MINI REVIEW

Evolution of the capacity to evolve

R. L. CARROLL

Redpath Museum, McGill University, Montreal, Canada

Keywords:

developmental pathways;
evolution of life;
exon shuffling;
gene recruitment;
genome evolution;
horizontal genetic exchange;
origin of life;
regulatory elements;
RNA world.

Abstract

During the past two decades, the fields of molecular biology and genetics have enabled study of far broader and more detailed aspects of evolutionary change than were possible when the evolutionary synthesis was elaborated in the mid-twentieth century. The capacity for complete sequencing of both genes and proteins of all groups of organisms provide, simultaneously, the means to determine both the patterns and processes of evolution throughout the history of life. Increased knowledge of the genome documents the changing nature of its composition, mode of transmission, and the nature of the units of selection. Advances in evolutionary developmental biology demonstrate the conservation of genetic elements throughout multicellular organisms, and explain how control of the timing, position and nature of their expression has produced the extraordinary diversity of living plants and animals. The next generation of evolutionary biologists will benefit greatly from the increased integration of these new fields of research with those that are currently emphasized in the standard textbooks and journals.

Introduction

> …we can see that the Modern Synthesis of the mid-20th century was but a stage…in the elucidation of the history of life on Earth. In all likelihood, the past decade and the coming ones will prove equally significant as a second distinct stage in this quest (Wilkins, 2002, p. 523).

The study of evolution has been a progressively expanding field of research. It was based initially on diverse observations of the earth's current biota, but became integrated through a succession of explanatory hypotheses and tested through experimentation in the field and increasingly by analysis in the laboratory. The work of Darwin (1859) concentrated primarily on aspects of the natural history of living organisms, supported by knowledge of breeding of plants and animals, biogeography, and patterns of embryological development. The conclusions he drew from the integration of these subjects resulted in the rapid and general acceptance, by the scientific community, that all organisms on earth had ultimately evolved from a common ancestry through gradual change over hundreds of million to billions of years. However, it was not until the discovery of the nature of inheritance by Mendelian genetics and the subsequent formulation of population genetics (Fisher, 1930; Haldane, 1932) that Darwin's theory of natural selection was accepted as a means for evolutionary change through interactions between the production of genetic variations within populations and their differential survival from generation to generation.

The next major expansion of evolutionary thought was the integration of the genetical theory of natural selection with increased knowledge of the nature of species and speciation (Dobzhansky, 1937; Mayr, 1942) and of the fossil record (Simpson, 1944), which together formed the basis for a new evolutionary synthesis between the 1930s and 1950s (Mayr & Provine, 1980). Information and concepts brought together by this synthesis have been extremely effective in the analysis of evolutionary change at the level of populations and species, especially among sexually reproducing plants and animals. Larger scale evolutionary phenomena have generally been assumed to be explicable by comparable processes, occurring over much longer periods of time. This approach has continued to form the basis for a wealth of papers published in major journals, including *Evolution, Journal of Evolutionary Biology, and Biological Journal of the Linnean Society* (subtitled: *A Journal of Evolution*). The subject matter of these

Correspondence: Robert L. Carroll, Redpath Museum, McGill University, 859 Sherbrooke St. West, Montreal, H3A 2K6, Canada.
Tel.: (514) 398-4086-4090; fax: (514) 398-3185;
e-mail: robertc@shared1.lan.mcgill.ca

journals has, in turn, served as the primary basis for the production of textbooks that have served to educate the current generation of evolutionary biologists.

However, as documented in Mayr & Provine (1980), many important aspects of evolution were not included in the evolutionary synthesis. Notably missing were the consideration of the patterns and processes of evolution among micro-organisms, the problem of mass extinction, and the apparent rapidity of major radiations. The most striking omission was the divorce from embryology, which had been such an important factor in the acceptance of evolution in the nineteenth century (Hamburger, 1980). Most of these problems had not been considered at the time the synthesis gained its dominance, and indeed few were capable of solution on the basis of the knowledge and technologies then available.

The study of evolution now stands at a point comparable with that in the 1930s. A great range of new research programmes bearing on evolution have since developed and a very extensive literature is accumulating, but this information has not yet been fully synthesized or integrated into the widely used text books. The most significant advances are in the capacity to sequence genes and proteins, which provides a reliable means both for establishing the interrelationships of all organism (Doolittle, 1999; Peterson & Eernisse, 2001), and to document the precise mechanisms by which evolutionary change occurs. From geology and palaeontology have come a vastly increased knowledge of the fossil record and the capacity for much greater accuracy of dating to document the time of occurrence and duration of major evolutionary events (e.g. Bowring *et al.*, 1993; Knoll & Carroll, 1999).

Areas of study that have particularly benefited from these advances include investigation of the origin of life (e.g. Zubay, 2000), understanding of the nature of the genetic material in prokaryotes, its mode of transmission, and patterns of evolution (Thomas, 2000; Staley & Reysenbach, 2002), and analysis of how pathways of genetic expression control development in multicellular organisms and how specific changes in the regulatory elements of these genes have led to evolutionary change (Carroll *et al.*, 2001; Davidson, 2001; Wilkins, 2002). Journals specifically dealing with these subjects include *Origins of Life and Evolution of the Biosphere, Journal of Molecular Evolution, Molecular Biology and Evolution, Molecular Phylogenetics and Evolution, Evolution and Development, and Development, Genes and Evolution.*

The integration of these new fields is clearly a large task, for the areas of evolutionary research are vast, diverse and involve an extensive new technical vocabulary. Nevertheless, they offer the basis for understanding some of the most fundamental questions of evolutionary biology.

This short essay will concentrate on a single approach to the study of evolution – investigation of the changes that have occurred in the nature of the genome and its means of transmission and regulation that have resulted in an increased capacity for evolutionary change over the history of life. It is intended as but an example of the explosive growth in new ways by which evolution can be studied (Fig. 1).

The RNA World

We can begin with the earliest and most primitive molecules that were capable of carrying genetic information, but which can be traced directly to all subsequent forms of life (Joyce, 2002). The simplest hereditary molecule that is capable of accurate self replication is RNA. Its constituent ribose sugars, nucleotides, and phosphates could have formed and polymerized randomly under the conditions hypothesized for the atmosphere and surface waters of the earth as soon at it had cooled to temperatures below 100 °C, between 4.2 and 3.8 billion years ago (Zubay, 2000) (Wilde *et al.*, 2001). At this point, a major problem in understanding the evolution of life is to explain how RNA could have replicated without the prior existence of appropriate catalysts. Accurate synthesis of RNA in modern cells requires a pre-existing template and the presence of specific protein catalysts. If the sequence of monomers of both the RNA and the catalyst were dependent on the other, how could either have evolved initially? The solution to this problem was suggested by Woese (1967) and Crick (1968): that RNA itself should be capable of serving as a catalyst. In particular, RNA's capacity for complex folding would facilitate its close integration with molecules required for polymerization and other biochemical processes.

This hypothesis led to the concept of the RNA world (coined by Gilbert, 1986), a period during which the only nucleic acid was RNA, and when critical organic reactions were catalysed by regions of the RNA itself (Gesteland *et al.*, 1999). Evidence for RNA acting as a catalyst in modern organisms was first gained from studies of its splicing, which was shown to occur in the absence of enzymes in some organisms (Kruger *et al.*, 1982; Zaug & Cech, 1986). From this, it was assumed that a segment of the RNA (termed a ribozyme) was serving a catalytic function. This has been recently confirmed (Valadkhan & Manley, 2001). Naturally occurring ribozymes have since been found to catalyse a number of other biochemical reactions (Doudna & Cech, 2002). The most important is that of RNA, acting independent of proteins, making up the active site in the ribosomes for protein synthesis (Newman, 2001). Other ribozymes catalyse phosphoester transfer and cleavage and polynucleotide ligation (Joyce & Orgel, 1999), and govern the nicotinamide biosynthetic pathway (Cleaves & Miller, 2001). The ability to recover RNA segments that are capable of catalysing a number of essential biochemical processes from large, randomly produced arrays of RNA,

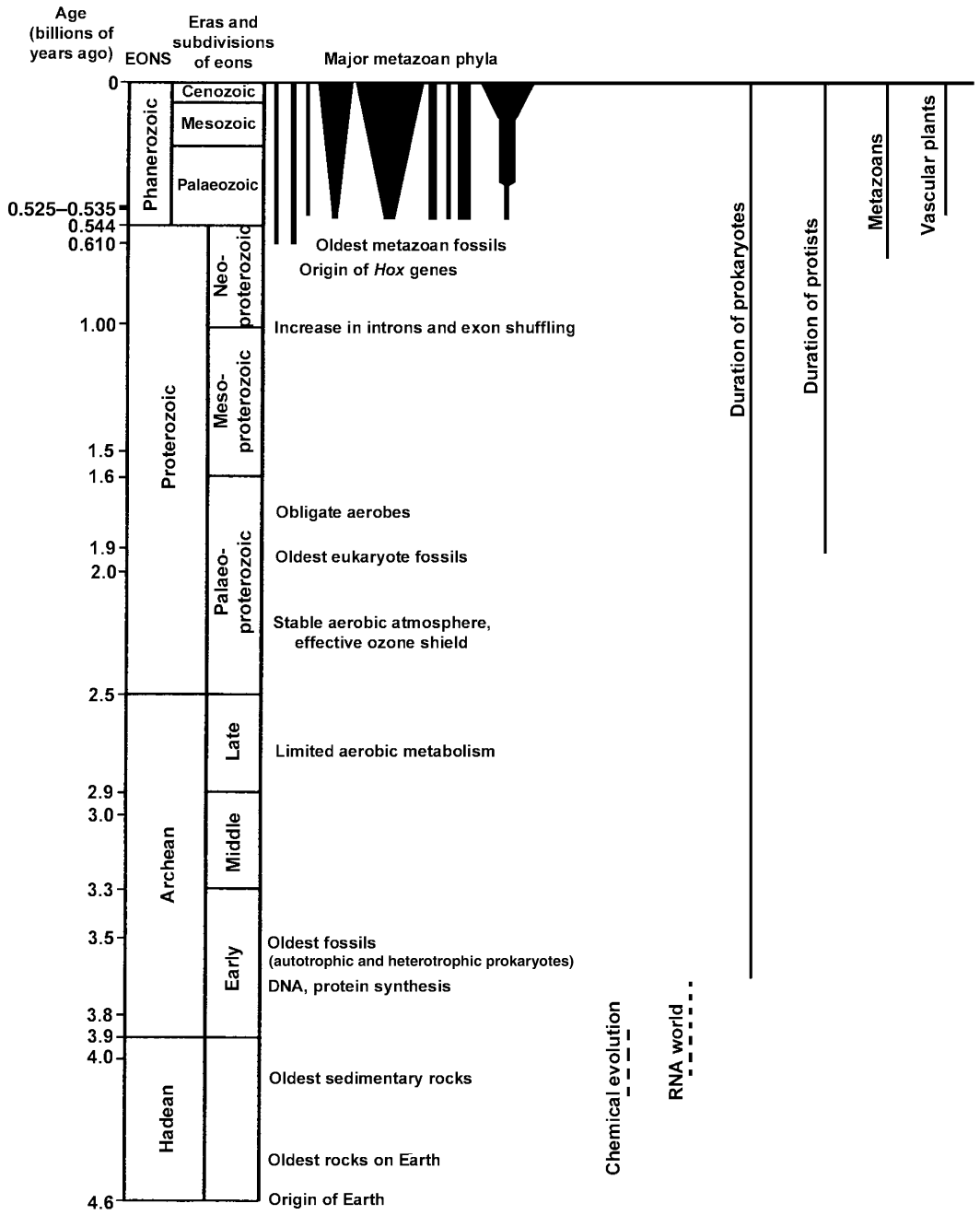

Fig. 1 Key events in the history of life associated with the evolution of the genome, from the consolidation of the earth's crust to the present.

supports the hypothesis that the origin of ribozymes, early in the history of the earth, had also resulted by chance (Bartel, 1999).

Prior to the appearance of RNA sequences that were capable of facilitating their own polymerization and replicating the initial nucleotide sequence, their synthesis was governed by chance and the relative concentration of the available nucleotides. Once a particular sequence of nucleotides appeared that had the capacity to facilitate accurate replication of the entire molecule, those molecules would have replicated more effectively, thus giving them a selective advantage over others. At this point, natural selection emerged as the primary means of maintaining sequence stability and governing the direction of change in all information bearing molecules.

It was initially thought that life evolved in open water, with the molecules free to move in any direction, providing a great range of possibilities as to what chemical combinations might occur. On the other hand, the capacity for combining a number of useful molecules in a small volume would be much greater if they were loosely attached to some surface. Both clay minerals (Zubay, 2000) and iron pyrite (Maynard Smith & Szathmáry, 1995) have been suggested as plausible substrates. Eventually, however, the common ancestor of all modern forms of life came to be contained in a phospholipid membrane. Although phospholipid membranes can be formed in the laboratory, this requires conditions that are inimical to other organic molecules such as RNA and amino acids. However, both RNA and phospholipids are sufficiently stable that they could be synthesized in different areas and then brought together by water currents.

Experiments by Deamer (1997), further discussed by Segre *et al*. (2001), demonstrate that drying followed by rehydration can result in the incorporation of RNA into phospholipid spheres. This implies the relative ease of passage of nucleic acids through phospholipid membranes, which has important consequences in early modes of genetic exchange. Random combinations of a variety of RNA strands would result in different kinds of cells with varying capacities for reproduction and metabolism. Maynard Smith & Szathmáry (1995) pointed out the selective advantage of synchronized replication that would result from the formation of multigene chromosomes.

Once primitive self-replicating molecules were incorporated within a cell membrane, the potential for accumulating a variety of organic monomers and elaborating additional biochemical pathways increased immensely. It was presumably only at this stage that regulated polymerization of amino acids into enzymes became possible (Zubay, 2000).

Although self-replication of RNA and the catalysis of other biosynthetic processes by ribozymes may mark the beginning of truly evolutionary processes, RNA as a basis for the transmission of hereditary information was inherently limited by its chemical instability, the great potential for accumulating errors during replication and limited ability of repair (Maynard Smith & Szathmáry, 1995). Very early in the history of life, there must have been selection for mutations resulting in the removal of oxygen from ribose sugar to form deoxyribose and methylation of uracil to produce thymine (which also occur during the synthesis of DNA in modern cells; Gilbert & de Souza, 1999). This led in a much more stable chemical configuration for DNA compared with RNA. In addition, the obligatory pairing of complementary strands of DNA and the proof reading capacity of DNA polymerase result in a much lower mutation rate. Whereas strands of RNA are generally limited to less than 5000 nucleotides, with the capacity to code for about five protein-based enzymes, DNA extends the entire length of chromosomes, encompassing tens of thousands of genes. With DNA serving the primary role of replication, RNA molecules became specialized for various aspects of protein synthesis.

Evolution among prokaryotes

The first direct evidence of living organisms is provided by fossils from Australia dated from approximately 3.5 billion years ago (Schopf, 1999). These include cells less than 3.5 μm in diameter (preserved in chert) comparable with those of modern heterotrophic bacteria and larger cells associated with laminar deposits of calcium carbonate that resemble those formed by living bluegreen algae or cyanobacteria. The size and presumed photosynthetic activity of these cells suggest that they had already evolved many of the metabolic pathways common to living bacterial species. This implies a relatively rapid rate of biochemical evolution during the previous 300 million years, since the first extensive evidence of liquid water on the surface of the earth.

On the other hand, the structural evolution of bacteria in the subsequent 3.5 billion years seems to have been extremely slow, to judge by the near identity of cell size and details of surface features visible in fossils representing a multitude of lineages, as well as the similarity of community assemblages (Schopf, 1995). This is in marked contrast with the very rapid rate of evolution observed among living bacteria in their adaptation to changing environments and ways of life.

Prokaryotes, including both Eubacteria and Archaea, show a very limited range in the size of their genome relative to that of eukaryotes. From the smallest to the largest bacterium, the range is only about 26-fold, whereas that of eukaryotes is approximately 80 000-fold (Graur & Li, 2000). Most prokaryotes are limited to a single circular chromosome that has only a single point for initiating replication. Because of its small size, bacteria can replicate very quickly, but this advantage would be lost if the size of the genome were significantly increased. The minimum number of chromosomal genes necessary for the basic processes of metabolism and reproduction among living heterotrophic bacteria may be as small as about 260 (Itaya, 1995). This presumably corresponds with the complement present in ancestral prokaryotes. However, a large number of genes that are *not* integrated into the chromosome are also present in many modern bacteria that have been studied.

It was long assumed that bacteria could evolve rapidly because large numbers of mutations would accumulate over a short period of time among the enormous number of progeny. However, the fact that antibiotic resistance, achievement of toxicity, and capacity to feed on different nutrients frequently appeared nearly simultaneously in different strains and species suggested that some other method of evolution is involved than the progressive

accumulation of new mutations in each of the lineages. It is now recognized that many of these changes are not occurring within the chromosomes of individual lineages, but are the result of exchange of genes from sources outside the cell, which are not integrated into the chromosome (Summers, 1996; Ochman *et al.*, 2000; Thomas, 2000; Bushman 2002). This is referred to as horizontal genetic transmission, as opposed to the vertical transmission that occurs from generation to generation within lineages. In this way, many ecological and physiological adaptations among modern Eubacteria arise in a fundamentally different manner than those of multicellular eukaryotes, but show similarity with the random entrance of RNA into phospholipid spheres hypothesized for the RNA world.

There are several ways in which DNA can be exchanged among bacteria and between bacteria and higher organisms: (1) *Transformation* involves the uptake of naked DNA from the environment. (2) *Transduction* occurs via the introduction of a bacteriophage into a bacterium [upwards of 100 kilobases (kbs) may be carried in a phage capsid]. (3) *Conjugation* requires immediate physical contact between donor and recipient cells. This usually involves transfer between prokaryotes, but also, more rarely, between bacteria and yeast and vascular plants. Numerous bacteria have specific uptake systems to receive DNA from the general environment or from other organisms (Summers, 1996). Sonea & Mathieu (2000) speak of the essentially universal horizontal exchange of genetic material throughout the bacterial world.

These additional genes occur within the cell in a variety of forms. The most conspicuous are the plasmids, which may be in as large as 1.7 million base pairs in size, and can carry all the genes necessary for any of a variety of pathways, including drug resistance, protection from toxic metals, nitrogen fixation, and for rhizobial associations with plants (Downie & Young, 2001). The complex of genes in a plasmid shares a basic replicon that coordinates the timing of replication with that of the host genome, and limits the number of copies to avoid overtaxing the host. The presence of particular plasmids will be selected against if the environment changes and they are no longer beneficial to the host (Espinosa *et al.*, 2000; Lawrence, 2001). Other elements that can enter the cells of prokaryotes are those of the phage genome, and a variety of transposons. The amount of extrachromosomal genetic material in bacterial cells is commonly underestimated because sequencing is typically limited to the chromosomal portion.

More permanent integration of foreign DNA within the bacterial genome can occur through insertion into the chromosome via transposable elements. Transposons may lie within the genome of phages, may incorporate entire plasmids, or they may more closely resemble the transposons in eukaryotes, which are very important in generating recombination within chromosomes while replicating themselves and adjacent elements. By these means, some bacteria achieve a high degree of recombination among chromosomal genes without sexual reproduction (Merlin *et al.*, 2000).

The long-term consequences of horizontal genetic exchange among prokaryotes is demonstrated by the amount of genetic material that has a different source than most of the chromosomal genes. Among bacteria in which the entire genome has been sequenced, the amount of horizontally acquired DNA is conservatively estimated at 12.8% in *Escherichia coli*, 16.8% in *Synechocystis,* and 7.5% in *Bacillus subtilis* (Ochman *et al.*, 2000). This is based on atypical G + C content and pattern of codon usage. Twenty-four per cent of the hyperthermophilic Eubacteria *Thermotoga maritima's* open reading frames display great similarity to those of archaeal bacteria. Conservation of gene order between *T. maritima* and Archaea in many of the clustered regions suggests that lateral gene transfer occurred between the two major groups of prokaryotes (Nelson *et al.*, 1999). According to Woese (2002, p. xxvi) 'There can no longer be any doubt that horizontal gene transfer is not only a force in genome evolution but a major determining one.'

The potential for much larger chromosomes in eukaryotes obviates the necessity for plasmid like structures that can be picked up and discarded as needed, but transposons and the numerous repetitive elements that they generate remain as significant components of the chromosomes of advanced eukaryotes (Lawrence, 2001).

Eukaryotes

Origins

The greatest dichotomy among living organisms, in both structure and means of reproduction and evolution, lies between prokaryotes and eukaryotes. The first fossil evidence of eukaryotes is from about 1.9 billion years ago, based on fossils of cells approximately 10 times the diameter of those of prokaryotes (Schopf, 1999). Judging by the most primitive members of each of the major living eukaryote groups, their common ancestors in the Proterozoic had probably already acquired a phospholipid membrane separating the nucleus from the rest of the cell, linear chromosomes capable of replication at many points, microtubules capable of forming a mitotic figure, and a cytoskeleton that maintained the integrity of the cell in the absence of a cell wall.

The relatively larger size of ancestral eukaryotes and the absence of a cell wall may have enabled them to feed by engulfing prey, or to have their cells invaded by smaller prokaryotes. One or other of these processes presumably led to the endosymbiotic relationship between nearly all eukaryotes and mitochondria, whose origin can be traced to free living aerobic, purple-nonsulphur Eubacteria, and the chloroplasts of plants, derived from blue-green algae (Moreira *et al.*, 2000).

Endosymbiosis may be considered as an extreme example of horizontal transfer of genetic material, in which the entire genome of one type of organism is transferred to another. This clearly has little in common with the successive accumulation of point mutations within lineages or the long-term selection of alternative alleles that are thought to be the primary means of evolutionary change in eukaryotes. Somewhat in analogy with the plasmids of bacteria, the mitochondria and chloroplasts retain their own, essentially bacterial genome, separate from the nuclear chromosomes, but in most species a large number of genes are transferred into the nucleus (Selosse *et al.*, 2001). Although subsequent endosymbiotic events can be recognized among various eukaryotic lineages, the specific events incorporating mitochondria and chloroplasts appear to be unique, based on the great similarity of the genes that are retained in these organelles among all animals and plants that possess them, and the near certainty of their identity with those of particular groups of living prokaryotes.

Unicellular eukaryotes

Maynard Smith & Szathmáry (1995) provided informative models of how the patterns of chromosome replication changed between prokaryotes and eukaryotes, leading to modern mitosis and meiosis. Of particular importance is the fact that disjunction of the circular chromosomes of prokaryotes is impossible if they have undergone crossing over, effectively precluding recombination. It is only among eukaryotes with paired, linear chromosomes, that genes can be readily exchanged between homologous chromosomes during meiosis. Such recombination is the basis of sexual reproduction in higher organisms, and only after it has been achieved can we speak of evolution as proceeding via selection between alternative alleles.

The capacity for sexual reproduction may have evolved among the common ancestors of all eukaryotes. However, many groups of protists (Anderson, 1988) and fungi reproduce primarily asexually, and others alternate between sexual and asexual reproduction, depending on environmental factors. Sexual reproduction is only the rule (although not universal) among multicellular plants and animals. Regular recombination may have been a requirement for the origin of multicellular organisms with a high degree of cell differentiation, although the correlation between these processes is not well established (Butterfield, 2000).

On the other hand, another mode of evolutionary change became very important among single cell eukaryotes. This is the phenomenon of exon shuffling (Graur & Li, 2000). It was long thought that genes acted essentially as single units, to code for particular proteins. It is now recognized that most genes in eukaryotes are composed of several functionally distinct elements. The portions of genes that code for proteins are termed exons, between which are noncoding regions called introns. Other units, adjacent to the protein coding portion, the promoter and regulatory elements, govern the activation and timing of expression of the gene. Each of these elements can function on its own or in combination with many others (Fig. 2). The presence of introns between the translated portions of the chromosomes enables them to change position within a single chromosome, switch to different chromosomes or be duplicated (Gerhart & Kirschner, 1997; Graur & Li, 2000). The specific configuration of the introns facilitates excision and insertion of the exons without producing phase shifting and loss of function of adjacent exons. Changes in the position of

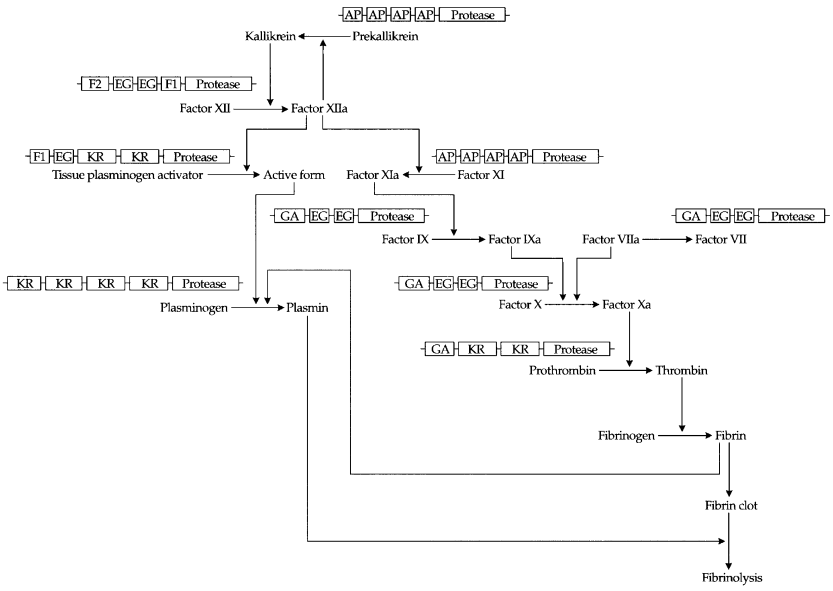

Fig. 2 Simplified diagram of blood coagulation, involving several mosaic proteins. Each series of boxes illustrates the structural modules of an individual protein, that have been assembled through exon shuffling. Each one of the modules may have been originally coded by individual genes, but they are now linked to one another in a variety of different ways. Many proteins with different functions have been formed from a smaller number of initially independent elements. AP = apple module; EG = epider-epidermal growth factor module; F1 = fibronectin type-1 module; F2 = fibronectin type-2 module; GA = γ-carboxy-glutamate domain; KR = kringle module; Protease = a serine proteinase region homologous to that of trypsin. Diagram reproduced from Graur & Li (2000).

the exons are facilitated by other elements, the transposons, which can insert in a variety of positions on the chromosome and promote insertion, deletion or duplication of the exons (Thomas, 2000).

Knoll (1995) found that the fossil record of protists during the Proterozoic shows a progressive appearance of lineages whose living members show larger and larger numbers of introns. This suggests that introns became more and more prevalent during the evolution of unicellular eukaryotes, and that this is associated with more and more complex proteins. In addition, changes and multiplication of regulatory elements would have facilitated greater degrees of interaction between genes, for a single gene may activate many others, or be activated by many others. Exon shuffling can occur through unequal crossing over during meiosis, but whether the organism is haploid or diploid, it can also occur via transposons, among organisms that are primarily or entirely asexual. By these means, the genomes of many primarily unicellular lineages became increasingly complex throughout the Proterozoic.

Evolution and development among multicellular organisms

According to Woese (2002), it has only been possible within the last 30 years to establish the patterns of relationship and processes of evolution at the bacterial level. In contrast, multicellular, sexually reproducing plants and animals have long served as the primary models for the study of evolutionary patterns and processes. However, even within these groups there has been an enormous surge of new information based on growing knowledge of the genome and especially the genetics of development, but within an even shorter time span.

We may take the studies of the Galápagos finches by Rosemary (Grant & Grant, 1989) and Peter Grant (Grant, 1999) as models of extremely thorough and detailed studies of evolutionary change in living populations and species, making use of the concepts that were elaborated within the context of the evolutionary synthesis. They focused on changes in the dimensions of the beak related to environmentally induced changes in food supplies. Modifications in the size and shape of the beaks were attributed to selection for alternative alleles of genes that coded for quantitative traits. They accepted that quantitative traits were governed by a number of essentially similar genes with primarily additive effects, although techniques available late in the twentieth century had not revealed the actual nature of such genes (Lynch & Walsh, 1998).

It was long assumed that modifications in the number and nature of genes for quantitative traits were the primary force behind long-term and large-scale evolutionary changes as well as those that could be studied in living populations. However, without detailed knowledge of their nature, the means by which they governed change could not be established. This altered dramatically when it became possible to establish the nucleotide sequence of the entire gene in multicellular organisms. This led not only to the ability to describe and compare the protein coding elements of the genes in all groups of metazoans, but showed that their expression was determined as much by regulatory elements associated with each gene as by the nature of the exons. These discoveries now make it possible to establish the specific way in which genes control development, and also how changes in these genes have affected evolution throughout the history of multicellular animals and vascular plants.

Some degree of aggregation between cells of a single type occurred as early as the first appearance of prokaryotes in the fossil record (Schopf, 1999). Living cyanobacteria have separate photosynthetic and nitrogen fixing cells (Gerhart & Kirschner, 1997), and the myobacteria and the protist *Dictyostelium* form fruiting bodies with spores that are differentiated from vegetative cells in both shape and function. Numerous cell types can be recognized in brown and red algae (Bell & Mooers, 1997), but they lack a regular body form. On the other hand, the choanoflagellates, hypothesized as close to the ancestry of both metazoans and fungi, lack cell differentiation, but may be colonial. The ultimate ancestry of land plants also lies among unicellular forms, the charophyceans within the green algae (Graham *et al.*, 2000). This indicates that cell differentiation evolved separately within the primitive lineages leading to multicellular animals and terrestrial plants.

Multiplication of cell types, organized into a specific body form, requires a novel system of genetic control and organization not present in unicellular organisms, although the potential for this type of control can be seen at the bacterial level in the function of the *lac*-operon (Jacob & Monod, 1961; Jacob, 1977). Because nearly all cells in each multicellular organism contain the same DNA, they have the potential, early in development, for differentiating into any of the different kinds of cells present in that animal. The specific way in which each cell differentiates depends on which of its genes are activated. This is determined by specific kinds of proteins, termed transcription factors, that are produced by regulatory genes. Wilkins (2002) summarized recent work showing how these genes are organized into sequential genetic pathways that control networks of development. This information forms the basis for understanding how development proceeds, and how modification of these pathways and networks lead to evolutionary change.

A major controlling element of the genetic pathways in all multicellular animals are the *Hox* genes (Carroll *et al.*, 2001). Their equivalent in plants are the MADS box genes (Graham *et al.*, 2000). *Hox* genes are also referred to as master control genes (Gehring, 1998), for they regulate the expression of structures along the main anterior-posterior axis of the body. The *Hox* genes do not directly control the

nature or shape of structures, but regulate the expression of a complex hierarchy of other genes that do so.

Hox gene evolution presumably began with the modification of particular regulatory genes, similar to homeobox genes in yeast, that gained the capacity to control where and when in the developing embryo specific cell types differentiate. This capacity may have been first applied to determine the oral and aboral end of the body in sponges and cnidarians (corals and hydra) (Finnerty & Martindale, 1998), and may initially have involved only a single primordial *Hox* gene. That gene underwent repeated duplication and divergence, to give rise to three regions of body control in bilaterians (protostomes and deuterostomes): anterior, central, and posterior, in a linear sequence. Other duplications occurred subsequently in each of these domains.

The most obvious aspect of change within the genome of metazoans was the increase in the number of *Hox* genes from one or two in the sponges and cnidarians, to about seven in the common ancestors of protostomes and deuterostomes and up to 13, carried on a single chromosome, in higher metazoans. Vertebrates were unique in undergoing a succession of duplications of the entire *Hox* complement, resulting in four *Hox* clusters on different chromosomes in birds and mammals and up to seven in some fish (Ram *et al*., 2001). The overall increase in anatomical complexity between sponges and most protostomes and deuterostomes, and between primitive chordates and advanced vertebrates may be attributed to the increase in the number of *Hox* genes, but this does not explain the extreme diversity of body plans that were evident among the many distinct phyla that appeared in the explosive radiation of metazoans at the base of the Cambrian.

Wilkins (2002) argues that the capacity for generating a diversity of body plans must have been present in the immediate ancestors of all bilaterians, and may have been initiated at the cniderian level, within which are present most of the basic molecules necessary for generating the body form in the more advanced phyla and the initial appearance of developmental pathways. All development among bilaterians is governed by hierarchical networks of regulatory genes. The *Hox* genes appear at the top of a network that regulate the position and timing of development of major body regions and structures. Expression of *Hox* genes leads to the activation of an anastomosing cascade of down-stream genes that initiate such functions as cell division and differentiation, adhesion and mobility. More local regulation may, for example, determine the exact timing and position of mesenchymal condensation, chondrification and ossification that produces bones of a particular size and shape.

Even if it has not yet been possible to determine all the details of developmental pathways, the expression of specific genes, or groups of genes during early development have already elucidated such key features as the identity of genes that determine major elements of the body plan in the arthropods, echinoderms, hemichordates and chordates. This demonstrates the basic homology of the gut, aspects of the central nervous system and the determination of dorso-ventral polarity that unite protostomes and deuterostomes (Arendt *et al*., 2001; Tagawa *et al*., 2001).

All elements of the genetic pathways are subject to mutational change that has the potential to lead to anatomical and functional modification. It was long thought that changes in the protein coding sequence were the primary means by which such changes occur. An excellent example was recently described by Galant & Carroll (2002) and Ronshaugen *et al*. (2002), in which changes in the protein coded by the *Hox* gene *Ubx* were associated with the loss of abdominal limbs in insects, compared with their more primitive arthropod ancestors.

On the other hand, Davidson (2001), Carroll *et al*. (2001), and Wilkins (2002) have shown that changes in regulatory elements within genes are also extremely important in determining when, where, and to what degree particular proteins are expressed, and may provide a more ubiquitous and flexible means for genetic and evolutionary change at all levels (e.g. McGregor *et al*., 2001). Davidson and Carroll both emphasized the importance of the *cis*-regulatory elements, also termed enhancers. There are typically several such elements associated with each gene that serve as the binding site for various transcription factors that can activate the gene in different ways. They may either promote or inhibit this response, depending on their nature, number and position relative to the coding portion of the gene. The *cis*-regulatory elements are small and easily inserted, deleted, transposed, or duplicated as a result of unequal crossing over or the activity of transposons. As their active sites may consist of only a few base pairs, they may be readily generated from segments of the introns. Selection from among different forms and positions of the *cis*-regulatory elements could explain many of the fluctuating changes that are observed in modern populations and species. In this way, *cis*-regulatory elements fill the functional role previously taken by the concept of genes for quantitative traits.

Another phenomenon that has the potential for large scale developmental and evolutionary consequences is that of gene recruitment, described in detail by Wilkins (2002). Early in metazoan evolution, developmental pathways may have consisted of no more than two or three interacting genes, but these have been augmented through the incorporation of additional genes and integration of originally distinct pathways. The primary cause of recruitment may be mutational change in an enhancer of a gene belonging to a previously established pathway that enables a different transcription factor to bind to it. The gene producing this transcription factor may be newly evolved via duplication and subsequent change, in which case it simply adds to the existing pathway, or it may have already been part of a pre-existing pathway, in

which case the two pathways become integrated. In either case, the developmental function of the pathway may be significantly altered.

A striking example of gene recruitment is provided by the expanding function of the gene *distalless* (*Dlx*), which is widely expressed at the distal end of appendages. Wilkins (2002, p. 86) illustrates a sequence of structures, key to advances in vertebrate anatomy, that can be attributed to a succession of recruitments of *distalless*: tripartite brain, paired sensory organs, pharyngeal skeleton, dorsal and anal fins with fin rays, true teeth, paired fins, and jaws. The origin of feathers and the evolution of the tetrapod limb may also be explained in terms of a succession of gene recruitments involving both transcription factors (specifically of the *Hox* genes) and signal transduction systems (Wilkins, 2002).

Much research remains to be pursued (especially on vascular plants), but these new techniques provide the potential to explain the specific genetic basis for many of the anatomical changes that have occurred within the history of multicellular organisms, in relationship to their adaptation to various ways of life.

Conclusions

Although the most striking discoveries so far revealed by the new techniques of molecular evolution and phylogenetics are associated with large scale patterns and processes such as the origin of life, the differing modes of transfer of genetic material, the assembly of complex proteins, and the origin and radiation of metazoans, they are equally applicable at the level of populations and species. With these tools, it is now possible to document the unbroken continuity of evolutionary change from prokaryotes through the great diversity of multicellular organisms. At the same time, these capabilities provide more focus and possible solutions to the continuing field and laboratory investigations of both living populations and fossil material that have formed the basis of our current understanding of evolutionary principles and processes. It is the responsibility of the current generation of evolutionary biologists to make the newly acquired information and techniques more widely recognized by our students through lectures, papers and new, more fully integrated text books. Only in that way can we fulfil the potential of the next stage in the evolution of the study of evolution.

Acknowledgments

I wish to thank Catherine Boisvert and Mary Ann Lacey for technical help with the manuscript and figures, Tom Bureau for reviewing the manuscript, and the Natural Sciences and Engineering Research Council of Canada for grant support.

References

Anderson, O.R. 1988. *Comparative Protozoology*. Springer-Verlag, Berlin.

Arendt, D., Technau, U. & Wittbrodt, J. 2001. Evolution of the bilaterian larval foregut. *Nature* **409**: 81–85.

Bushman, F. 2002. *Lateral DNA Transfer: Mechanisms and Consequences*. Cold Spring Harbor Laboratory Press, NY, USA.

Bartel, D.P. 1999. Re-creating an RNA replicase. In: *The RNA World* (R. F. Gesteland, T. R. Cech & J. F. Atkins, eds), 2nd edn, pp. 143–162. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Bell, G. & Mooers, A.O. 1997. Size and complexity among multicellular organisms. *Biol. J. Linn. Soc.* **60**: 345–363.

Bowring, S.A., Grotzinger, J.P., Isachsen, C.E., Knoll, A.H., Pelechaty, S.M. & Kolosov, P. 1993. Calibrating rates of Early Cambrian Evolution. *Science* **261**: 1293–1298.

Bushman, F. 2002. *Lateral DNA Transfer: Mechanisms and Consequences.* Cold Spring Harbor Laboratory Press, NY, USA.

Butterfield, N.J. 2000. *Bangiomorpha pubescens* n. gen., n. sp. Implications for the evolution of sex, multicellularity and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* **26**: 386–404.

Carroll, S.B., Grenier, J.K. & Weatherbee, S.D. 2001. *From DNA to Diversity*. Blackwell Science, Malden, MA.

Cleaves, H.J. & Miller, S.L. 2001. The nicotinamide biosynthetic pathway is a by-product of the RNA world. *J. Mol. Evol.* **52**: 73–77.

Crick, F.H.C. 1968. The origin of the genetic code. *J. Mol. Biol.* **38**: 367–379.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. Murray, London.

Davidson, E.H. 2001. *Genomic Regulatory Systems*. Academic Press, San Diego, CA.

Deamer, D.W. 1997. The first living systems. *Microbiol. Mol. Biol. Rev.* **61**: 239–261.

Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.

Doolittle, W.F. 1999. Phylogenetic classification and the universal tree. *Science* **284**: 2124–2128.

Doudna, J.A. & Cech, T.R. 2002. The chemical repertoire of natural ribozymes. *Nature* **418**: 222–228.

Downie, J.A. & Young, P.W. 2001. The ABC of symbiosis. *Nature* **412**: 597–598.

Espinosa, M., Cohen, S., Couturier, M., del Solar, G., Diaz-Oreja, R., Giraldo, R., Jánniere, L., Miller, C., Osborn, M. & Thomas, C.M. 2000. Plasmid replication and copy number control. In: *The Horizontal Gene Pool* (C. M. Thomas, ed.), pp. 1–47. Harwood Academic Publishers, Amsterdam.

Finnerty, J.R. & Martindale, M.Q. 1998. The evolution of the Hox cluster: insights from outgroups. *Current Opinion Gen. Dev.* **8**: 681–687.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Galant, R. & Carroll, S.B. 2002. Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**: 910–913.

Gehring, W.J. 1998. *Master Control Genes in Development and Evolution: the Homeobox Story*. Yale University Press, New Haven, CT.

Gerhart, J. & Kirschner, M. 1997. *Cells, Embryos, and Evolution*. Blackwell Science, Oxford.

Gesteland, R.F., Cech, T.R. & Atkins, J. (eds) 1999. *The DNA World*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Gilbert, W. 1986. Origin of life: the RNA world. *Nature* **319**: 618.

Gilbert, W. & de Souza, J. 1999. Introns and the RNA World. In: *The RNA World* (R. F. Gesteland, T. R. Cech & J. F. Atkins, eds), 2nd edn, pp. 221–231. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Graham, L.E., Cook, M.E. & Busse, J.S. 2000. The origin of plants: body plan changes contributing to a major evolutionary radiation. *PNAS* **97**: 4535–4540.

Grant, P.R. 1999 (second printing). *Ecology and Evolution of Darwin's Finches*. Princeton University Press, Princeton.

Grant, B.R. & Grant, P.R. 1989. *Evolutionary Dynamics of a Natural Population*. University of Chicago Press, Chicago, IL.

Graur, D. & Li, W.-H. 2000. *Fundamentals of Molecular Evolution*, 2nd edn. Sinauer Associates, Sunderland, MA.

Haldane. J. B. S. 1932. *The Causes of Evolution.* Longman, Green, London.

Hamburger, V. 1980. Embryology and the modern synthesis in evolutionary history. In: *The Evolutionary Synthesis: Perpectives on the Unification of Biology* (E. Mayr & W. B. Provine, eds), pp. 97–112. Harvard University Press, Cambridge, MA.

Itaya, M. 1995. An estimation of minimal genome size required for life. *FEBS Lett.* **362**: 257–260.

Jacob, F. 1977. Evolution and tinkering. *Science*. **196**: 1161–1166.

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

Joyce, G.F. 2002. The antiquity of RNA-based evolution. *Nature*. **418**: 214–221.

Joyce, G.F. & Orgel, L.E. 1999. APPENDIX 3: reactions catalyzed by RNA and DNA enzymes. In: *The RNA World* (R. F. Gesteland, T. R. Cech & J. F. Atkins, eds), 2nd edn, pp. 687–690. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Knoll, A.H. 1995. Proterozoic and Early Cambrian protists: evidence for accelerating evolutionary tempo. In: *Tempo and Mode in Evolution* (W. M. Fitch & F. I. Ayala, eds), pp. 63–83. National Academy Press, Washington, DC.

Knoll, A.H. & Carroll, S.B. 1999. Early animal evolution: emerging views from comparative biology and geology. *Science* **284**: 2129–2137.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. & Cech, T.R. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**: 147–157.

Lawrence, J.G. 2001. Catalyzing bacterial speciation: correlating lateral transfer with genetic headroom. *Syst. Biol.* **50**: 479–496.

Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Malden, MA.

Maynard Smith, J. & Szathmáry, E. 1995. *The Major Transitions in Evolution*. Oxford University Press, Oxford.

Mayr, E. 1942. *Systematics and the Origin of Species*. Columbia University Press, New York.

Mayr, E. & Provine, W.B. 1980. *The Evolutionary Synthesis*. Harvard University Press, Cambridge.

McGregor, A.P., Shaw, P.J., Hancock, J.M., Bopp, D., Hediger, M., Wratten, N.S. & Dover, G.A. 2001. Rapid restructuring of *bicoid*-dependent *hunchback* promoters within and between Diperan species: implications for molecular coevolution. *Evol. Dev.* **3**: 397–407.

Merlin, C., Mahillon, J., Nešvera, J. & Toussaint, A. 2000. Gene recruiters and transporters: the modular structure of bacterial mobile elements. In: *The Horizontal Gene Pool* (C. M. Thomas, ed.), pp. 301–361. Harwood Academic Publishers, Amsterdam.

Moreira, D., Le Guyader, H. & Phillippe, H. 2000. The origin of red algae and the evolution of chloroplasts. *Nature* **405**: 69–72.

Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K.A., McDonald, L., Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M., Steward, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A., Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D., Eisen, J.A., White, O., Salzberg, S.L., Smith, H.O., Venter, J.C. & Fraser, C.M. 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**: 323–329.

Newman, A. 2001. RNA enzymes for RNA splicing. *Nature* **413**: 695–696.

Ochman, H., Lawrence, J.G. & Groisman, E.A. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.

Peterson, K.J. & Eernisse, D.J. 2001. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. *Evol. Dev.* **3**: 170–205.

Ram, J.L., Swalia, B.J., Malaga-Trillo, E. & Meyer, A. 2001. HOX clusters and the evolution of morphology – genome duplications and accelerated evolution of Hox genes and cluster architecture in teleost fishes. *Integrative Comparative Biol.* **41**: 676–687.

Ronshaugen, M., McGinnis, N. & McGinnis, W. 2002. Hox protein mutation and macroevolution of the insect body plan. *Nature* **415**: 914–917.

Schopf, J.W. 1995. Disparate rates, different fates: Tempo and mode of evolution changed from the Precambrian to the Phanerozoic. In: *Tempo and Mode in Evolution* (W. M. Fitch & F. J. Ayala, eds), pp. 41–61. National Academy Press, Washington, DC.

Schopf, J.W. 1999. *Cradle of Life*. Princeton University Press, Princeton.

Segre, D., Ben-Eli, D., Deamer, D.W. & Lancet, D. 2001. The lipid world. *Orgins Life Evol. Biosphere* **31**: 119–145.

Selosse, M.-A., Albert, B. & Godelle, B. 2001. Reducing the genome size of organelles favours gene transfer to the nucleus. *Trends Eco. Evol.* **16**: 135–141.

Simpson, G.G. 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.

Sonea, S. & Mathieu, L.G. 2000. *Prokaryotology*. Les Presses de L'Université de Montréal, Montréal.

Staley, J.T. & Reysenbach, A.-L. (eds) 2002. *Biodiversity of Microbial Life*. Wiley-Liss, New York.

Summers, D.K. 1996. *The Biology of Plasmids*. Blackwell Science, Oxford.

Tagawa, K., Satoh, N. & Humphreys, T. 2001. Molecular studies of hemichordate development: a key to understanding the evolution of bilateral animals and chordates. *Evol. Dev.* **3**: 443–454.

Thomas, C.M. (ed.) 2000. *The Horizontal Gene Pool: Bacterial Plasmids and Gene Spread*. Harwood Academic Publishers, Amsterdam.

Valadkhan, S. & Manley, J.L. 2001. Splicing-related catalysis by protein-free snRNAs. *Nature* **413**: 701–707.

Wilde, S.A., Valley, J.W., Peck, W.H. & Graham, C.M. 2001. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gry ago. *Nature* **409**: 175–181.

Wilkins, A.S. 2002. *The Evolution of Developmental Pathways*. Sinauer Associates, Inc., Sunderland, MA.

Woese, C. 1967. *The Genetic Code. The Molecular Basis for Gene Expression*. Harper & Row, New York.

Woese, C. 2002. Perspective: microbiology in transition. In: *Biodiversity of Microbial Life* (J. T. Staley & A.-L. Reysenbach, eds), pp. xvii–xxxii. Wiley-Liss, New York.

Zaug, A.J. & Cech, T.R. 1986. The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science* **231**: 470–475.

Zubay, G. 2000. *Origin of Life on the Earth and in the Cosmos*, 2nd edn. Academic Press, San Diego, CA.

Received 12 December 2001; revised 25 March 2002; accepted 3 July 2002