

REMARKS ON THE FOUNDATIONS OF BIOLOGY

Seán Ó Nualláin

ABSTRACT: This paper attempts, inevitably briefly, a re-categorization and partial resolution of some foundational issues in biology. An initial ground-clearing exercise extends the notion of causality in biology from merely the efficient cause to include also final and formal causality.

The Human Genome Project (hgp) can be looked on as an attempt to ground explanation of the phenotype in terms of an efficient cause rooted in a gene. This notion gives rise to the first section discussing the computational metaphor and epigenesis and suggesting ways to extend this metaphor. The extended notion of causality alluded to above is necessary, but not sufficient, to demarcate a specific explanatory realm for the biological. While the universe can ultimately, perhaps, be explained by quantum fluctuations being computed through the laws of nature, the origin of life remains a mystery. The ground-clearing exercise refers to coincidences that motivate the cosmological anthropic principle, before raising an alert about the possibility of similar thermodynamic laws facilitating the emergence of life.

Life itself seems to involve symbolic operations that can be described by the grammatical rules within tightly defined limits of complexity. The nascent field of biosemiotics has extended this argument, often in a Peircean direction. Yet, even here, the task involved needs to be specified. Is the organism creating proteins to launch an immune counter-attack? Alternatively, is a pluripotent stem cell generating an entire organism? We consider what these separate tasks might look like computationally.

The paper ends with further delimitation of the specifically biological. At what point in the infinitesimal does life refuse to reveal its secrets? Conversely, at what specific levels in increasing size and complexity do boundary conditions emerge with hierarchy becoming immanent?

KEYWORDS: Anthropic principle, emergence, computation, hierarchy, causality, biosemiotics, thermodynamics, chaoplexity, Waddington, disease

1. INTRODUCTION

To say Biology is in “crisis” is, paradoxically, a complement. It is to state that the discipline has progressed to the point where contradictions between its putative founding principles, actual practices, and results are apparent enough to suggest salutary root-and-branch reform. This paper will issue such categorical prescription in the area of Genomics. Much of the rest of what follows is simply reflecting best practice in various subfields of biology.

Yet that is a non-trivial task, particularly as biology becomes invaded by researchers from other disciplines like statistics or, as in the case of the present writer, computer

science. A first stumbling-block is the nature of causal explanation in biology. We refugees from the informational sciences tend to think only in terms of Aristotle's "efficient cause". Wiser heads have pointed out that the final, teleological cause was necessary as an explanatory gambit for explaining the role of the heart in the circulation of the blood; similarly, the role of whole-properties associated with entities like the cell, echoing Aristotle, has given rise for about a half-century to the notion that organization might itself be a cause in the matter.

Keller (1995) famously explicated some of the metaphors underlying genomics, at a time when the HGP was quickening its pace. While the intellectual fireworks surrounding the argument about gene-as-Cartesian-homunculus were indeed spectacular, their glare perhaps blinded us to the apparently more prosaic speculations about the nature of the gene-as-computer-program. This is particularly the case as "computation" has undergone a deconstruction at least as significant as that of "cause", particularly in the work of Brian Smith (1996); see also Ó Nualláin (2007a). Echoing again the Aristotelian distinctions, what underlies genomics is the notion of efficient computability; the purely syntactic operations it assumes cannot yield phenotypes while maintaining their formal purity as context-independent operations. Ó Nualláin (2004, p. 174) summarises the situation as an essential tension between the formal requirements of syntax and the real-world exigencies of intentionality, reference in the real world, be that world a perceptual one, as cognitivism would be concerned with, or the biological phenotype. Yet, as we shall see below, that is just the beginning.

Elsewhere (Ó Nualláin, 2007d), this author specified the lacunae in current biology. The failure of the cognitive, neural and social sciences to explain cognition and consciousness; intelligent design/creationism versus "Darwinism"; the HGP and its problems, wherein biochemistry examples can be given to show the critical importance of metabolism (Strohman, 2003). On a more fundamental level, we have the observer paradox in biology, whereby reductionism results in losing life itself in a mass of physicochemical detail. The "Where is the program?" theme addresses itself to the massively complex interactions of hox (to be described below), epigenetic factors, types of rnas called sirna and microrna, both of which modulate gene expression, and so on. All these have consequences for the university-industrial complex, and how to rectify matters there in terms of faculty hires and other strategies.

With respect to health and ageing there is ongoing work by Bortz (2005), Veech and others (2003) on metabolism (particularly as it interacts with diabetes 2, a new epidemic based on insulin resistance) and health. In both his academic and more popular work, Bortz (forthcoming) argues that the attempt to find a "silver bullet" drug for diabetes 2 by targeting the biochemical networks involved in getting glucose into the cell is misguided; insulin itself is unnecessary if the organisms's metabolism is, aided by exercise, functioning as it should. Of course, there now is evidence that exercise prompts neurogenesis, thus alleviating depression.

In cancer research, there is an ongoing tension between oncogene (the notion that cancer is due to gene mutations) and aneuploid (the notion that cancer is due to gross

abnormalities at the chromosomal level) stories, perhaps soon to be resolved by the examination of carcinogenesis in the very aneuploid sufferers of “Gulf war disease”. The metaphysical ground specified as above, it should be stated that the purpose here is avoiding methodological blind alleys, and orienting toward appropriate sciences and technologies of medicine and agriculture. Too often has the discussion ended here with an invocation of chaoplexity concerns, and deep regret about the direction of biological research. The path here is to be different.

We are also concerned with bioethics, and the difficult issue of how precisely to think about our ecosystem in order to preserve it. According to Eliot Sober (Ó Nualláin, 2004, 143-149), there can be no bioethics simpliciter, no derivation of “ought” from “is”; rather, in the manner of Peter Singer (*ibid.*) with preference utilitarianism, one approaches the phenomena from traditional moral standpoints. Obviously, we need a bioethics and indeed a green politics. This writer believes that we will not be able to prescribe any lifestyle and political organisation in our lifetime from a science recognisably continuous with that of tradition. There are massive consequences of this, including a privileged position for freedom of conscience that precedes any scientific fact in the future. So this is also a counterargument to eliminative materialism; the political will forever precede the “scientific”

Pace titles like “From molecules to metaphor” (Feldman, 2006), let alone the strange political momenta seen therein in the abandonment in 2998 of the George Lakoff Rockridge institution, we are highly unlikely to have descriptions of mind in our lifetimes of sufficient cogency for political prescriptions. However, Green parties can exploit the formal limits that ecology sets on market capitalism to propose more just and sustainable societies. This will take a larger role in our concluding discussion.

It will be taken as given that the entire apparatus of that heterogeneous body of theory and disparate findings we label, variously, as “chaoplexity”, “dynamical systems”, “emergence”, and so on and discussed below applies to biological systems as it does to the realm dealt with in physics. Biology has many better claims for special treatment. First of all, there indeed are codes in biology, and the nascent science of biosemiotics is well-begotten. Secondly, while phase transitions of course exist in the non-biological world, it does seem the case that both they and phase boundaries are interrelated in biology with a fundamental notion of hierarchy. Thirdly, the nature of life does seem to invite a biological uncertainty principle, whereby the very existence of investigative devices at the nano level and smaller might require us eventually to specify at what scale the investigation is purely physics, rather than biology. This, of course, requires us to give an account, however tentative, of what life itself is.

This is an enormously ambitious project for a short paper. Luckily, Keller (*ibid.*), Strohmman (1993, 2003) and many others have covered the scholarly part of the project brilliantly. What this paper will focus on are four themes; first of all, computation and its reference to epigenetic explanation; secondly, evolution, metabolism and thermodynamics; thirdly, symbols, recursion and phenotypes; fourthly, boundary conditions, emergence, and the putative biological uncertainty principle. Then, having

depicted a specifically social realm discontinuous from the biological, we go on to examine consequences in terms of worldview and possible ethical echoes.

2. EPIGENETICS AND COMPUTATION

Deoxyribonucleic acid (DNA) is a polymer of four molecular compounds with subunits called nucleotides, each of which has a base, a sugar, and one to three phosphates. It is read 3 bases at a time, and an amino acid is specified through a complicated process involving RNA. Amino acids, in turn, are bound together by peptide bonds to form proteins. Gene expression is the transfer of biological information from the gene to the protein; an rna copy of the gene in mrna is made by a polymerase; the mrna binds to the ribosome; and there the proteins are constructed from amino acids. Three bases specify an amino acid. Each sequence of three nucleotides in dna is called a codon. These code for proteins, which are assembled on the ribosome. Regulatory genes control the operation of other genes.

A gene's most common allele, or variant, is called the "wild" allele. Transcription factors initiate the transcription of genes and are themselves regulatory proteins. Retroviruses perform reverse transcription into dna. This seems a priori a counterexample to Crick's deliberately portentous "central dogma" that dna generates rna generates proteins.

Similarly, the Beadle/Tatum dogma (1941) of one gene, one protein/enzyme was found utterly impracticable; there must be a generative approach like that for modern linguistics which we explore below. Molecular biology is ultimately a synthesis of informational (including Gamow-like cryptographic) and biochemical ("wet lab") approaches. It attempts to use the processes of physics and chemistry in genetics. It has been in many ways spectacularly successful; yet, even after the HGP, we have straightforward monogenetic correlates of only 2% of diseases. The analogy with Machine translation by computer, where brute-force methods are currently being touted by researchers, after 50 years of other methods, seems apt.

In Paris in the early 1960's, Jacob (et al, 1962), Monod, and their colleagues produced evidence for genes that work at a meta-level, by turning other genes on and off. In particular, genes can initiate the activity of sets of other such in particular metabolic environments. In the 1970's the work of Roberts, Sharp, and others indicates that much dna is "silent", and a consequence is that "alternative splicing", a phenomenon like prepositional phrase attachment in natural language processing, occurs with genes (Ó Nualláin, 2007b). At that time Watson (1977) hastened to insist that it is possible that a number of proteins could be generated from single "genes" (ibid.). This cautionary lesson was lost in the hype about the HGP.

Since the advent of multi-cellular organisms 600 million years ago, master genes have had to determine what type of expression would take place in different contexts. They in turn are governed by ?gSpindles?h in the chromatin. The chromatin is a set of specialised molecules that protect and control DNA. In embryonic cells, therefore, the

master regulatory genes are simultaneously repressed and readied for action through marks on the "spindles" in the chromatin. The developmental programs directing a cell to specialise into, for example, a neuron or liver cell, are initiated by these master regulatory genes. These genes produce proteins called transcription factors that bind to special sites on the dna and control the activity of lower-level target genes.

Of course, the issue of what regulates the master genes comes to mind. The answer seems to be related to the spools of protein called histones around which the DNA strand is looped 1.5 times. This hierarchical form is common to language and music. These spools can make the DNA accessible to transcription or not, as the occasion demands. For example, a complex known as the polycomb tags the spools at a site called K27, which signals another set of proteins to make the dna inaccessible. Conversely, spools tagged on K4 allow the cell to activate the local genes.

Histone proteins and the nucleosomes they form with dna are the fundamental building blocks of chromatin. We are obviously referring to nucleated organisms. Histone modifications may affect chromatin structure. Barbieri (1998, 2002) hypothesises that a histone "language" may be present that is read by proteins.

A recent article (2006) by Bernstein, Lander, et al, in "Cell" looked in detail at the chromatin at the points where the master regulatory genes were themselves in turn being regulated. They found that the chromatin contained both types of tags, as if the genes were being simultaneously readied for action and silenced. This makes sense, as most master regulatory genes will be unnecessary, but one will eventually be required. Likewise, mature cells have resolved into carrying just one or more of the K4 tags. So the "bivalent" state in embryonic cells is keeping them poised to go in a number of different directions. This has been established for mouse, human, and dog cells. The specifics of control seem to involve a network of oct4, sox2, and nanog, known to be associated with the embryonic state.

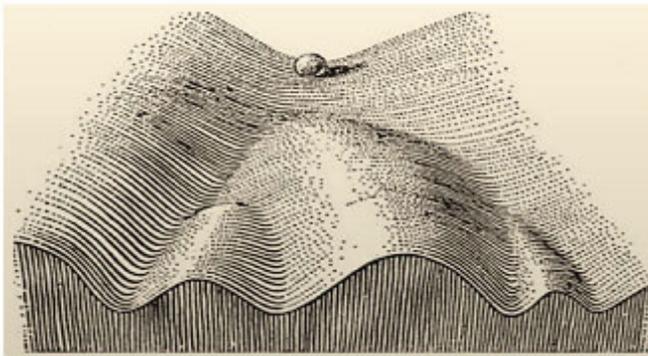
The chromatin may give us our first non-human code not specifically related to the genetic code, which links dna and amino acids. We will see several more such below.

Epigenetics is simultaneously in danger of becoming a catch-all term, and finding itself chained in perpetuo to single lab observations like methylated cytosine DNA markings or acetylated histone markings. For once, Scylla and Charybdis truly beckon. The sea-monster has been thrown up with the idea that there exists another logical level of genetic explanation, that of the epigenome, and that the epigenome manifests itself as the genome. So the HGP – or, more correctly, the HEP - should be done again, but with a broader jurisdiction and terms of reference. The contrasting danger is that of identifying epigenetics solely with one observed phenomenon, be that related to stress reaction or obesity.

Hidden in the discussion of epigenetics is the issue of where is the program for the phenotype. The answer from the boosters of the HGP was the disingenuous "in the genes". The reply from researchers like Atlan, Koppel, and Nijhout (Keller, op. cit., 28-29) was vociferous; the DNA may be merely a data structure within the overall computational architecture of the cell. In fact, inheritance may best be regarded as an

epigenetic phenomenon in their scenario. It behooves us to unpick this latter lock first.

Conrad Waddington (1966) was concerned, at least at one stage in his varied career, with outlining a scenario in which effectively inheritance of acquired characteristics could be squared with adherence to the “Central dogma” of molecular biology that dna makes rna makes proteins make the phenotype, rather like the more famous Baldwin effect. His specific mechanisms of epigenetic assimilation and canalization of development are of less interest to us here than the notion of the epigenetic landscape itself.



The above is the famous depiction of the “epigenetic landscape” and is currently being more narrowly interpreted in terms of cell fate with respect to organs, where the ball “fell” into one or other chreod, or groove. In fact, Waddington’s concept affords a perspective in which organism and environment can coherently be considered as a single unit over time (Maynard Smith, 1958, 1993). This conceptual breakthrough admits of the possibility of a converse to “epigenetic assimilation”, the process by which useful characteristics are incorporated in to the genome; that is, “environmental assimilation”, a process whereby the genome allows mechanisms to be inherited through the environment. The grooves, and indeed the whole landscape, can themselves change shape. It is likely indeed, given his later posthumous book on chaoplexity, that Waddington (1977) would be sympathetic to an idea that saw discontinuities arise in evolution as a result, for examples, of catastrophes as one fold in the landscape touched another. (Finally, this author’s Berkeley colleague Richard Strohman has written extensively on process structuralists like Brian Goodwin and I do not wish to pre-empt Strohman’s publishing this work by outlining his as yet unpublished arguments).

An example of “environmental assimilation” might perhaps be the Galapagos finch species *scandens* and *fortis* (Grant et al, 2006). The single species demarker is the song; yet this is learned from the father, and has no genetic component. Evolutionary biologists will recognize the sympatric/allopatric sequence. Similarly, monarch butterflies manage a multi-generational migration which seems to depend on the disposition of milkweed over an an enormous geographical space over eons. If the environment is stable, it may be more computationally efficient to allow it rather than the genome to hold

the necessary information. Indeed, we might find the organism and its environment perpetually together at the edge of chaos (Kauffman, 2000), as we discuss in the next section.

What of computation? Here we might learn much from the programming language Lisp, based as it is on lambda calculus with all the connotations that has for the essence of computation itself; lambda calculus is a formally equivalent alternative to the turing machine. The general format of Lisp is (function arguments); thus (+ 3 4) will return 7. (Please note that the parenthesising form is ubiquitous in Lisp, leading to the derogatory nickname “Lots Of Irritating Silly Parentheses”) Each form gets handed over to an interpreter called “Eval” which evaluates it. (The addition of a compiler would strain the analogy beyond breaking-point). We may, however, wish to prevent this evaluation and can do so by prefacing the structure with a quote; ‘(+ 3 4) will simply return (+ 3 4). This is useful if we are dealing with text. We can write procedures if, for example, we wish repeatedly to add 3 and 4 by using defun;

```
(Defun add34 ()
(+ 3 4))
```

We can invoke it as follows; (add34) which will simply return 7. And so we have a correlate to the mapping from sequences of nucleotides to specific amino acids. Let’s call this level 1 computation.

In common with most programming languages, Lisp allows creation of more complex functions. These may be triggered by preconditions which we symbolise by “cond”; in general, the syntax is

```
(Cond (critical condition) (Action if critical condition is met)
(T (Action if critical condition is not met).
```

Using this schema, we can engage with the Jacob/Monod apparatus of operators and operons. So we wish beta-galactoidase to be made if lactose is present; otherwise, we wish nothing to be done (with a report ‘null)

```
(defun lactose ()
(Cond ((sugar (setq beta-galactoidase ‘t)
(t ‘null))))))
```

At any point, then, we can all the function lactose to probe the environment to see if a particular type of transcription should start. Let us call this level two.

So far, then, we have automatic generation of amino acids and conditional initiation of transcription. The many brilliant minds behind Lisp, however, did not leave matters there. While functions can easily be written, it may be the case that we want them to be applied to one static piece of text and a variable to be instantiated with a value that depends on the context. A technical apparatus called defmacro allows this; briefly, the symbol ` requires that the text be left inviolate while ~ (or in “Common” Lisp dialect,) requires instantiation. Now we have ways of nuancing evaluation whereby the same list can be data, a function, and partially or wholly evaluated as we wish. The point is that if we are to make use of the computational analogy in DNA transcription, it helps if it is

a more sophisticated one that allows the same stretch of dna to be a transcription unit, ignored, or part of an encompassing task.

It is fair to say that we can encompass even the more radical suggestions of Nijhout with a conception of the genome as now program, now data structure, now a mixture of both; now catering only to context-dependent concerns, and now working at a level of great abstraction. The ultimate controller is not to be found in any Cartesian homunculus, Maxwellian demon, or Schroedingerian deity, to cite three of Keller's (ibid) metaphors; rather, it is ordering principles immanent in the relation of organism, species and population in situ over time.

This writer predicts that it will take many of the greatest minds of the twenty-first century much of its span to untangle the chain of control involving the polycomb, operators, promoters, micornas, sirnas, and so on with respect to the environment. (In the meantime, the straw man of Darwin, innocent of genetics, will confront that of YHWH, innocent of the whole of modern science, and great will be the din.) Issues of control of gene expression must be confronted in an evolutionary context; it may be the case, for example, that sirnas prevent too much change to an evolutionarily stable phenotype by one-off expression anomalies.

So far, we have a computational metaphor for automatic and conditional transcription. The latter can handle immune reaction, changes due to metabolic factors, and so on. It is obviously of much greater interest to try develop an analogy for generation of an entire organism. We know (Ó Nualláin, 2007a) that Hox genes work at two levels, at least. At the first level, genetic switches belonging to Hox itself specifies expression in different longitudinal segments of the body; at another, recognition by hox proteins varies gene expression. These are discussed in greater detail in section 5. We need to be able to handle factors at a greater level of abstraction using the technical apparatus supplied by macros. In particular, we should have a generic macro that can specify how a generic body part is made, and have this be instantiated into legs, arms, and so on which then have functions to generate them. Lisp allows all of this;

```
(defmacro make-body-part (part)
  `(defun,part,specification-of-part))
```

So we can then set up the details for variable called "arm" which comprise the information that Hox uses and then (make-body-part arm) will require that a function be generated called "arm" that, in self-referential fashion, works on its own definition of itself qua data structure..

I wish to emphasise that all of the above is metaphorical; the cell is not a lisp interpreter, and the functions, macros, etc do not provide anything like the requisite complexity needed. However, Lisp is a much better computational tool for thought than the various concepts Keller (1995) so ably exposes as the core of late twentieth century thinking about the gene and then, equally ably, deconstructs; Maxwell's demon, Descartes's homunculus, and so on.

In this section, then, we looked at some basic mechanisms of gene expression, and attempted to consider the computational environment in which such expression occurs.

Waddington's notion of the "epigenetic landscape" was returned to the larger context in which it was devised; that of the co-evolution of organism and environment over time.

3. EVOLUTION, METABOLISM AND THERMODYNAMICS

Strohman (2003) stressed several issues about thermodynamics and biology. First of all, no gene expression can operate as a *deus ex machina*; it must conform to the laws of thermodynamics and kinetics. Indeed, the theory of evolution must need be incomplete without continual reference to such laws. Conversely, the organism can handle great genetic insult if the ancient biochemical pathways maintain their integrity. He lauds the emphasis by physician-researchers like Walter Bortz(2005) on metabolism in maintaining health.

Remarkably, the argument about metabolism can be extended to consideration of the origin of life itself. (This is distinct from the set of arguments about one versus two membranes *ab initio*; see Cavalier-Smith, 2006). It is perhaps clear that the chicken/egg quandary concerning the fact that replication requires a critical number of nucleotide replicants which cannot get together without replication already being in place is unresolvable from first principles. While Shapiro (2007) successfully compiles a list of arguments against dna-first and rna-first theories, Orgel's (2008) posthumous rebuttal argues that the perceived inadequacy of these theories is insufficient to motivate adoption of metabolism-first. In particular, he stresses the current paucity of non-enzymatic thermodynamically stable chemical pathways that are thus demonstrably abiotic; yet he leaves the door open that some such may be discovered.

In his Dublin exile (for immersion in what he called "this remote and beautiful island" he and his wife regularly, if quietly, thanked the Fuehrer) Schroedinger chanced on the notion of negentropy as a defining factor of life. While Keller (1995, 66-78) is rightly keen to invoke Maxwell's demon in her exegesis of Schroedinger, we now have a superior vocabulary of dissipative systems, and chaoplexity has gifted us, *inter alia*, with the realisation that even abiotic systems like hurricanes can exist, negentropically, far from thermodynamic equilibrium.

Several other ordering principles are carried over from the physical realm. It will never do simply to invoke the deity to explain life, or even gratuitously in a scientific context *à la* Schroedinger. However, it is as well to say that life could not have emerged without the masses of the electron and proton at about 1836 (Barrow, 2002, P. 166). Likewise, immanent ordering laws are present not just in Feigenbaum's constant in chaoplexity, wherein the ratio between orbits at that rate seems to herald the onset of behaviour driven by a strange attractor, but also in the simple fact that a log scale of mass plotted against size in animals gives a nearly straight line (Barrow, 2002, P. 47). Similarly, the fine structure constant might be tweaked slightly without many adverse result, but make it much bigger and there can be no atoms (op. cit. 141-2). Finally, reduce nuclear forces slightly and there can be no biochemistry (ibid.). The number of photons per proton, the ratio of dark versus luminous matter, the specifics of the expansion of the

universe all bear similar witness (op cit. 182)

The coincidences that proponents of the anthropic principle rely on continue further. Fred Hoyle, the major opponent of the theory he derisively labelled “big bang”, theorised and then established that carbon exists only because the production of oxygen is non-resonant (op. Cit 153-154). Likewise, Kauffman (2000, 157) cites theory from Horowitz that the biosphere is at an energy per unit volume that permits the maximum expansion of molecular diversity.

Shapiro (ibid) stresses the critical importance for metabolism-first models of the prior existence of the following components; a boundary (ie a primitive membrane), an energy source, a linking mechanism(so that energy can be used), and a chemical network. He adds that the network must grow and reproduce and will at some point welcome a replicator., thus allowing death by “wearing out” to enter nature. At this point, we can perhaps dare to call the result “life”, while remaining attentive to Orgel’s caveats.

We should perhaps investigate the use of concepts from complexity and dynamical systems theory in general. Obviously, many of these terms are an expression of skepticism about the current state of our knowledge, epitomised by the word “chaos” itself. In a dynamical system, a fixed rule describes the time dependence of a point in geometrical space. “Emergence” refers to complex pattern formation from simple rules. A complex system is one with emergent properties. Chaotic motion is bounded, sensitive to initial conditions, and has dense periodic orbits. We can talk of basins of attraction like those that the brain settles into;

Fixed point ; dampened pendulum

Limit cycle; heartbeat when resting

Strange attractor; has non-integer dimensions and its dynamics are chaotic

Kauffman (ibid.) argues that nature always drives the nexus of organism and environment over time to the “edge of chaos”, maximum creativity. His is one of the most serious recent attempts to create new foundations for biology continuous with physics and chemistry, yet honouring the incredible complexity of life

Kauffman states that life emerges from complex chemical reactions. He then introduces the notion of “ autonomous agents” that can replicate and do one work cycle, normally far from thermodynamic equilibrium. They will be crucial, and he regards them as a new ontology of processes and events. He argues that there is a current crisis of explanation; an explanatory circle involving work, constraints, measurements, energy records, processes, events and - particularly – organization. We have adequate accounts of matter, energy, entropy, and information. Kauffman. is strongest when he draws consequences from chaoplexity.

In similar vein to his invocation of Horowitz above, who argues that the biosphere is at an energy per unit volume that permits the maximum expansion of molecular diversity, he argues that networks in cells exist near a phase transition between order and chaos. So the number of attractor states in the cell is the root of the number of genes (about 158) and the cell can cycle around all these states in a couple of days. He

argues in this context that communities of molecular autonomous agents may evolve to apparently discrete phase transitions like going to the edge of chaos or a self-organized criticality. The ordered and chaotic regimes and phase transitions between them are characteristic of a class of non-linear parallel dynamical systems.

Ecology is a fortiori a happy hunting-ground for proponents of emergentism (Williams et al, 2000) and indeed power laws (Williams et al, 2001) wherein linearity is left behind as a template for species distribution. There are many competing paradigms in ecology, which will have the above apparatus of chaoplexity repeatedly used on it.. For example, metabolic theory relates metabolism to body mass and temperature and derives useful items like speciation. At a higher logical level, spatial macroecology looks at how patterns in the distribution and abundance of species vary across spatial scales. Thermodynamic considerations include the fact that things never self-organise in a closed system Entropy production is maximised in irreversible processes; from this fact, as some argue, the other phenomena like species area laws fall out naturally. For example, is the greenness of earth maximising dissipative heat flow?

In this section, then, we have been careful to consider emergent ordering principles throughout all of nature, particularly when these are of a non-linear nature.

4. SYMBOLS, RECURSION, AND PHENOTYPES

The new field of biosemiotics stresses that with the advent of life comes also the existence of codes that are by definition not informationally dependent on the appearance of their carrier. The strong biosemiotics position is that biology is above all about communication. Codes include the genetic code; codes for apoptosis, or programmed cell death; histone codes; and so on. Witzany (Ó Nualláin 2007c and d) specifies code-editing as the essence of life, and a process that links us humans to the biosphere as a whole.

Let us first look at some of the basic tenets of biosemiotics (see Barbieri 1998, 2002, 2007). (Bios=life; semion=sign). Sebeok pointed out that signs used by animals are processed in the same way as humans' signs; his term, zoosemiotics, was later extended to plants as well and thus "biosemiotics"; coined by Rothschild in 1962, became common currency. Much of the theoretical infrastructure comes from Peirce who asserted there is a trio of sign, object (meaning), and interpretant in any act of signification. Signs must signify something; conversely, meanings require signs for their completion. A semiotic system connects meanings and signs through a code; all three elements are necessary. This contradicts Saussure, for whom a semiotic system was "signifier and signified", with the values of each deriving from their patterned differences with other such (The author wishes to thank an anonymous reviewer for this formulation),

The biosemiotics credo a la Barbieri is that organic coding requires signs, meaning, and an adaptor. This is a variation on Peirce. The physicalist notion is that "biological information" is a metaphor. This is denied by most biosemioticians; they say, for example, that genes and proteins are artifacts, made by molecular machines, and this

artifactual property is the essence of life. Thus, it refutes this physicalist idea

Genes and proteins are molecular artifacts because they are created by molecular machines. Life is “artifact-making”. Artifacts require entities like sequences and codes to be characterised. Nevertheless, organic information and organic meaning are not metaphors, but as real as any “natural” process. We call their results “nominable” entities which require an ordered listing of their elements for identification. Shannon’s concept of information, based on Boltzmann’s equation about the relationship between entropy, microstate, and macrostate, does not specify sequence subunits. However, biological information does, and is thus a nominable entity. Organic meaning mediates between molecules

Any organic code links two independent worlds (e.g. Genes and proteins) by a third world (e.g. Rna). “Sign, meaning, and adaptor” is pertinent rather than “sign, meaning, and interpretant”. According to Barbieri, therein lies a crucial distinction between objective and subjective. Early in the history of the biosphere, chemical “bondmakers” got created; some of them acquired the ability to join nucleotides together with respect to a template, and are “copymakers”. As proteins require mrna, transfer rna (trna), and the ribosome, they are more complex than copymakers. There is no necessity in the relationships between dna and amino acids, or between proteins and their eventual destination in the cell; we therefore can speak of codes.

A code is a set of rules that establishes correspondence between elements of two independent worlds. Barbieri (2007) has repeatedly argued that trna, with two spatially and informationally distinct loci that represent codons and amino acids is the beginning of the emergence of an arbitrary code. Organic information, it bears repetition is, according to biosemioticians, objective and irreducible. There are a plethora of codes in nature including;

The genetic code; the mapping from dna nucleotide sequences to amino acids

Signal transduction codes in cells: Cells continually respond to their environment; yet, the hundreds of possible “first” messages are transformed into combinations of only 4 “second” messages

The spliceosome features recognition of either end of scores of introns for each “gene”. So we can talk about “splicing” codes

The cytoskeleton is anchored to the cellular structure in an arbitrary fashion

There also are sugar codes, apoptosis codes, and so on.

So we can produce some summary statements as follows: The cell is a semiotic system with genotype, phenotype, and ribotype. The basic processes in life are coding and copying. The basic processes in evolution are natural selection and natural convention. Semiosis is defined by “coding”, not interpretation. Signs and meanings are codemaker-dependent. They are nominable, that is they can be specified by naming their components in their natural order. Rna and proteins also are codemaker-dependent. The translation apparatus is a semiotic system.

In classical semiosis; A (the interpretant) interprets B (the object) as representing C (the “meaning”).

This is called the “interpreter” model. The “codemaker” model has A (the adaptor, like tRNA), operating on B (the “sign”, like the genes), to produce C (the “meaning”, like amino acids). (The scenario to be outlined below reserves meaning for function in the environment). Since codes are arbitrary, learning their application is context-dependent. Genes are made by copymakers, proteins by codemakers.

Several variations on what biosemioticians do exist. For Sebeok, life and sign science imply each other. Communication is exactly what distinguishes living from nonliving. An organism is a device which communicates its structure to its offspring. For Emmeche, it is a branch of general semiotics, and its place in nature has yet to be determined. For Hoffmeyer, unification of biology depends on emphasising the semiotics nature of life. Sharov, largely in agreement, contends that it should be viewed both as biology and semiotics, not as a branch of the latter. Pollack introduces how our understanding of the genetic code, which burgeoned between the 1950’s and 1970’s, has unified the notion of text and organism. Pattee added the notion that communication is the essential characteristic of life, and Hoffmeyer the notion that the organism itself is a message. Uexkuell, following Piaget, wrote about organisms as interpreters of their environment.

Pattee (2001) is one of the chief theorists in the area. He argues that “semantic closure” is the crux for autonomy of systems. Thus, they reproduce themselves in the future and define their identity through the process of self-reproduction. “Self” is a semiotic term, but is handled well in immunological theory.

He emphasises the epistemic cut, the separation of rate-independent symbols from the rate-dependent dynamics they control. The converse is measurement, the coding of dynamic processes into symbols.

For him, there is a general issue about bridging the gap between the observer and the observed, the controller and the controlled, the knower and the known, mind and brain; the epistemic cut.

In perhaps an excessive move (despite the fact of Pattee’s original formation in theoretical physics), he argues that the process of measurement in QM falls under the same rubric as biological processes. (In a forthcoming paper in *Biosemiotics*, I explain why I find the move excessive).

Pattee extends the epistemic cut to a variety of distinctions; observer and observed, knower and known, genotype and phenotype. So the answer to the question about the distinction between living and nonliving is complex. The “motion of inorganic corpuscles” is rejected; there is no merely physical delimiter. There is an epistemic cut, with both constraints and dynamics. In short, local and unique heteropolymer constraint determined the origin of life.

The perspective in this paper is that the biosemiotic position, as enunciated by Barbieri, has merit with respect to novelty in evolution and the overall perspective afforded. In particular, the nexus of histones, epigenesis, sirnas and metabolism described above can best be approached using suitably sophisticated linguistics. Let us try and grapple with this subject, using a generative approach.

We should distinguish between Grammar, the totality of a person's linguistic knowledge, and "grammar" (lower-case), which we'll restrict to syntax. A grammar in the generative tradition is a set of re-write rules that generates (ideally, all and only) the sentences of a language (l). The vocabulary $V = a,b,c,\dots$ is the tokens forming the elements of the l. A sentence S (e.g. bca) is a string of these tokens

L is all sentences each of which is a string over V. We also have non-terminals like NP and VP, etc, to complement the terminals a,b,c,...

The Chomsky hierarchy posits languages of different levels of formal complexity; we can formulate general grammatical rules at each level. The top level, level 0, languages modelled by non recursively-enumerable sets and the next level, level 1 recursively enumerable such, are outside our scope here (Ó Nualláin (2007b, 253-254). Type 2, context-sensitive grammars, have rules of the form S (Sentence, or protein sequence) is rewritten as X if A precedes X and B follows it

$$S \rightarrow X/AB$$

We can also have Hox genes generating contexts or "fields" in this manner. With respect to our language analogy, "Sonic hedgehog" type genes allow context-sensitive rules. Carroll (2005, 42) cites experimental work by Saunders showing that transplantation of a chunk of tissue from a posterior to an anterior part of developing chicken's wing-bud resulted in "fingers" being developed in the reverse order to what they would have been. Thus if "a" and "p" stand, respectively, for anterior and posterior contexts, we can write the following rules;

$$S \rightarrow a/_{3,4}$$

$$S \rightarrow p/_{4,3}$$

Hox genes, then, allow for context-sensitive operations. It cannot be proven that any randomly chosen level 2 language can be recognised in finite time; likewise, computational problems at this level cannot have posited of them a solution in finite time. These connections are reminiscent of the unexpected connections between fractals and chaos; a priori, there is no reason to anticipate this type of phenomenon. Daley et al (2003) provide further light on this intriguing area. Natural languages in general seem susceptible to formalisation by an indexed grammar, which caters to the context-sensitive examples found by various linguists in human language (Ó Nualláin, 2003, 101-128)

Phenomena like alternative splicing (Ast, 2005) indicate that there is some degree of ambiguity, and thus complexity in gene expression. One hypothesis is that proteins are semantic primitives, rather than "meaning". At the syntactic level, Bcl-x, which governs cell death, can be alternatively spliced into Bcl-x (S), a promoter of cell-death, and Bcl-x(L), a suppressor thereof (Ó Nualláin (2007b, P. 251). This is structural ambiguity; to be more specific, it is a problem of prepositional phrase attachment, which needs access to the semantic or a deeper level to be resolved. So "the man shot the girl with the gun" leaves issues about his culpability. The nature of this layer has been approached by Bentilola (2005) who stresses the fact that living organisms have immense, dynamic

memories; each act, reflects all previous acts.

Let us introduce type 3 grammars with an example; playing tennis and “plugging” one’s opponent’s backhand before playing a cross court to the forehand. We do not want to generate bf; that is over-generation, but all and only the winning rallies. We need at least two backhands, so bbf, bbbf, bbbbf, etc are winning rallies. To write the grammar, we need to introduce a non-terminal E (Upper-case for these). We can define $E \rightarrow bf$. We also write a recursive rule $E \rightarrow bE$. If we add $S \rightarrow bE$, where S is a full rally (or sentence) we now have the full grammar.

A critical non-human code is the starling’s song, analysed by Tim Gentner et al (2006), and found to conform to a recursive grammar. It took 15,000 trials with differential reinforcement to get this effect

Birds were asked to pick out songs with inserted rattling or warbling phrases. 9 of 11 succeeded 90% of the time. The same task was attempted, and failed by tamarin monkeys used by Mark Hauser. He argues that the starlings do not have semantics, even if the recursive phenomenon is correct

Type 4, regular grammars, look like keyword systems; this/that is/was alive/dead which can generate just the sentences “this is alive”, “that was dead” and so on. The HGP was implicitly based on type 4, or finite-state automaton, even at the HGP’s most enlightened moments. The rest of the time, the HGP behaved as if the genome was similar to the parody of language apparent in the pattern-matching programs of the 1960’s. Lisp programs were programmed in these systems with preprogrammed scripts like “I have problems with my x”, to which they would reply “Tell me more about x”. It is likely that parsing the genome is infinitely more complex than this. Ó Nualláin (2007b) explores this issue in detail

In gene expression then, as in language, it is suggested here that contexts, are idiosyncratic interactions between linguistic and operational knowledge, which require knowing the precise relationship between the words (the genes) and semantic formalism (the metabolic context) in order for correct processing to occur (in order to predict what proteins will be generated). It is likely indeed that all of this holds for gene expression (Ó Nualláin 2007b)

Specifically, context seems to deform the layers of language as it becomes restricted in much the same way that gravity deforms space-time as one approaches the surface of a planet. (ibid.). The HGP worked on the assumption that the context was always going to be sufficiently restricted for single words to work, and therein lies its failure. It is a valuable lexicographic tool, and therein lies its success. However, we also need syntax, semantics, discourse pragmatics, and enumeration of contexts if nl is anything to go by. The HGP may be looked at alternatively as having elicited context-independent semantic primitives, or discovering unambiguous collocations, or as doing a lexicon. Time will tell which is the most fruitful perspective.

We must try and understand the metabolic context for the particular cases that Veech et al (for example, in their 2001 paper) examine or the “cytoplasmic regulatory protein components” (Bentilola, 2005). Then comes issues of “discourse structure”; how

tasks like building and maintaining the integrity of the organism act as high-level goals affecting the minutiae of protein generation, which we outlined above

With respect to gene expression research then, exemplified by the exaggerated claims being made in HGP research, it looks like remaking all the same mistakes made in AI. Indeed, the search for an “epigenome” where all will be revealed corresponds precisely to the search for an interlingua in machine translation, which was an almost complete failure (Ó Nualláin, 2003). Likewise, the attempt to introduce Bayesian nets between genotype and phenotype resembles nothing so much as expert systems which have been revealed to be either context-dependent or else of very little practical use.

5. BOUNDARY CONDITIONS, EMERGENCE, AND A BIOLOGICAL UNCERTAINTY PRINCIPLE

Following Maynard smith, Barbieri (2007, P. 16) outlines a set of “major transitions”; genes, proteins, first cells, eukaryotes, embryos, mind and language. There is an emergentist ethos here that cannot be gainsaid. We can also discern a hierarchy of natural law regimes from the level of the biosphere to that of populations to single species to groups, individuals, organs and cells; and now plasmon rulers, optical tweezers and so on allow investigation at scales hitherto inconceivable. At the nanometer level, X-ray crystallography is appropriate; above 200 nm, light-ray microscopy. It is possible that characteristics of life itself will not emerge at these dimensions. As we progress up the hierarchy from the genes to the biosphere, there are boundary conditions at each level. For example, the genome can be considered as the set of boundary conditions for the making of proteins.

For the moment, let’s look at an intermediate dimension, that studied by evolutionary developmental biology (evo-devo). Just as CERN attempts to intuit aspects of cosmogenesis by studying subatomic interactions, so evo-devo attempts to study evolution by looking at laboratory incidents of gene expression. Hox genes, which we’ve alluded to several times, belong to the homeobox gene family, a gene sequence that determines how the body develops from the first stages of embryogenesis. After the anterior-posterior and dorsal-ventral axes of the body are established, hox genes pattern the body into distinct segments, and within-segment patterns of cell fate determination. Thus, hox genes determine where limbs and other body segments grow during development. The homeobox is a 180-basepair sequence of dna and the polypeptide domain it encodes is called the homeodoman, about 60 amino acids long. Homeobox genes encode transcription factors.

There are four families of hox genes for different body areas in mice and all other mammals as in fruit flies. The evolution of body form in both arthropods and vertebrates has been achieved by shifting hox genes up and down the main axis. “If the impossibility of formation of a complex organ through a series of small changes was ever to be proven my theory would have certainly collapsed. However I could not find such an organ...” (Darwin, 1964, page 189.) Along with alternative splicing (short et al, 2008), hox provides

a realm of plausibly non-small changes.

We have seen that Hox genes function with switches at two levels; one set belongs to the hox genes themselves, and specify their expression in the different segments of the animal, and the other involves recognition by hox proteins to vary gene expression. The 1980's and 1990's saw the identification of Hox genes in drosophila. For example, drosophila's front wings are large, flat, venated, and powered for flight. The hind wings are balloon-shaped, smaller, and used for balance. The difference is due to a hox gene called Ultrabithorax (Ubx). This is activated in the hind wing to ensure that venation does not occur. Ubx switches off genes that encourage venation. Tool kits like these can be used again and again.

In fact, much recent evolution can be thought of in terms of the shifting of Hox genes up and down the body axis. The bat's wing, horse's leg, whale's flipper and human hand are homologously formed by homologous genes along homologous developmental pathways. The pax-6 or eyeless gene regulates eye development from our camera-type eyes to the fly's compound eyes. When the gene is manipulated to turn it on in other parts of the fly, we get eye tissues in the wings, legs, and so on. This also applies if the mouse eyeless gene is turned on in the fly.

Returning to *Drosophila*, bithorax mutations produced flies with an extra pair of wings. Antennapdia mutation transformed antenna into legs. These confirmed that there might be genes that controlled large-scale patterns of whole-body architecture, including positional information on the body axis. A mutation in human hox d13 produces an extra finger. Homeosis describes modifications of the anatomical themes and variations in body form. These genes have remained relatively unchanged throughout evolutionary history, and the complex may have evolved from a single ancestral hox gene.

In flowers, we get the same theme of serially repeated structures arranged in a systematic order, with repetition in linear-concentric symmetry rather than linear-bilateral symmetry. The multifunctionality of tool kit genes provides evidence for the notion of evolutionary descent from a few common ancestors. So we find that the BCMP gene, expressed in different contexts, can govern development of ribs, outer ear, and thyroid cartilage. Mc1r and agouti can both cause melanism. Mc1r is known to cause melanism in a set of light and dark pocket mice in Arizona. 475 miles away in New Mexico, it is caused by an as yet unknown mechanism. Note what we have been saying; the biological correlate to "meaning" in language is not proteins generated, but function in the environment, and again we emphasise context.

While the concept of the "biological uncertainty principle" has recently been misused for "silent" DNA, it has a more venerable proponent;

"In every experiment on living organisms there must remain an uncertainty as regards the physical conditions to which they are subjected, and the idea suggests itself that the minimum freedom we must allow the organism will be just large enough to permit it, so as to say, to hide its secrets from us. On this view, the very existence of life must in biology be....(like the quantum action) taken as a basic fact that cannot be derived from ordinary mechanical physics" (Bohr, 1933)

It is possible that life needs an interlocked self-contained metabolic system protected by some membrane that also includes the possibility of self-replication and therefore some kind of, paradoxically vicarious, survival after death. The resulting complex has emergent properties that several of the tools available today are too tiny to detect; precisely Bohr's point.

Let us leave this section with a brief foray into the ID/Darwin minefield. Fodor (1975) successfully argued that all human knowledge must be innate on totally logical grounds; a new concept is by definition discontinuous from the old (Ó Nualláin, 2003 91-92). The Iders will continue to use this type of rhetorical device successfully, particularly if their opponents insist on referring to Darwin, whose notion of inheritance actually precluded evolution insofar as it was at all coherent. Indeed, they will probably win arguments specifically about speciation, let alone the major transitions referred to above. Evolution, with its multitude of mechanisms including endosymbiosis and genetic drift, must be rescued from Darwin.

6. MIND AND THE DISCONTINUITY OF THE SOCIAL AND BIOLOGICAL

Witzany (2006, 2007) has laudable goals in his *Weltanschauung*; they are to establish the social as continuous with the biological, and inculcate a reverence for life. The argument to be made in this short section is that this goal is misguided. In the first place, it is a category error; the social is a different category to the biological. Ó Nualláin (2003, 169) cites work by Dyer that testifies to the extremely difficult path from biological process to symbolic behaviour. Yet that is still within the realm of the biological; what is not is a notion of a norm.

This author's work (Ó Nualláin, forthcoming) attempts, *inter alia*, to provide a foundation in cognitive science for social science concepts. It begins with the concept of selfhood, which is posited in the writer's experimental work as being founded on data-compression in the brain. In particular, our experience of selfhood qua identification derives from exclusion of data regarded by the organism either as irrelevant or as dangerous to its integrity. Our experience of ourselves as agents is largely a fiction in that we narrate continually, ascribing to ourselves the origins of actions that occurred automatically. However, there is a core of agency within each of us, while manifest perhaps in a conscious "won't" rather than "will".

The social is a higher category to the cognitive, which refers to the processing of information in the individual brain, and to the non-cognitive biological, yet is rooted in both. Insofar as the individual can construe herself as being the object of a norm or value, she is immersed in the social. This can be mediated as intersubjective, and not necessarily conscious, or fully conscious. Social pressures can be exerted on the individual, with often devastating results. Some such are artifacts of colonisation; yet that is getting ahead of the story.

The expression of the social is primarily verbal, and that will be our first port of call. It can be argued that all human institutions are created by speech acts, whereby concepts

previously abstract acquire rights and duties, and power over the individual; deontic powers. The particular types of speech acts involved are called performative utterances, which take the form “X institution shall...”. And so it can be said, for example, that a certain limited company shall exist within a certain legislative framework for a certain amount of time, including for eternity.

Yet the words alone will not unravel the web in which critical theory finds that humans live. While it is perhaps excessive to state that behind the structure of modernity lies the mediaeval, there is undoubtedly ample room for projection, and indeed scapegoating. The contemporary critic might indeed go further and point to the access of torture and Christian crusade in the early years of the 21st century to suggest something altogether darker. Admixed with freedom of conscience are themes a great deal lower in the brain than the cortex.

Therefore, the social realm can be expressed as laws, norms and values; it is a mark of legislative failure that biological concerns intervene to destabilise a regime. It is a symptom of maladaptation, whether cause by himself or the society, that an individual should have to focus on the facts of his biological being, rather than objective conditions, in order to resolve a personal crisis.

Finally, one can consistently follow Barbieri (2007) who, perhaps by oversight, omitted consciousness from the major transitions. It can consistently be argued that consciousness existed eternally, and that what is happening in the case of our individual experience is that the intentional structure of the mind is “bathed” by consciousness. This allows us to create the largely fictitious selves that we have through narration to ourselves; yet such narrations are the personal correlate of those inevitable forces that we call “social”

7. THE FOUNDATIONS OF BIOLOGY AND CONSEQUENCES

Let us now recapitulate with some summary points. 21st century biology will possibly be as consequential to its century as physics was to its predecessor:

1. Darwin must be sacrificed for the sake of the stupendous theory of evolution which is emerging, which draws its evidence from the subatomic as from Hox genes.
2. Some kind of anthropic principle will always be invokable to explain the origin of life, of multicellularity and all the other major transitions as it is for apparent coincidences like the value of the fine structure constant
3. Recursive symbolic function is a fact of nature.
4. There does exist a biological uncertainty principle at a scale much larger than the physical one.
5. Some aspects of the emergence of novelty will remain forever mysterious.

6. A new, more elaborate version of computation must be used for cellular function.
7. In a post-dissipative systems world, it is helpful to define life also in terms of replication after death.

The interaction between metabolism and symbol-processing that is present at all levels of life that makes biology so difficult, has enormous consequences. Gene-expression is subservient to ancient metabolic pathways; GMOs may indeed be filtered out by pure thermodynamics (but not before they have done much damage). Agriculture undoubtedly needs to focus more on the integrity of the ecosystems that it undoubtedly disrupts. In medicine, while prevention, including exercise, is undoubtedly currently more efficient than cure, the advent of “prospective medicine”, which predicts illnesses 50 years away, suggests a different argument. While an emphasis on metabolism, with its integrity undamaged by insult to the phenotype is correct, a return to socialized medicine a la Europe will work even better. Corporate medicine will always find a way to exploit the individual.

So we end with the political infringing on the biological. In this context, it is worth saying that even a cursory look reveals a stupendous epic of evolution resulting in our current state. It is indeed just a story; yet so also the rest of our knowledge is subject to the restrictions of these evolved brains. It is as ridiculous to impose metaphysical censorship on ourselves as we, with awe, contemplate our origins as it is to accept as Gospel—well, the Bible. It is possible to develop a politics based on the integrity of ecosystems as long as it is realised that our capacity for symbol use, and our very selves are also part of nature. To assert a “green” politics is also to assert the finest heights of human culture, and its extraordinary perennial search for the absolute grounds of its own existence.

Seán Ó Nualláin
visiting scholar, molecular and cell biology
Berkeley
Nous Research
Dublin
Ireland

REFERENCES

- Ast, G. (2005) “The alternative Genome” *Scientific American*, April 2005, 58-65.
Barbieri, Marcello (1998). The Organic Codes. The basic mechanism of macroevolution. *Rivista di Biologia-Biology Forum*, 91, 481-514.
Barbieri, Marcello (2002). Has Biosemiotics come of age? *Semiotica*, 139, 1/4, 283-295.

- Barbieri, Marcello (2007). "The mechanisms of evolution" in Marcello Barbieri (ed.) *The codes of life* (Springer) Pp 15-35.
- Barrow, J. (2002) *The constants of Nature* NY: Pantheon.
- Beadle, G.W. and E. Tatum (1941) "Genetic control of biochemical reactions in *Neurospora*" *Proc. Natl. Acad. Sci. USA* 27, 499-506.
- Bentolila, S. "'Live memory' of the cell, the other hereditary memory of living systems" *Biosystems* 80 (2005), 251-261.
- B. Bernstein, T. Mikkelsen, X. Xie, M. Kamal, D. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells." *Cell*, Volume 125, Issue 2, Pages 315 - 326.
- Bohr (1933) "Light and life" *Nature* 131, P. 458.
- Bortz, W. (2005) "Biological basis of determinants of health" *American journal of public health*, Vol 95 No 3 389-392.
- Bortz, W. (forthcoming) *New Medicine* NY: OUP.
- Beadle, G.W. and E. Tatum (1941) "Genetic control of biochemical reactions in *Neurospora*" *Proc. Natl. Acad. Sci. USA* 27, 499-506.
- B. Bernstein, T. Mikkelsen, X. Xie, M. Kamal, D. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath (2006) "A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells".
Cell, Volume 125, Issue 2, Pages 315 - 326.
- Carroll, Sean (2005) *Endless forms most beautiful* NY: Norton.
- Cavalier-Smith, T. (2006). "Rooting the tree of life by transition analysis". *Biol. Direct* 1: 19.
- Darwin, C. (1964) *The Origin of Species: A Facsimile of the First Edition*, Harvard University Press, 1964.
- Daley, M., O. Ibarra, and L. Kari (2003) "Closure and decidability of some language classes with respect to ciliate bio-operations" *Theoretical computer Science*, Vol. 306, Issue 1-3, Pp. 19-38.
- Di Giulio, M. (2005) "The origin of the genetic code: theories and their relationship, a review" *Biosystems* 80 (2005), 175-184.
- Feldman, J (2006) *From Molecules to metaphor* MIT press.
- Gentner, Timothy, Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum, "Recursive syntactic pattern learning by songbirds", *Nature*, 27 April 2006.
- Grant, P. and B. Rosemary Grant (2006) "Evolution of Character Displacement in Darwin's Finches", *Science* 14 July 2006: Vol. 313. no. 5784, pp. 224 - 226.
- Fodor, J. (1975) *The language of Thought*. New York: Crowell.
- Jacob, F., R. Sussman, and J. Monod (1962) "Sur la nature du répresseur assurant l'immunité des bactéries lysogènes" *Comptes rendus des Academies des Sciences*, 254, 4214-16.
- Kauffman, S. (2000) *Investigations*. New York: OUP.
- Keller, E F (1995) *Refiguring Life*. NY: Columbia.
- Maynard Smith, J. (1958, 1993) *The Theory of Evolution*. London, Penguin Books.

- Ó Nualláin, Seán (2003) *The Search for mind*; third edition. Exeter: England.
- Ó Nualláin, Seán (2004) *Being Human: The Search for order*; second edition. Exeter: England.
- Ó Nualláin, Seán (2007a) “Code and context” in Marcello Barbieri (ed.) *The codes of life* (Springer) Pp 347-356.
- Ó Nualláin, Seán and R. Strohman(2007b) “Genome and natural language” in Witzany (ed.) *Biosemiotics in transdisciplinary contexts: Proceedings of Biosemiotics 2006*. Helsinki; Umweb Pp. 249-260.
- Ó Nualláin, Seán (2007c) Review: Günther Witzany (2006) *The logos of the bios 1*. *tripleC* 5(1): 1-3, 2007.
- Ó Nualláin, Seán (2007d) Review: Günther Witzany (2006) *The logos of the bios 2*. *tripleC* 5(3): 1-3, 2007.
- Ó Nualláin, Seán (forthcoming) “Subjects and Objects” is in press with *biosemiotics journal*, Volume 2.
- Orgel LE (2008). The Implausibility of Metabolic Cycles on the Prebiotic Earth. *PLoS Biol* 6(1): e18.
- Pattee, H. (2001). “The physics of symbols: bridging the epistemic cut” *Biosystems* Volume 60, Issues 1-3, May 2001, Pages 5-21.
- Shapiro, R. (2007). A Simpler Origin for Life. *Scientific American* February 12, 2007.
- Short, S. And L. Holland (2008). “The Evolution of Alternative Splicing in the Pax Family: The View from the Basal Chordate Amphioxus “ *J Mol Evol.* 2008 May 14.
- Smith, B. (1996). *The origin of objects* MIT Press.
- Strohman, R. (1993). “Ancient genomes wise Bodies, unhealthy people” *Perspectives in biology and Medicine* 37(1): 112-145.
- Strohman, R. (2003). “Thermodynamics-. old laws in medicine and complex disease” *Nature Biotech*, May 2003, Vol 21, Pp. 477-79.
- Veech, R.L., B. Chance, Y. Kashiwaya, H. Lardy, and G. Cahill (2001) “Ketone Bodies, potential therapeutic uses” *IUMB Life* 51:241-247.
- Waddington, C (1966) *Principles of development and differentiation* NY: Macmillan
- Waddington C(1977) *Tools for Thought: How to Understand and Apply the Latest Scientific Techniques of Problem Solving* London: Jonathan Cape.
- Watson, J (1977) Cold Spring Harbour annual report.
- Williams, R. J. and N. D. Martinez. 2000. Simple rules yield complex food webs. *Nature* 404:180-183.
- Williams, R. J., N. D. Martinez, E. L. Berlow, J. A. Dunne, and A-L Barabási. 2001. Two degrees of separation in complex food webs. *Santa Fe Institute Working Paper* 01-07-036.
- Witzany, G (2006) *The logos of the bios 1*. Helsinki: Umweb.
- Witzany, G (2007) *The logos of the bios 2*. Helsinki: Umweb.