. USA*, **88**, 2093–2097.
- Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Bidick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei,

- M., Wides, R., Xiao, C., Yan, C., et al. (2001) The sequence of the human genome. *Science*, **291**, 1304–1351.
- Wächtershauser, G. (1992) Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Prog. Biophys. Mol. Biol.*, **58**, 85–201.
- Walbot, V. (1996) Sources and consequences of phenotypic and genotypic plasticity in flowering plants. *Trends Plant Sci.*, **1**, 27–32.
- Wang, W., Brunet, F.G., Nevo, E. and Long, M. (2002) Origin of sphinx, a young chimeric RNA gene in *Drosophilamelanogaster*. *Proc. Natl. Acad. Sci. USA*, **99**, 4448–4453.
- Watson, J.B. and Sutcliffe, J.G. (1987) Primate brain-specific cytoplasmic transcript of the Alu repeat family. *Mol. Cell. Biol.*, **7**, 3324–3327.
- Weiner, A.M., Deininger, P.L. and Efstratiadis, A. (1986) Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.*, **55**, 631–661.
- Weismann, A. (1892) *Das Keimplasma: Eine Theorie der Vererbung*. Gustav Fischer, Jena, Germany.
- Weismann, A. (1893) *The Germ-plasm: a Theory of Heredity*. Walter Scott, London, UK.
- Weismann, A. (1902) *Vorträge über Descendenztheorie gehalten an der Universität zu Freiburg im Breisgau*. Gustav Fischer, Jena, Germany.
- Westheimer, F.H. (1986) Polyribonucleic acids as enzymes. *Nature*, **319**, 534–535.
- Williams, G.C. (1966) *Adaptation and Natural Selection*. Princeton University Press, Princeton, NJ.
- Wills, C. and Bada, J. (2000) *The Spark of Life: Darwin and the Primeval Soup*. Perseus Publishing, Cambridge, MA.
- Wilson, D.S. and Szostak, J.W. (1999) In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.*, **68**, 611–647.
- Woese, C.R. (1967) *The Genetic Code: The Molecular Basis for Genetic Expression*. Harper and Row, New York, NY.
- Woese, C.R. (1980) Just So Stories and Rube Goldberg machines: Speculations on the origin of the protein synthetic machinery. In *Ribosomes: Structure, Function, and Genetics* (Eds. Chambliss, G., Craven, G.R., Davies, J., Davis, K., Kahan, L. and Nomura, M.). University Park Press, Baltimore, MD, pp. 357–373.
- Woese, C.R. (2001) Translation: in retrospect and prospect. *RNA*, **7**, 1055–1067.
- Woese, C.R. (2002) On the evolution of cells. *Proc. Natl. Acad. Sci. USA*, **99**, 8742–8747.

## Glossary

**RNA world**, a very early stage of cellular life, in which RNA biopolymers encoded genetic information and had catalytic and structural functions in the cell. Consequently there was no separation of phenotype and genotype.

**RNP world**, RNA/protein world, a subsequent stage of cellular life, in which genetically encoded polypeptides joined RNA as catalytically and structurally functional biopolymers.

**DNA world**, a further stage of cellular life, in which RNA as genetic material has gradually been exchanged and superseded by DNA, presumably by enzymatic conversion of RNA into DNA.

**Retroposition**, the conversion of RNA into DNA by reverse transcriptase. The resulting complementary DNA (cDNA) is more or less randomly integrated into the genome. Apparently, this proceeds rather intensely and unabated in many lineages to this day.

**Exaptation**, a nuon, gene, organ, or other features that now enhance fitness but were not built by natural selection for their current role (Gould and Vrba, 1982). Bird's feathers are exaptations as initially they served for temperature regulation and some were, at later stages, co-opted for flight. Another example is sex. Its major evolutionary advantage has been exchange of genetic material. It has also been exapted in many species for bonding between individuals, thus providing additional advantage to both the offspring and in some instances entire social groups. Charles Darwin already described the process of exaptation in his 'Origin of Species', page 398 (Darwin, 1872):

'Again, an organ may become rudimentary for its proper purpose, and be used for a distinct one: in certain fishes the swim-bladder seems to be rudimentary for its proper function of giving buoyancy, but has become converted into a nascent breathing organ or lung. Many similar instances could be given.'

**Gene**, originally a unit of heredity. The term underwent numerous modifications as knowledge in Molecular Biology accumulated. For example, its regulatory regions are sometimes included, sometimes it

refers only to the DNA that is transcribed into RNA. Today, the gene concept is fuzzy.

**Nuon**, any discrete segment of nucleic acid, RNA or DNA (Brosius and Gould, 1992, 1993). A nuon can be defined by sequence similarity (repetitive elements) or function (exon, intron, promoter, enhancer, splice site), by its biogenesis (see retronuon), evolutionary fate or effect (e.g. naptionuon to indicate current non-aptation or xaptionuon for an exapted nuon). An advantage of nuon is that it not only denotes a purely genotypic nucleic acid but, also a biopolymer that unites both phenotype and genotype. This is not possible when using the term gene and hence very important for discussing the RNA and RNP worlds. For further information see: <http://exppc01.uni-muenster.de/expath/retronuons.htm>.

**Retronuon**, any nuon generated by reverse transcription of any RNA, viral RNA, messenger RNA or untranslated RNA yielding short interspersed repetitive elements (SINES), long interspersed repetitive elements (LINEs) etc.

**snmRNA**, small (~50–500 nucleotides in length) non-messenger RNA. Some snmRNA species can be highly efficient templates for reverse transcriptase and are the founders for SINES.

**Gene Duplication**, Susumu Ohno described in his book (Ohno, 1970) the importance of gene duplication for the evolution of genes, genomes and organisms. Although, by some, only the gene is considered the unit of selection (Williams, 1966; Dawkins, 1982), the concept might be expanded to signify 'unit' duplication if applied to nuons, cells, organs, organisms, social groups and states. As Ernst Mayr pointed out in 1960 (Mayr, 1960), Darwin already recognized the principle of duplication on page 147 (Darwin, 1872):

'Again, two distinct organs, or the same organ under two very different forms, may simultaneously perform in the same individual the same function, and this is an extremely important means of transition...'

Amplification of units constitutes, together with random variation and selection, the basis of Neo-Darwinian evolution.

**Darwinism**, species have evolved from simpler ancestral types by the process of natural selection acting on the variability found within a population.

**Neo-Darwinism**, Darwinism merged with the Mendelian laws of genetics; makes use of the modern knowledge of chromosomes and genes to explain the source of the genetic variation upon which selection works. Darwin's (and Ohno's) concepts of unit duplication might be included to yield the evolutionary chain of amplification, variation (including recombination) and selection.

**Lamarck**, less well known for his theory of speciation through gradual change that had been published five decades before Darwin's, than for 'his' error that evolution took place through inheritance of modifications caused by the environment, and the effects of use and disuse of organs. In the strict sense such effects act directly and non-randomly on the gene(s) responsible for the respective phenotypes. For facts and fiction concerning Lamarck's ideas and contributions, see Ghiselin (1994).

**Neo-Lamarckism**, as understood today is more broadly defined as 'the evolutionary mechanism of the inheritance of acquired characteristics' (Bowler, 1992). Examples of Neo-Lamarckism emerge, wherever Weismann's Barrier recedes if one allows for its definition as the inheritance of acquired characteristics but barring directed modification of genetic material (see also Jablonka and Lamb, 1995).

**Darwinian Threshold**, a hitherto unrecognized phase transition in the evolutionary process (Woese, 2002). It corresponds to the emergence of a new level of order in terms of cell organization. Cellular machineries became more and more complex and interconnected. As a consequence, the interchangeability of genetic material became gradually less feasible and thus curbed the acquisition of novelties via horizontal gene transfer. Consequently short lived entities were replaced by more permanent ones - species. Hence, 'the Darwinian Threshold truly represents the origin

of speciation as we know it' (Woese, 2002). Woese's concept might be misunderstood as an assault on the foundations of Darwinism. In contrast, it is solidly anchored on the three Neo-Darwinian pillars of amplification, modification and selection (see above). With his recent proposal, Woese has lifted the origin of species, always guided by Neo-Darwinian principles, out of the muddy layer of communal evolution, across the Darwinian Threshold onto a level where separatism allowed for the establishment of relatively stable 'species'. In molecular terms, Woese sees this transition occurring when templated translation produced ever more sophisticated protein molecules including the ones that are involved in the translational machinery itself.

**Weismann's Barrier**, recognized that the germ plasm remains unaffected by any changes affecting somatic cells during the life time of an organism. Weismann's germ plasm theory (Weismann, 1892, 1893, 1902) based on Mendel's laws (Mendel, 1866, 1870) debunked Lamarckism. At the same time Weismann cemented the Neo-Darwinian doctrine. Yet, Weismann's Barrier applies only to a segment of all known forms of life, namely those that feature a germ line and do not permit genetic 'contributions' from somatic cells. Weismann's Barrier does not apply to unicellular organisms or multicellular organism that do not maintain a germ line or do so only temporarily by alternating between sexual and clonal reproduction, such as many plants (Pollard, 1984; Walbot, 1996). An acquired mutation in a somatic cell that is beneficial for the respective plant may be passed on to future generations by somatic reproduction and subsequent switching to sexual reproduction.

**Central Dogma of Molecular Biology**, states that DNA makes RNA makes protein, in that order (Crick, 1958). This was amended twice; when S. Spiegelman showed that RNA could be replicated from RNA templates (Haruna and Spiegelman, 1965) and H. Temin and D. Baltimore discovered reverse transcriptase (Baltimore, 1970; Temin, 1970). If its often considered a molecular application of Weismann's Barrier.