to accurately reflect 'immunocompetence' since these cells are usually only affected in particular chronic infections. Moreover, measuring immune-system parameters after manipulating the host's load of co-evolved parasites (usually 'worms') is equally uninformative because the responses such parasites generate in their hosts are complex and poorly understood[8].

In short, as well as considering the nature of the biochemicals mediating the relationship between immune-system function and signal traits, behavioural ecologists need to reach a consensus about the definition of 'immunocompetence' and subsequently design accurate assays for this phenomenon. Only then can we hope to conduct sound empirical tests of the 'immunocompetence handicap' theory[1].

Michael T. Siva-Jothy

Dept of Animal and Plant Sciences, PO Box 601, University of Sheffield, Sheffield, UK S10 2UQ

References

1 Folstad, I. and Karter, A.J. (1992) *Am. Nat.* 139, 603–622
2 Owens, I.P.F. and Short, R.V. (1995) *Trends Ecol. Evol.* 10, 121–122
3 Møller, A.P. (1995) *Trends Ecol. Evol.* 10, 121
4 Owens, I.P.F. and Short, R.V. (1995) *Trends Ecol. Evol.* 10, 44–47
5 Roitt, I., Brostoff, J. and Male, D. (1993) *Immunology*, Mosby
6 Besedovsky, H., del Rey, A., Sorkin, E. and Dinarello, C.A. (1986) *Science* 233, 652
7 Dieken, E.S. and Miesfeld, R.L. (1992) *Mol. Cell. Biol.* 12, 589–597
8 Maizels, R.M., Bundy, D.A.P., Selkirk, M.E., Smith, D.F and Anderson, R.M. (1993) *Nature* 365, 797–805

The evolution of information storage and heredity

Eva Jablonka
Eörs Szathmáry

Many important transitions in evolution are associated with novel ways of storing and transmitting information. The storage of information in DNA sequence, and its transmission through DNA replication, is a fundamental hereditary system in all extant organisms, but it is not the only way of storing and transmitting information, and has itself replaced, and evolved from, other systems. A system that transmits information can have limited heredity or indefinite heredity. With limited heredity, the number of different possible types is commensurate with, or below, that of the individuals. With indefinite heredity, the number of possible types greatly exceeds the number of individuals in any realistic system. Recent findings suggest that the emergence and subsequent evolution of very different hereditary systems, from autocatalytic chemical cycles to natural language, accompanied the major evolutionary transitions in the history of life.

Eva Jablonka is at The Cohn Institute for the History and Philosophy of Science and Ideas, Tel-Aviv University, Tel-Aviv 69978, Israel; Eörs Szathmáry is at the Institute of Zoology, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, and the Ecological Modelling Group, Dept of Plant Taxonomy and Ecology, Eötvös University, Ludovika tér 2, H-1083 Budapest, Hungary.

Replicators, objects that pass on their structure more or less intact[1], qualify as units of evolution: they multiply (objects give rise to more objects), have heredity (type A objects produce more type A objects, type B objects produce more type B objects, and so on), and show variation (heredity is not exact). If there is hereditary difference in fecundity and/or survival of the different types, then evolution by natural selection can take place[2]. In such a population, evolution may go on indefinitely.

Yet evolution quickly comes to a halt, or ends in a series of recurrent states, if the number of types is limited. For the commonly known replicators, this is not the case: there are $4^{99} = 10^{60}$ different nucleotide sequences of length 99, which can code for a polypeptide with 33 amino acids[3]. To the contrary, imagine hexanucleotide replicators: the number of all sequences is a mere 4096, which is easily realizable even in a minute molecular system. The notion of limited versus indefinite (unlimited) hereditary replicators was introduced to emphasize the difference between these respective systems[3,4]. For the latter, the number of types is much higher than the number of objects (individuals) in any realistic system; for the former, the two numbers are commensurate, or the number of types may be much less than that of the individuals.

Important evolutionary transitions may have been associated with novel means of information storage and transmission. Catalytic cycles, ribonucleotide replication, translation into proteins, epigenetic inheritance and natural language are all means of information storage and transmission. Their emergence and mobilization during evolution has been an essential component of major evolutionary transitions. Although the systems for storing and transmitting information discussed here are very different from each other, they have all evolved from limited heredity towards indefinite heredity. The attainment of indefinite heredity within a system points to a major transition in the history of life (see Ref. 4).

Replication and information before templates

From the chemical point of view, the essence of replication is autocatalysis[5], that is, when a compound catalyses its own formation. Since all catalysts work in cycles, it is natural that autocatalysts undergo cyclical processes[6,7]. Some autocatalytic cycles are well known from biochemistry. For instance, the Calvin cycle or the reductive citric acid cycle both serve to fix carbon. By doing so, the internal compounds of these cycles produce more internal compounds: in the Calvin cycle, for example, three molecules of 3-phosphoglycerate are necessary to ensure the production, after completion of one turn, of a fourth molecule[6]. The snag with these systems is that the elementary steps of the cycles are catalysed by enzymes that are not autocatalytic themselves. This would seem to render such cycles uninteresting for evolution. Gánti has shown that this is not the case[6]: hereditary changes may result from the intracellular competition of two rival cycles consuming and producing the same materials. While the integration into the whole of metabolism would remain unimpaired, the cycle with the higher net growth rate would competitively displace the other. Once all intermediates of a cycle are lost, and they cannot be regenerated from salvage pathways, its absence is inherited, even if the necessary enzymes are still available.

Forty years ago, Ycas suggested that evolution of cycles could have been important in the origin of life[8]. There is a non-enzymatic system that is thought by many to have played a central role, namely, the formose reaction (the autocatalytic production of sugars from formaldehyde[9]). More recently, an archaic version of the reductive citric acid cycle, together with some other alternatives, was proposed by Wächtershäuser in his theory of surface

metabolism[10] (a hypothetical organization before cellularization).

Clearly, intermediates of alternative cycles seem to qualify as units of selection: they multiply and have heredity[11]. Their evolutionary role would be more important if it could be shown that they are also variable. For this, they should be changeable by 'mutation' in some broad sense. Although experimental evidence is still lacking, such hereditary variants of autocatalysts have been suggested[12,13]. Traces of the autonomous evolution of such cycles may have been preserved in the metabolic map of coenzyme biosynthesis[12].

The appearance of a novel hereditary cycle would be more like the formation of a new species by 'macromutation', rather than that of a new sequence by a point mutation[3]. Autocatalysts of such cycles, where replication proceeds piecemeal rather than modularly[13], are limited hereditary replicators devoid of microevolution (see glossary; Box 1). The information that they carry can be termed analog information[14]. Nevertheless, owing to their intrinsic exponential growth tendency, they obey darwinian selection laws[11].

Synthetic replicators with limited heredity

Another type of system with limited heredity is a short template system with modular replication. The first such replicator was constructed by von Kiedrowski[15] in 1986, and several alternatives have been elaborated since. These systems are either short oligonucleotide analogs, or otherwise still modularly replicatable small molecules with template (complementary) surfaces, for example, a molecule with shape ⌐| can catalyse the ligation of the building blocks ¬ and | to |¬. By varying the ⌐ parts of these molecules, the creation of variants is possible, the replication of which is not exact[16].

The peculiar nature of the growth (in the absence of enzymes) of these replicators in a medium is that it is slower than exponential (malthusian): it is subexponential or parabolic[15]. This has a profound consequence for the dynamics of selection, resulting in 'survival of everybody'[11,17]. It is plausible that such replicators existed at a very early stage of the evolution of life. However, further evolution required the appearance of novel, indefinite hereditary replicators with modular replication[13], having the potential to carry digital information[14] and obeying the darwinian dynamic.

From the RNA to ribonucleoprotein world: the genetic alphabet and the genetic code

Ever since 1967–1968, many researchers have considered that the modern biological world was preceded by a world based on RNAs serving as information carriers as well as enzymes (reviewed in Ref. 18). Although it is unlikely that RNA was the primary genetic material, sooner or later it could have appeared and taken over the control of metabolism by its dual functionality. This represents a major evolutionary transition.

An important question about such a world is the size of the genetic alphabet: why are there two pairs of nucleotides (A:U/T and G:C), and not less or more[19]? This is especially relevant since the demonstration that novel types of base pairs (with alternative H-bond configurations) can be designed and synthesized, and they are accepted by some polymerases[20]. Two hypotheses have been put forward as an explanation, both of them arguing that two base pairs were an optimal character state in the RNA world, which became frozen in our RNA–protein world.

According to one explanation, the secondary structure of molecules composed of one base pair is very sensitive to mutation – they collapse easily, which would be detrimental for an enzymatic function[21]. On the other hand, random sequences composed of three base pairs do not readily fold into stable secondary structures. Hence, the two base pairs are a compromise between stability against mutation and thermodynamic stability[22]. This does not explain, however, why the state could not be invaded, through evolution, by nonrandom sequences with three base pairs.

The other hypothesis demonstrates that (1) with the increasing size of the alphabet, point-mutation rates increase faster than exponentially, and (2) the catalytic efficiency of RNA enzymes (ribozymes) increases slower than exponentially with the size of the alphabet, hence (3) there is an evolutionary optimum with two base pairs[23,24]. This is valid under the assumptions that the RNA world was metabolically complex, that the growth rate of unicellular ribo-organisms was limited by the flux of a critical metabolic pathway, and that the fitness of such a cell was proportional to the product of overall metabolic efficiency and genomic-copying fidelity. Both ideas are experimentally testable and they are not mutually exclusive.

The genetic code allowed for the radical expansion of the so-called catalytic alphabet (the list of proteinogenic amino acids[23]), by linking this to the genetic alphabet. The core of the code is a list of assignments of base triplets to amino acids. Based on the theory discussed above, it can be stated that the improvement on catalytic efficiency was impossible by simply extending the genetic alphabet[23]. Evolution must have sought another route that led to the biosynthesis of enzymes using the genetic code. The snag is that evolution is short-sighted, and therefore the transition cannot be explained by arguing that the end result was advantageous.

Many ideas, therefore, refer to preadaptation as a possible solution: amino acids – first in uncoded, later in coded form – must have served some other, useful function in the cell. Such could have been the facilitation of RNA replication with a 3'-bound amino acid[25], or the synthesis of useful dipeptides by peptide-specific ribosomes[26]. The most radical such idea is the hypothesis of coding coenzyme handles[27], suggesting that amino acids were used as coenzymes of ribozymes, and were equipped with oligonucleotide handles (Fig. 1).

The emergence of epigenetic inheritance

When cells divide, they frequently transmit their phenotypic characteristics to daughter cells. A cellular phenotype is sometimes inherited through many cell divisions even when the stimulus that first induced it is no longer present. In other words, not only can cells with identical DNA sequence differ in phenotype, these phenotypes can also be inherited. The inheritance of functional or structural

Box 1. Glossary

Analog information: the diffuse information stored by the the molecule as a whole, the latter being the intermediate of an autocatalytic cycle.
Digital information: the information stored by templates made of letters of a genetic alphabet. It is alterable, digit by digit.
Modular replication: the copying of a message consisting of the letters of the genetic alphabet.
Catalytic alphabet: the list of molecular species serving as building blocks of enyzmes.
Coenzyme handles: currently, many coenzymes (e.g. ATP, NAD, FAD) are equipped with nucleotide parts, mostly inactive in catalysis *per se*. They act, and used to act, as handles by which enzymes can and could grab them. This is especially obvious with enzymes made of RNA (ribozymes).
Cellular memory: the maintenance of functional or structural cellular phenotypes in cell lineages.
Epigenetic inheritance system (EIS): a system that enables the transmission of the phenotypic state of a cell or an individual to the next generation.
Steady-state inheritance system: the inheritance of functional states through self-sustaining metabolic networks, maintained by positive feedback.
Structural inheritance: the inheritance of new cellular structures that are arranged and ordered under the influence of existing cell structures.
Chromatin-marking inheritance system: the system responsible for the inheritance of states of chromatin, for example, for the inheritance of methylation patterns, or patterns of DNA-bound protein complexes.
Ripping: a process whereby repeated DNA is either mutated, excised or modified.

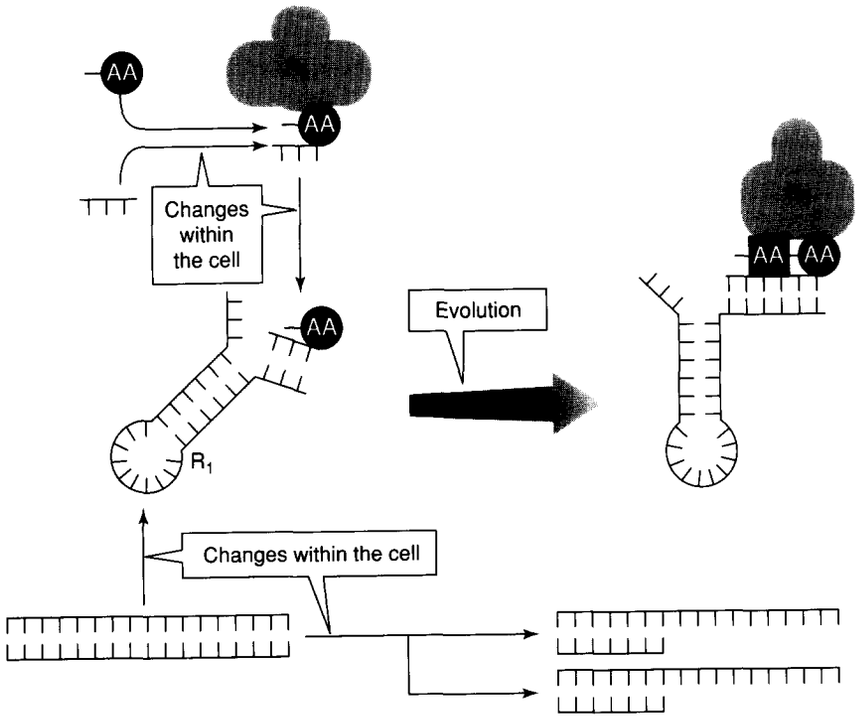

Fig. 1. The origin of the genetic code according to the coding coenzyme handle hypothesis[27]. In the context of an RNA world[18], different processes went on in a cell: nucleic acids were copied (genetic function), but they also acted as enzymes (metabolic function). Some ribozymes (R_2) were able to charge trinucleotides specifically (i.e. in a coded fashion) with an amino acid (AA), and the trinucleotide-AA compound acted as a coenyzme of ribozymes (R_1). Coenzymes were recognized and bound by their handles through base pairing. In evolution, R_1s were mostly replaced by messenger RNAs and corresponding protein enzymes, and R_2s were substituted by protein aminoacyl-tRNA synthetases. A new function (R_3) evolved to catalyse the coupling of amino acids to each other, currently performed by the ribosome. The original trinucleotide handles correspond to the anticodons of present-day tRNAs. *From Ref. 4, with permission.*

states in cell lineages is often referred to as cellular memory[28].

The transmission of cell phenotypes requires mechanisms that store and transmit information in ways additional to that based on DNA sequence. These mechanisms have been termed epigenetic inheritance systems (EISs), or dual inheritance systems[29]. Three such systems are recognized[30]: the steady-state inheritance, the structural inheritance, and the chromatin-marking systems (Box 2).

When and how in evolutionary history did EISs first emerge? The origin of steady-state systems was probably very simple: they depend on the regulatory properties of gene products, and these regulatory properties were probably associated with some proteins in the first cells. Variations in the quantity and the cellular localization of these proteins would affect the stability of the transmission of cellular phenotypes and their resistance to environmental perturbations, and therefore would be modulated by natural selection.

The evolutionary origin of structural inheritance is more obscure, but the universality of mechanisms that ensure the maintenance of cell form, and that ensure the transmission of cytoskeletal organization and of cortical organization, suggests that this EIS is also very ancient, accompanying cellular organization from the very start[34].

In bacteria, the main role of the methylation marking system is as part of the restriction-modification system. This is a protection mechanism against DNA parasites, in which a special enzyme methylates host DNA in a sequence-specific manner, so that the bacterial genome is distinguishable from foreign parasitic DNA, which is not methylated, and consequently is degraded by the host's restriction enzymes. Bestor suggested that the restriction-modification system was the evolutionary precursor of the methylation system adopted for the regulation and transmission of states of gene activity in eukaryotes[35]. The adoption of DNA methylation for the control of gene activity is in fact already seen in some bacteria: in uropathogenic *E. coli*, the expression of pilus proteins depends on the methylation patterns of adenines in two GATC sites in the regulatory region of the gene[36]. However, use of differential methylation in prokaryotes seems limited; in most cases, either all or none of the DNA sequences of a particular type in the genome are methylated. Locus-specific methylation seems to be the exception in prokaryotes. The type of selection pressure that may have led to the evolution of regulative hereditary states in unicellular organisms is described in Box 3.

The evolution of epigenetic inheritance

The evolution of epigenetic inheritance to an indefinite potential was important for the evolutionary transition to multicellularity in animals, plants and fungi. In multicellular organisms with many different cell types, the determined and differentiated state of various cell lineages has to be maintained. Maintenance of the differentiated state usually does not involve changes in DNA organization, but rather the operation of highly efficient EISs. High fidelity of transmission is required in such organisms because many are large and long-lived, and have extensive cell turnover. The chromatin-marking EIS seems more suitable than the steady-state EIS, in these cases, since it is more resistant to minor fluctuations in intracellular concentrations of regulators[34]. The maintenance of form during growth is poorly understood, but it seems that the structural properties of the extracellular matrix binding and coordinating a group of cells are very important[39]. Without efficient cellular memory, multicellular organization and complex ontogenies could not have evolved and could not be maintained[4,38,40].

During the evolution of multicellular organisms, the methylation marking system seems to have acquired a particularly important regulatory role. Some idea of the route that led to the role of methylation in gene regulation is indicated by the way the methylation system is preserved as a protection mechanism in higher eukaryotes.

Box 2. Epigenetic inheritance systems

The **steady-state inheritance system** is based on auto-regulation in which gene products act as positive regulators of gene expression. The simplest system is one in which a gene produces a product that ensures its own continued activity. The transmission of the functional state depends on the distribution of a sufficient quantity of regulatory gene products to daughter cells following cell division. Many such steady-state systems are known in prokaryotes and in eukaryotes[31]. In the **structural inheritance system**, a three-dimensional structure acts as a template for an identical structure in the daughter cell. This EIS was first described in the ciliated protozoa, where Sonneborn showed that experimental or accidental variations in cortical architecture can be inherited through mitosis and meiosis[32]. The **chromatin-marking system** is based on the inheritance of chromatin marks such as DNA methylation patterns (Fig. 2), or patterns of proteins bound to DNA. A particular DNA sequence can have several different heritable methylation patterns, or patterns of protein complexes, imposed on it[33].

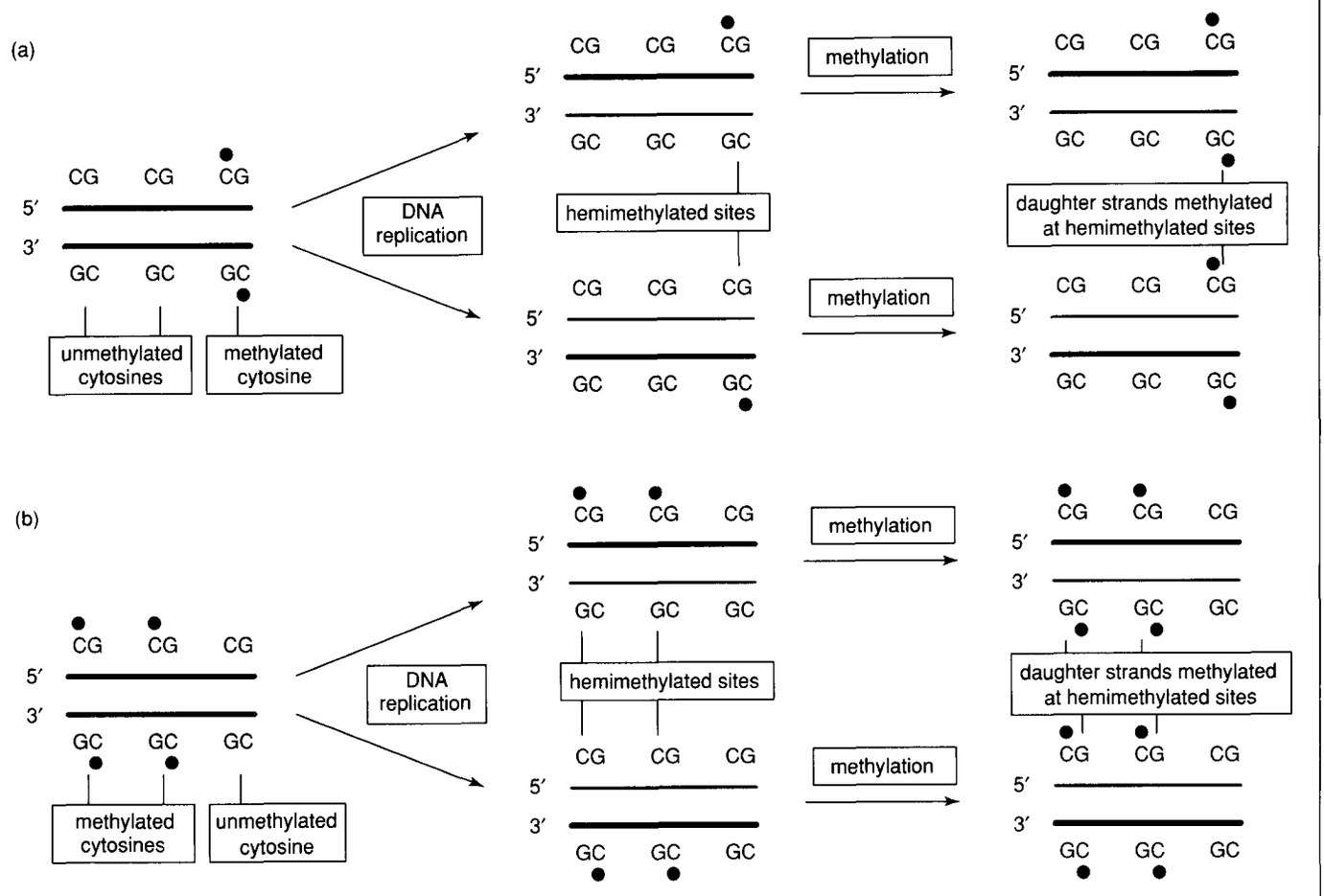

Fig. 2. The inheritance of methylation patterns. (a) and (b) are two identical DNA sequences with a different pattern of methylation at CpG sites (sites in DNA where cytosine, C, is followed by a guanine, G; p denotes the phosphate group so that C is at the 5′ position relative to G). Methylation sites and patterns are symmetrical on the two strands of the DNA duplex. After DNA replication, the unmethylated sites remain unmethylated, whereas the methylated sites are unmethylated on the new strand. This asymmetry is recognized by the enzyme methyltransferase (maintenance methylase), which methylates the CpG on the new strand opposite the methylated site on the old strand. Two identical methylation patterns are thus transmitted to daughter cells.

A process known as 'ripping' was first described in fungi[41]; during this process multiple copies of foreign DNA or duplicated copies of endogenous genes are extensively methylated and the methylated state is transmitted to daughter cells. In some fungi, such as *Neurospora crassa*, the methylated DNA is subsequently mutated or rearranged[41], whereas in other fungi, methylation merely inactivates DNA. Similar inactivation of foreign DNA by methylation – a kind of ripping – is also found in mammals[42] and plants[43].

It is not difficult to see how inactivation by DNA methylation, which was originally targeted at repeated foreign or endogenous sequences, could be modified to inactivate any endogenous gene. At first, endogenous gene inactivation was probably limited to loci with repeated sequences, or regions of chromosome with a repeated structure comparable to heterochromatin. Subsequently, establishing methylation patterns on the basis of local chromatin configurations determined by locus-specific, DNA-bound proteins provided a locus-specific mechanism of gene regulation. In animals, CpG sequences, and in plants, CpNpG sequences of endogenous genes, can be methylated; the extent and pattern of DNA methylation sites can be inherited and often affect the transcriptional activity of the gene[44]. Once this type of locus-specific heritable marking had evolved, the number of possible methylation phenotypes a cell can assume increased dramatically. The number of possible methylation patterns is determined by the number of loci with CpG sites and by the number of heritable functional states that every locus can assume. The potential of epigenetic inheritance became very powerful in higher eukaryotes.

The emergence of protolanguage: a cultural inheritance system with limited potential

The evolutionary transition to a culturally integrated group was made possible by a new method of storing and transmitting information – communication through natural language. Although animal groups (especially birds and mammals) can have limited group identity formed by socially transmitted, shared behavioural patterns, language has greatly enhanced group unity and cohesion, and allowed the formation of the complex, variable and long-term cultural units that characterize our species.

Many different scenarios have been suggested for the evolution of human language[45]. The large gap between human language and animal modes of representation and communication, combined with the coordinated and intricate structure of syntax, make the gradual evolution of a specific language-adaptation difficult to envisage and reconstruct. These problems have led some people to suggest that the evolution of language is a pleiotropic effect of an increase in general intelligence. However, few evolutionists believe that natural language sprang fully formed, like the goddess Athene, from the head of *Homo sapiens*. Most authors would agree that some simpler form of language, with a poorly developed grammar and lexicon, preceded the mature form of language that all extant healthy people in normal societies share. Recent discussions have focused on the initial conditions and selective pressures that started this evolutionary process, and on the number and nature of

Box 3. The evolution of regulative heritable states in unicellular organisms

The sophistication and elaboration of EISs as transmitters of regulative states seen in eukaryotes probably evolved as a response to recurring, fluctuating, environmental conditions. One aspect of the environment to which unicellular organisms are exposed is the periodical cycle (e.g. day and night, tides and seasons), which are often somewhat longer than the generation time of the organism. The two recognized ways of responding adaptively to changes in the environment are either through a change in DNA sequence (a classical mutation) or through a short-term, non-heritable change in gene expression. However, neither of these responses seems adequate if the cycle is longer than the generation time of the individual, but not as long as the time required for the occurrence and fixation of a rare advantageous mutation.

Organisms can adapt to such intermediate-term changes if they transmit phenotypes either through unconventional changes in DNA (as seen in many types of phase variations), or through EISs. A unicellular organism that can transmit its phenotype in such a fluctuating environment will have an advantage, because its progeny avoid some of the cost of being transiently in a non-adaptive state. Selection ensures that the transition rate from one epigenetic state to another reflects the environmental periodicity. Locus-specific transition rates (i.e. locus-specific memory spans) could thus evolve gradually[37]. Another preadaptation of EISs may have been associated with sex: yeast has different cell types, cellular memory and hierarchical gene control in association with mating types and the sex habit[38].

the possible intermediate stages before mature language emerged[45–47].

An example for a candidate protolanguage is provided by Bickerton[47]. He argues that human babies younger than two years, adults deprived of exposure to language during the critical developmental stages, apes instructed in language, and communities of adults deprived of a common language, all develop a relatively simple form of verbal communication. He suggests that this form of communication represents the evolutionarily ancient protolanguage. In this type of language, concepts of things and actions are labelled, yielding protonouns and protoverbs, and these are linked to simple propositions with no systematic word order and almost no grammatical structure. Grammatical items, such as *the*, *of*, *his*, *above*, *therefore*, *-ing*, are absent. As a result, many sentences have ambiguous meaning.

What were the initial conditions and the selection pressures that allowed the emergence of protolanguage? The assumption that a high level of social intelligence and social learning was one of the necessary preconditions for the evolution of language is widely shared (see Ref. 48 for a recent discussion). It is also accepted that the evolution of the vocal apparatus played an important role in the evolution of language[49], although its precise role is controversial. Darwin suggested that a rudimentary form of song was the first, essentially emotional, protolanguage[50]. Kimura proposed that tool-use led to gestural communication, and that vocal speech came later[45]. In the same vein, Corballis argued that articulate speech had to be built on the control of motor skills[51], because both are founded on generativity – a general ability to represent entities and actions by combining their component parts. Generativity first emerged with tool-making and was generalized to include vocal and visual generativity. Donald stressed the demand for increasingly better memory, and better voluntary access to memories, and suggests a mimetic stage involving all modalities before the emergence of spoken language[45]. All these factors are suggested as major contributors to the transition of the type of protolanguage seen when apes are taught language.

Although the amount of information that even a very simple protolanguage can transmit and store is enormous, relative to other vocal communication systems, complex sentences are impossible, and the type of transmissible information is limited relative to that allowed by a mature grammatical language.

The emergence of mature language with universal grammar

The transition from protolanguage to a rich, fully grammatical, mature language is no less obscure than that of protolanguage, and is the subject of much speculation and discussion. Following Chomsky's theory about the innate quality of syntactic structures[52], some linguists and evolutionists argue that the evolution of syntax is the most important process in the evolution of mature language and that syntactic structures must be a result of biological evolution and not of cultural evolution (reviewed in Ref. 53). Others believe that syntax is the emergent consequence of cultural evolution based on gradual lexical invention[45]. The lack of even remotely similar structures in other animals suggests to some that a macromutation (of unprecedented magnitude) has produced syntactic structures, while to others it suggests that it was the emergent pleiotropic threshold effect of the gradual evolution of general intelligence in a species that already possessed a protolanguage. Still others believe that gradual evolution of adaptive syntactic structures occurred[46].

A plausible way for the gradual evolution of syntax, from a mental structure that depended on extensive learning into an innate mental structure, could have been[46] via a process known as the Baldwin effect[54], or genetic assimilation[55]. This process of darwinian selection requires that the environment has a double role, being both the inducer of the adaptive phenotype and the selective agent. In the extreme case, natural selection leads to the transition from a stimulus-dependent physiological or behavioural response to a stimulus-independent response. A less-extreme case of genetic assimilation is a transition from a response that depends

Box 4. Learning and the evolution of behaviour[56]

Hinton and Nowlan constructed a genetic assimilation model showing how learning can facilitate adaptive evolution. The model is a neural network with 20 potential connections, each specified by a different genetic locus. Each of the 20 genes can be in one of three allelic forms: form 1 (connection present), form 0 (connection absent), or form ? (connection unspecified, and modifiable by learning). There is only one network that is 'correct', all other networks being equally 'incorrect' (thus we have a flat adaptive landscape with a needle-like peak).

The simulation starts with 1000 randomly generated organisms, so that each organism has 10 of its loci fixed in a 1 or 0 position (with equal probability), and the 10 other loci at the ? position, that is, left for learning to specify. Once an organism has acquired the correct setting in a ? locus, it is not altered subsequently. Each organism can perform 1000 trials during its lifetime. The organisms are sexual, having one offspring per mating, and to generate the next generation, 1000 matings are performed. The probability of being chosen as a parent increases if the organism has learnt, and is proportional to $1 + 19n/1000$, where n is the number of trials that remain after the organism has acquired the correct network. Thus, an organism that was born with all the correct fixed connections, and does not need to learn at all, has a fitness of 20, whereas an organism that has never learnt has a fitness of 1. The sexual reshuffling of genes allows the formation of new allelic combinations, and selection eliminates the fixed alleles at the incorrect positions and increases the number of the fixed alleles at the correct positions.

It is important to note that although there is only one correct network ('the needle in the haystack'), different genotypes can end up having the right network: different combinations of fixed alleles in the correct positions and different ? loci that assume the correct position following learning can produce the right network. As with other epistatic systems, there is selection for allelic combinations rather than for alleles. Selection leads to a higher frequency of 'learners' and to more-efficient learning, depending on fewer trials. The point is that, without learning, having 15 connections correct would be no better than having five correct; with learning, however, the former case decreases the time of learning, and, in turn, increases fitness. Ultimately, more and more correct couplings that had to be learnt at the beginning become genetically assimilated.

on a prolonged exposure to stimuli or to many learning trials, to a response that depends on a shorter exposure to the stimulus, or on fewer learning trials. Hinton and Nowlan have modelled this process and showed how learning can lead to the partial genetic assimilation of the response[56], so that learning, and the consequent adaptive behaviour, depend on fewer trials (Box 4).

Naturally, if one asserts that elements of syntax are innate, and that they have undergone evolution by natural selection, one would also expect that there must be genetic variation for grammatical skills. However, a recent analysis by Gopnik revealed[57] that a so-called feature-blind grammatical dysphasia in a family most likely results from a dominant gene. Affected members of this family are completely normal, but they cannot automatically generalize to obtain the rule that plurals are generated by *-s* and past tense by *-ed* (barring exceptions such as *went*); for them, the regular plural of a word is a separate word that they have to memorize independently. Although this is not a violation of universal grammar *sensu* Chomsky[52], it shows that (1) there is genetic variation for grammatic capacity, (2) there can be intermediates between protolanguage and proper languages, and (3) having an imperfect language is much better than having none at all[4].

The consequence of the acquisition of syntax is an enormous extension in the types of sentences that can be stored and communicated. Indeed, the number of possible, unambiguous sentences becomes infinite[52]. The rules of grammar allow the construction of hierarchical sentence-structures (which convey much complex information) and the formation of intricate narratives. With the evolution of mature language, cultural evolution, based on oral traditions of information-transmission, has become the major driving force in human evolution.

Acknowledgements

We thank Marion Lamb and Geva Rechav for helpful comments on an earlier draft of the manuscript, and Günter Wagner for inviting us to the symposium where we met. Support by Rüdiger Wehner to E.Sz. in the form of a guest professorship is gratefully acknowledged. This work was supported in part by the Hungarian National Scientific Research Fund (OTKA). The authors spent a year at the Collegium Budapest.

References

1 Dawkins, R. (1976) *The Selfish Gene*, Oxford University Press
2 Maynard Smith, J. (1987) in *The Latest on the Best: Essays on Evolution and Optimality* (Dupré, J., ed.), pp. 119–131, MIT Press
3 Szathmáry, E. and Maynard Smith, J. (1993) *Abstr. Bot.* 17, 197–206
4 Maynard Smith, J. and Szathmáry, E. (1995) *The Major Transitions in Evolution*, Freeman
5 Orgel, L.E. (1992) *Nature* 358, 203–209
6 Gánti, T. (1979) *A Theory of Biochemical Supersystems and its Application to Problems of Natural and Artificial Biogenesis*, Akadémiai Kiadó/University Park Press
7 Gánti, T. (1987) *The Principle of Life*, OMIKK
8 Ycas, M. (1955) *Proc. Natl Acad. Sci. USA* 41, 714–716
9 Cairns-Smith, A.G. and Walker, G.L. (1974) *BioSystems* 5, 173–186
10 Wächtershäuser, G. (1992) *Prog. Biophys. Mol. Biol.* 58, 85–201
11 Szathmáry, E. (1989) *Oxf. Surv. Evol. Biol.* 6, 169–205
12 King, G.A.M. (1980) *BioSystems* 13, 23–45
13 Wächtershäuser, G. (1988) *Microbiol. Rev.* 52, 452–484
14 Wächtershäuser, G. (1994) *Proc. Natl Acad. Sci. USA* 91, 4283–4287
15 von Kiedrowski, G. (1986) *Angew. Chem.* 98, 932–934
16 Rebek, J. (1994) *Sci. Am.* 271(1), 34–40
17 Szathmáry, E. (1991) *Trends Ecol. Evol.* 6, 366–370
18 Joyce, G.F. (1989) *Nature* 338, 217–224
19 Orgel, L.E. (1990) *Nature* 343, 18–20
20 Picirilli, J.A., Krauch, T., Moroney, S.E. and Benner, S.A. (1990) *Nature* 343, 33–37
21 Fontana, W., Griesmacher, T., Schnabl, W., Stadler, P.F. and Schuster, P. (1991) *Monatsh. Chem.* 122, 795–819
22 Schuster, P. (1993) *Orig. Life Evol. Biosph.* 23, 373–391
23 Szathmáry, E. (1991) *Proc. R. Soc. London Ser. B* 245, 91–99
24 Szathmáry, E. (1992) *Proc. Natl Acad. Sci. USA* 89, 2614–2618
25 Orgel, L.E. (1989) *J. Mol. Evol.* 29, 465–474
26 Weiner, A. and Maizels, N. (1987) *Cold Spring Harbor Symp. Quant. Biol.* 52, 743–749
27 Szathmáry, E. (1993) *Proc. Natl Acad. Sci. USA* 90, 9916–9920
28 Alberts, B. *et al.* (1989) *Molecular Biology of the Cell* (2nd edn), Garland
29 Maynard Smith, J. (1990) *J. Theor. Biol.* 143, 41–53
30 Jablonka, E., Lachmann, M. and Lamb, M.J. (1992) *J. Theor. Biol.* 158, 245–268
31 Blau, H.M. (1992) *Annu. Rev. Biochem.* 61, 1213–1230
32 Beisson, J. and Sonneborn, T.M. (1965) *Proc. Natl Acad. Sci. USA* 53, 275–282
33 Jablonka, E. and Lamb, M.J. (1989) *J. Theor. Biol.* 139, 69–83
34 Jablonka, E. and Lamb, M.J. (1995) *Epigenetic Inheritance and Evolution: the Lamarckian Dimension*, Oxford University Press
35 Bestor, T.H. (1990) *Philos. Trans. R. Soc. London Ser. B* 326, 179–187
36 Van der Woude, M.W., Braaten, B.A. and Low, L.A. (1992) *Mol. Microbiol.* 6, 2429–2435
37 Jablonka, E. (1994) *J. Theor. Biol.* 170, 301–309
38 Szathmáry, E. (1994) *J. Theor. Biol.* 169, 125–132
39 Ettinger, L. and Doljanski, F. (1992) *Biol. Rev.* 67, 459–489
40 Wolpert, L. (1990) *Biol. J. Linn. Soc.* 39, 109–124
41 Selker, E.U. (1990) *Annu. Rev. Genet.* 24, 579–613
42 Kricker, M.C., Drake, J.W. and Radman, M. (1992) *Proc. Natl Acad. Sci. USA* 89, 1075–1079
43 Jorgensen, R. (1993) *Philos. Trans. R. Soc. London Ser. B* 339, 173–181
44 Holliday, R. (1990) *Biol. Rev.* 65, 431–471
45 Donald, M. (1991) *Origins of the Modern Mind – Three Stages in the Evolution of Culture and Cognition*, Harvard University Press
46 Pinker, S. and Bloom, P. (1990) *Behav. Brain Sci.* 13, 707–784
47 Bickerton, D. (1990) *Language and Species*, University of Chicago Press
48 Dunbar, R.I.M. (1993) *Behav. Brain Sci.* 16, 681–735
49 Lieberman, P. (1984) *The Biology and Evolution of Language*, Harvard University Press
50 Darwin, C. (1871) *The Descent of Man, and Selection in Relation to Sex*, John Murray
51 Corballis, M.C. (1989) *Psychol. Rev.* 96, 492–505
52 Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press
53 Pinker, S. (1994) *The Language Instinct – The New Science of Langauge and Mind*, Penguin
54 Baldwin, J.M. (1896) *Am. Nat.* 30, 441–451
55 Waddington, C.H. (1942) *Nature* 150, 563–565
56 Hinton, G.E. and Nowlan, S.J. (1987) *Complex Syst.* 1, 495–502
57 Gopnik, M. (1990) *Nature* 344, 715