the origins of life

from the Birth of Life to the Origin of Language

J.Maynard Smith and E.Szathmary

January 8 - 12, 2006

Life is wet and messy. And what is it after all? Schrdinger asked the question famously in his 'What is Life' picking up a few typical characteristics of life which are shared by things that are obviously not life, as the example of fire, which has metabolism and an ability to multiply and grow. The authors take the point of view that life cannot be separated from evolution. Only that has life which has the potential to develope and evolve.

The evolution of life, which is nothing if not the one historical narrative on this planet, and one for which it is rather doubtful that it enjoys parallels anywhere else in the universe, is characterized by a few crucial turning points, the purpose of this book it is, to closer scrutinize and try to explain. The point of view of modern biology is the transmission of information, and this is also going to be the overarching perspective. As a presentation the book succeeds mostly on its delineation of complicated issues, dispelling many of the simplistic misconceptions many readers of popular biology may entertain; and not so much in its speculative musings on how things really came about.

One very remarkable thing about life is that it has, at least here on earth, a very precise characterization, namely the presence of DNA. There are, however, peculiar forms that lack DNA, (yet are intimately dependant upon those), known as viruses, and generally thought about halfway between life and inert matter. Those do not appear to be precessors to life, rather parasitic off-springs of it[1]. The principle of evolution, namely stable stages of developement is very general, general enough also to enclose not only the chemical world, but the subatomary. The building up of increasingly more elabourate atomic nuclei is an extreme example, and one which however is very limited in range. A more stringent definition, appropriate to life, involves notions like heredity and an increasing range of possibilities.

The Biochemistry of Life revolves around proteins, built up as chains of aminoacids, acquiring intricate, and mostly unpredictable (from first principles) 3-dimensional structures, critical to their various functions. Aminoacids occur spontaneously in, what has long been referred to as the soup of life, and supposedly present on the early earth. Construction of complex proteins would be too haphazard and slow a process would it not somehow be 'directed' and 'focused'. What does this is the emergence of the DNA molecule and the genetic code. One great mysetry to me is why there could not be alternate 'solutions' to this 'problem' each living alongside with DNA? One explanation could be that DNA is

[1] In the same way one may conceive future life-like machines, as not life on its own, but off-shoots from the great evolutionary river. Any discussion on life has to be set in a historical context, this being one of the major lessons of Darwin.

more or less canonical, in particular it would be a necessary component of life and whenever 'life' would appear on a different planet, DNA would have evolved. Another question is how likely its emergence would be. Maybe so unlikely that among the literally billions of possible sites, it has only emerged (so far) on earth. The extreme rarity of the event, could also explain the total predominance of DNA on earth. While rival solutions may in principle exist, each solution is so rare that so far only one has emerged. Finally there is the argument that DNA has somehow, by being first on the scene to argue for its intrinsic superiority is to reduce to the case of a canonical solution), exempted all alternatives, either by 'killing them off' or absorbing them. Against this scenario is the fact that there is a huge variation of life-forms on earth, able to co-exist with each other. Why should there not be differently based life-forms as well? An explanation would be of some 'bottleneck' effect. The basic information transmitting mechanism of life inherits some 'virtual' space that is quite small. Anyway, the authors do not at all formulate this question, let alone discuss it. Instead they do linger on why the genetic code looks as it is, and discusses its stability over time. This discussion appears to me a little bit puzzling, after all what is the basic mechanism of it? Each triplet of bases codes for a certain aminoacid, and how? Presumably because of the intricate 3-dimensional structures involved, allowing keys to fit into locks.

Another crucial evolutionary turning point is the development of the eukaryote cell from the more primitive prokaryote. The cell itself, is like a home, enclosing and concentrating chemicals to speed up chemical processes. Thus a more global and loose variation on the same principle on which the enzyme works[2]. The prokaryte cell has rigid walls, when it wants to appropriate some external object, it has to exude enzymes to 'loosen it up'. This is clearly a very wasteful process. The eukaryote cell has lost its rigid wal and acquired a very flexible one, which allows its form to change greatly, in particular allows it to engulf external objects. Clearly the internal ingestion is far more efficient, unused enzymes are not wasted, the whole process being enclosed. Other 'innovations' are the presence of subsidiary structures. The mitochondria being one example, engaged in the metabolic process of the cell by furnishing ATP molecules, the basic currency of energy in biochemical processes[3] It is now generally believed that the mitochondria historically used to be independant prokaryote cells, which somehow became trapped into a symbiotic relationship. Metaphors are often as misleading as they are suggestive. From the point of view of mitochondria, it simply means that those cells are just spatially confined to exist within their eukaryotic hosts, they have so to speak acquired a hull, but are still free to do what they are all about, divide and proliferate. Had the combination been detrimental, lineages had just died out.

[2] As I understand it, the enzyme, due to its intricate 3-dimensional structure, traps particular molecules on its surface, hence enhancing the rate of 'favourable' reactions.

[3] This incidentally points to another uniformity in life-chemistry, as other examples of which abounds, one may point to the various biochemical cycles. Such things are obviously selfperpetuating and thus deeply conservative and it is hard to see how alternatives would appear on the scene. What is not usually emphasized in biological texts is this hierachy of constraints, some are quite contingious, others seem eternal fixtures. In all science and rational thought, some things really have to be taken as inviolate, otherwise arguments would find no purchase and chaos would ensue. Another word for it is syntax.

The phenomenon points towards another important thing in biology, namely the notion of identity. Identity is an important philosophical concept of categorization, without which counting would be impossible. The cell is clearly a very basic biological identity, but what about multi-cellular organisms? In principle there is no distinction between a multi-cellular organism and a colony. A multi-cellular organism is not just an aggregate of cells, because if so we could in principle pick out any subset of cells, and call it an organism. Biology is about defining organisms, to make sense of what it means of different cells actually being connected to each other, the answer typically is phrased in terms of dynamic process to which each cell contributes by virtue of its specialized functions. It is always tempting to impose a design and purpose to those intricate processes (often likened to machines), the idea that they are just random goes against our instincts. In a very definite sense they are not random, as they are the outcome of history; yet it is instructive to think of them as blind. Now one can in principle in retrospect impose all kinds of underlying reasons to all kinds of perceived processes. A civilized society is obviously a good candidate for some kind of (meta) organism. Each individual perform many specialized functions for the good of the state, just like a cell. One may argue that the inherent complexity of the civilized state, in terms of thinking of the individuals as basic entities, is far less than that of each particular individual[4]. A far more succinct biological answer is that each cell in an organism is 'colored' in a distinct way unique (with obvious exceptions to be discussed below) to that organism. We are clearly talking about the genetic information stored in each cell. The reason for this uniformity is explained by history. Each organism is generated from only one cell. The case of natural cloning, i.e. the development of identical twins is interesting. Do they ultimately derive from just one single cell? If so, should they be thought of as one organism? From a dymanic point of view this is clearly absurd, once they have separated, physical connection is so, and the death of one twin does in no way compromise the integrity of the other. More interesting is the case when the twins are conjoined. Are they separate organisms? To some extent the test of physical separation could settle it. Yet there are many instances when this is not feasible and one still can talk about two different individuals, even when the brains are partially fused[5]. An insect colony shares many features with a multi-cellular organism. For one thing all of its members are genetically identical[6]. But unlike multi-cellular organisms (like the ants themselves) the 'embryological' development of the colony is exceedingly simple. There

[4] Even just in terms of numbers. A state counts its members in the millions, a human being its cells in the billions. Also a state, like a cell can be divided, and two seperate states can evolve. An advanced organism like a mammal, can not simply be cut in half, not even along its plane of symmetry, and expect each half to be viable.

[5] The typical conjoined twin is grossly asymmetrical, the other twin just being a more or less grotesque outgrowth of the other. Also one has to distinguish between each twin being feasible by itself, and the medical possibility (which is clearly time-dependant) of achieving this separation.

[6] This, as everything else in biology is only approximately true. The workers of an ant-colony, all being genetically identical to each other, are not so to the queen, only containg half of her material. In principle every cell of a multi-cellular organism can be turned into every other cell. A ant-worker cannot be turned into a queen. Also there are colonies with many genetically different queens, leading to interesting evolutionary strategies within the colonies lending themselves to mathematical game models,

are more sophisticated versions of colonies, namely the colonial medusa - *Nanomia cara*, in which each part, is a multi-cellular organism, each of them being far more specialized than the individual members of an insect colony. As always in biology it is hard to draw sharp boundaries, unless there are some very basic underlying principles.

Every organism starts out from one cell, and the subsequent division and organization, known as the embryological development, is a very complicated process, and dimly understood. It is not directed entirely by the genetic blue-print, but depends also very crucially on the environment. The often bizzare organisms that conjoined twins constitute, illustrate that the phenotype of an organism, is not determined by its genotype alone. On the cellular level it is based on supression of genetic material, through complicated pathways. This suppression clearly depends on what environment a cell will happen to find itself in. One should also be aware of that many structures unfold by no genetic instruction whatsoever, just following physical laws. It is assumed that the spotting and patterns of animals are the results of such purely physico-chemical processes[7].

Typical proliferation is by parthenogenisis. On the cellular level this simply means by division. And in the case of the the multi-cellular organism, one undifferentiated cell is the start of a new embryological process. Obviously in this way the progeny is genetically identical to the parent, and thus there is no developement just the dying out of lineages. Clearly this is absurd, and as everyone knows the solution to this paradox is simply that duplication is not perfect (as always creativity is a matter of the lucky error) nor is preservation eternal, mistakes or degenerations being known as mutations[8]. Mutations being random, most are detrimental to the functioning of the organism. Bad mutations are being weeded out by natural selection, preserving those who lead to higher proliferation. (The Darwinian almost tautoligy in its nutshell). The introduction of sexual replication has its obvious disadvantages but also subtle advantages. The recombination of genetic material means both a greater variety, not in the genepool as such, but in the possible combinations of genes. Also, when in pathogenisis, and mutations are weeded out, all the good ones in the lienage are as well. With sexual replication, bad genes go quicker, and the 'good' ones have a chance of escaping the same fate. Sexual replication highlights the classical problem, on what level does natural selection work, because on the level of individuality, the sexual element seems not to contribute anything at all. On the level of the single gene, the organism, or the species? One may make a very good case for the first[9], but the issue is basically one of philosophy, not biology, and careful consideration reveals that the issue is moot. Selection on the level of the gene, implies also higher-order selections at higher levels. Sexual replication introduces problems of its own. If we would randomly select parts of the genes from one parent each, the result would almost never be anything viable[10]. What happens in practice is that genes are strung up in chromosomes, each coming in pairs, associated with specific sites. Thus the new organism is sure to get

---

the provenence of Maynard Smith.

[7] as is the pattern of fingerprints? As I believe that identical twins have different fingerprints, a fact exploited by Mark Twain in one of his stories

[8] Clearly the most likely source of change is during the duplication process

[9] As exemplified by the 'selfish' gene

[10] The odds against it would be astronomically overwhelming.

packages of genes, each of them more or less sure to do the specific job required at a specific site. What is important for a functioning organism is to get viable combinations of genes, individual genes themselves make very little high-order sense[11] Genes are not shaken up, but structured packages of such. In practice things are not so simple. There are various phenomenon like cross-overs that violate the integrity of individual chromosomes, recombining them. Such effects should be classified as mutations, and are dealt with in a similar way. This also explains the confusion that often ensues as a result of sloppy writing. We all know that we share half the genes with either parents, and on the average one eights with each cousin. On the other hand we are told that we share 98 percent of all the genes with the Chimpanzee! Clearly we share half the chromosomes with either parents, and one eights on the average with our cousins etc. But the human chromosomes that are 'floating' around in the general 'gene-pool' are for each site-dependant very similar, in terms of what actual genes they contain. Thus you could simply have no simple chromosome in common with another human being, yet almost all the genes are the same[12]

> The elementary mathematical digressions involved with chromosome distributions are too sweet to pass up. In humans there are twenty-four sites, and hence twenty-four pairs of chromosomes. Given two parents with no chromosomes in common, they can than theoretically give rise to $(2^{24})^2 \sim 10^{15}$ genetically different individuals (only a tiny fraction of which are actually realised). In particular there are $2^{24} \sim 16$ million genetically different sperms a single man can produce, the total amount a man can produce may be estimated at about 3×10^{12} even that short of what would be needed, would every single sperm be used for fertilization. It is usually acknowledged that the actual off-springs of a couple are not-determined, there being so many possible combinations. Conversely it is claimed that on the other hand each human determine its parents. This is both true and false. False as the number of couples that could theoretically have parented a specific person is a far higher number than the possible progeny of a given couple, because chose a random complement to each of the choices of complementary halves of the given chromosomes, and as there are far more than fifty odd different chromosomes shared by humans, the number is staggering. True, because in practice, there will be just one couple that corresponds, almost none of the potential ever having existed (nor will ever exist). Furthermore call two people strangers if they share no chromosomes. It is possible, but very unlikely that you will be a stranger to a sibling[13]. On the other hand if your relatedness exceeds four steps, like second cousins, you tend to be strangers. Then of course even if strangers, your different chromosomes may be more or less related, but then the relationship become far more difficult to illustrate. One may ask questions as to how far back one need to go to find common ancestors, the more racially different, the higher the number, still it is reasonable to assume that any

[11] There are of course typical correspondences, like eye-color. Obviously higher order phenominon like intelligence cannot be reduced to the biochemical workings of a single gene. Nor for that matter to combinations thereof.

[12] Books are very different, yet they essentially contain the same letters.

[13] In fact as unlikely as you will be an identical twin, without being joined in uterus

pair of humans have common ancestors[14].

One may think that parthenogenisis only exists for primitive organisms. It is well-known that most plants will be able to grow from a shoot, and in fact many plants have a dual way of replication, sexual as well as asexual. It may then come as a surprise that there are lizards that are parthenogenic, although there are no mammals (or conifers), and most likely, unlike the case of reptiles, there is believed there is no way back for mammals. As mammals we are also told that sexual identity is genetically determined, but this is not necessarily true for reptiles. Whether an embryon is developed into a female or male depends on hormonic cycles. Those do not need to be genetically triggered, but can also be environmentally, say by the temperature. Yet another illustration that the relationship between genotype and phenotype is rather subtle.

The genetic material is a mess. The number of genes usually determine the complexity of an organism, but not necessarily. It has since its discovery always puzzled biologists that so much of the genetic code appears nonsense, coding for nothing, contributing nothing to the organism. Maybe we do not know the whole story, maybe all those nonsense segments have some hidden survival value? Still to certain types of computer-programmers, it should be natural. Why should nature edit its codes, if large part of a code is rendered inactive, why go to the trouble of eliminating it? And how would that be done physically? Genes do not ordinarily encode things that cut and paste genes themselves. And if such genes would develope, they could very quickly cause havoc.

The book concludes by a discussion on language and its possible genetic roots. Since the time of Chomsky it has been natural to assume that language is to a large extent innate, because after all we learn to talk as easily as we learn to walk[15]. All languages are also very similar, once one discards the accidental aspects of vocabulary and superficial grammatical structures. It is also notewowthy, that a child acquires the local language perfectly, regardless of its racial origin, testifying that human races are very similar. Yet, how did language arise? Any explanation is bound to be speculative, as their is a dearth of empirical evidence. Language does not fossilize, and there are no intermediate languages around to interpolate from the few proto-languages one may find among higher primates and dolphins. The advanced mind is supposed to be modular, i.e. very good at doing different parts. Socialization, manual problem-solving, knowledge about the natural word as to hunting and foraging. It is conjectured that a growing ability to use language allows an integration between the separated skills, enhancing them all, as well as allowing emergent features like abstraction and thus higher thought to evolve. It is, however, not very clear how to relate genetic developments to such a capability, obviously it is not the matter of some single 'language-gene' that has quickly and uniformly spread over the entire human population. Thus the discussion lacks the biological anchoring that has allowed the au-

---

[14] One assumes that all the surviving mitochondria stem from one 'Eve' some fifty-thousand odd years ago

[15] Adult speakers usually have great difficulty picking up foreign languages, and when not porperly instructed but nevertheless forced to use a strange tongue, they speak it very primitively, disregarding the finer points of grammar fashioning out a so called pidgin out of lingustic fragments; yet their children do manage unaided to impose a richer grammatical structure, creating what is lingusitically known as a creole, with essentially the same capabilities as a mature language

thors to write fascinatingly and with authority up to now. When it comes to philosophical issues, you are as likely as the other guy to come up with catching thoughts. The only attempt at some empirical comfirmation is a discussion of a strange language disorder in which the victims find it almost impossible to form plurals and other inflections of words. Each plural has to be learned as a separate item in a vocabulary, the idea of principle and transformation seems not to be present. Charting the deficiency in a family tree, one encounters the typical picture, that characterize say color-blindness or hemophilia. One noteworthy thing about language is that such dysfunctions seem to be very rare, unlike those that impose themselves in reading and writing. The reason being not too hard to understand, culture being such a recent phenomenon, that still does not entail more than say half of the total human population. Peoples ability to use language is remarkable. Although many native speakers may sin against the most superficial of grammatical conventions, they are unable to sin against the deeper syntax; and even people trying to learn a strange language do not make certain egregious mistakes. While peoples skills can differ greatly, their language skills seem not[16]. All mature languages are supposed to be equally capable, and the capabilities of languages do not develope, aside from the trivial extension of vocabulary, freezing so to speak the intellectual level of mankind. Dysfunction in mathematics and music seem far more common.

The evolutionary approach is a powerful one, and naively there seem to be no limits on its applicability. However, one should realize that there are constraints, and the mantra of reproductive advantage can be made to explain everything and hence nothing. There are orders of reproductive edge; some features are of first order so to speak, while others are of second and higher, contingent upon so many other things. Geological time is large, yet not infinite. Orders of speed do enter. This is in particular true for the very short intervals of human developement from humanoids. We are using units of tens of thousands of years, rather than millions[17]. This does not speak against language being a result of a genetic development, only the insurmountable difficulties in reliably charting its historical and biological narrative.

`January 14 - 16, 2006` **Ulf Persson:** *Prof.em, Chalmers U.of Tech., Göteborg Sweden* ulfp@chalmers.se

[16] Strictly speaking it is impossible to compare differences across modules of capabilities. We tend to take so much for granted, and thus simply discounting it, comparing mainly the differentials. Yet the average person is far more capable of judging and appreciating say Shakespeare, than they would be able to assess music or mathematics.

[17] Gould points out that there may be rapid evolvement on even fairly short time-intervals, put the nature of those are stochastic and d not point to any trends; just like the geoographical coastlines show finer and finer details, with the microscopic protuderances and inlets, have little significane to those that can be represented on maps.