

TOPICAL REVIEW

The evolutionary origin of form and function

Keith Baverstock and Mauno Rönkkö

Department of Environmental Science, POB 1627, Kuopio Campus, FI-70211 University of Eastern Finland, Kuopio, Finland

Abstract We regard the basic unit of the organism, the cell, as a complex dissipative natural process functioning under the second law of thermodynamics and the principle of least action. Organisms are conglomerates of information bearing cells that optimise the efficiency of energy (nutrient) extraction from its ecosystem. Dissipative processes, such as peptide folding and protein interaction, yield phenotypic information from which form and function emerge from cell to cell interactions within the organism. Organisms, in Darwin's 'proportional numbers', in turn interact to minimise the free energy of their ecosystems. Genetic variation plays no role in this holistic conceptualisation of the life process.

(Received 28 January 2014; accepted after revision 5 March 2014)

Corresponding author K. Baverstock: Department of Environmental Science, POB 1627, Kuopio Campus, FI-70211 University of Eastern Finland, Kuopio, Finland. Email: keith.baverstock@uef.fi

Introduction

Stimulated by the need to explain the phenomenon of genomic instability as it is induced by ionising radiation (Kadhim *et al.* 1992), which has resisted explanation in terms of genetic mechanisms, we have re-examined the foundations upon which biology is built (Baverstock & Rönkkö, 2008; Baverstock, 2013; Annila & Baverstock, 2014). In summary, we propose that the life process is based not on genetic variation, but on the second law of thermodynamics (hereinafter the second law) and the principle of least action, as proposed for thermodynamically open systems by De Maupertuis (Ville *et al.* 2008), which at the most fundamental level say the same thing. Together they constitute a supreme law of physics: biology based on this tenet is fully supported by reasoning and evidence (Annila & Baverstock, 2014). For example, we have been able to provide an explanation for the 'missing heritability' (Turkheimer, 2011) and a

rationalisation for the steady overall increase in complexity exhibited by 3.5 billion years of evolution.

In this revised model, genes do not play a prominent role: their sequences specify the sequences of amino acids in peptides, which fold into proteins, which in turn interact to give rise to the majority of phenotypic properties at the cellular level. These interactions upon which phenotype is contingent, are symmetry breaking and lead to cellular phenotype being an emergent property of the system (Anderson, 1972). Consequently, it is rarely possible to definitively relate phenotypic properties to specific genes. The exceptions occur when a single gene codes for a single peptide, which folds exclusively to a single protein and acts alone to provide a phenotypic property. Under these circumstances, there is an association between sequence and phenotype, but it is not causal in its nature.

Darwin insisted that evolution was a gradual process, but in his Recapitulation and Conclusions (Chapter 14) of

Keith Baverstock PhD, trained as a chemist, joined the Medical Research Council in 1971 with the dual remit of providing support to the MRC's Committee on Protection against Ionising Radiation and research. In 1991 he joined the Regional Office of the World Health Organisation in Rome, where he was instrumental in bringing to world attention the childhood thyroid cancer in Belarus resulting from the Chernobyl accident. Since retiring from WHO, he has been a professor at the University of Eastern Finland and is currently a Docent. His research interest is in theoretical biology. **Mauno Rönkkö** got his PhD in Computer Science from the Åbo Akademi University 2001. He is currently working at the University of Eastern Finland at the Department of Environmental Science. His research interest is in agent-based modelling and computational intelligence applied to biological and environmental ecosystems.



the *Origin* he reluctantly admits that the fossil record did not support this contention. Modern Darwinists regard the evolutionarily more recent life forms as arising through selection acting on genetic variation, with new variants deriving from mutation of existing genes – a slow and gradual process. However, Noble (2013) has systematically rejected the tenets of the Modern Synthesis. Eldridge and Gould (Gould, 2002) proposed that evolution was characterised by long, in evolutionary terms, periods of stasis punctuated by shorter periods of rapid development, the so-called theory of punctuated equilibrium. The reformulated biological foundation we have advanced (Annala & Baverstock, 2014) is more consistent with the latter than the former: we would equate the periods of stasis with replication of cellular phenotype true to form based on an evolutionarily conditioned attractor state (so called home attractor; Baverstock & Rönkkö, 2008). The periods of rapid change are associated with the less stable variant attractor state derived from the perturbation (usually by environmental stress) of the home attractor. Attractor states represent cellular phenotype (Baverstock & Rönkkö, 2008).

Theory

The importance of the ecosystem. In this reformulation form and function, extant and extinct, are the consequence of natural selection acting primarily upon the ability of organisms to extract energy (nutrient) from their environment, as pointed out in 1835, prior to the publication of *Origin*, by Edward Blyth (Blyth, 1835). Organisms are contingent on the ecosystem in which they thrive because this is the source of that energy. Darwin spoke in the *Origin* of the importance of the *proportional numbers* of organisms in an ecosystem in terms of that ecosystem's stability and gave explicit examples of how modifying those proportional numbers could cause dramatic changes in the ecosystem. For example in Chapter 3 of the *Origin* he describes heath land near Farnham in Surrey, populated with the occasional Scotch fir, yielding numerous such firs in a fenced off enclosure: browsing cattle were preventing the growth of the trees on the open heath land. From this profound observation we conclude that in terrestrial terms organisms are information-carrying components contributing to a steady state system driven by the availability of energy from the sun. Energy, in the form of nutrient, is consumed, thereby producing entropy, according to the second law in the most efficient way (least action) possible given the conditions. Under these circumstances, explicitly thermodynamically open systems, entropy is maximised in the form of organisation or complexity (Sharma & Annala, 2007) and not, as proposed by Boltzmann, disorder (Sharma & Annala 2007). In terms of the food chain, the entropy (bound

energy) of lower forms is available as free energy (nutrient) for higher forms.

Form and function in organisms. The energy and information transduction process described above was investigated in detail in a recent article (Ronkko, 2007). In that research, an entire ecosystem was artificially constructed based on a simple computational three-dimensional model. All elements of the ecosystem, including ground, water, clouds, plants, worms, and hunting beetles, were comprised of 'atoms' bearing information regarding their interactions (rules of engagement) with other 'atoms'. The interactions were either 'fixed' and gave rise to specific *forms* (the beetle, for example) or 'mobile', that is, motivated by pheromone based signalling from and to *functional* 'atoms' emitting or receiving the signal 'atoms': all 'living' entities needed energy in order to survive, and they made local decisions about reaching those energy sources based on the signalling. The ecosystem exhibited an ecological pyramid, where sunlight (in the presence of water) provided energy for the plants to grow (produce entropy or bound energy) through photosynthesis and the Calvin cycle, to produce energy rich nutrients such as carbohydrates. Grass acted as the energy source for the worms, which in turn acted as the energy source for the hunting beetles. Furthermore, all the interactions were physically restricted by a three-dimensional landscape. The modelling results established and validated the connection between form, function, and phenotypic behaviour. In particular, intelligent collective behaviour emerged from simple rules of engagement, indicating also the presence of implicit information transference. Analysis of the dynamics of the system in greater depth revealed the existence of conditioned attractor states; in the artificial ecosystem, ecologically stable dynamics were achievable only from specific initial conditions.

The above artificial life model illustrates how physical information distributed throughout a system comprised of structural and functional elements (atoms) leads to coherent emergent behaviour of the whole system. Our proposal is that cells comprising organisms, provided with phenotypic information derived from protein interactions, perform the role that 'atoms' play in the artificial life model. Organisms are attractor states based on cellular interactions in self-similarity to cellular phenotype envisaged as an attractor based on protein interactions, while also being the interacting information bearing components (attractor) for their ecosystem. We therefore postulate self-similar processes at three interacting levels, namely, proteins, cells and organisms that are analogous to the processes investigated and established earlier in the context of elementary organs, organisms, and ecosystems (Ronkko, 2007).

Discussion

We defend the argument that genetic variation is not the origin of the evolution of form and function on three levels as follows: the unrealistic nature of the prevailing genetic paradigm, the rational superiority of the thermodynamic tenet, and its explanatory power.

In 3.5 billion years of evolution of living systems, it was only some 600 million years ago that eukaryotic cells contrived to overcome the barrier to fully functional multicellularity. Opinion is divided on when single cell eukaryotes first appeared, but most think it was two or more billion years ago. However, the first primitive multicellular forms of bacteria, biofilms, were evident in the fossil record about 250 million years after the origin of life (Hall-Stoodley *et al.* 2004) and have complex dynamic structures comprising a number of differentiated cell types. It is therefore clear that pre-eukaryotic life had the two primary features that would form the basis of more complex multicellularity from an early stage. Furthermore, genome sequencing has revealed that although mammals, for example, deploy a few tens of thousands of discrete gene coding sequences, bacteria deploy several million (Yang *et al.* 2009).

We predicate the current proposal on a metabolism-first origin of life (Baverstock, 2013), in which proteins, free of DNA, were a form of proto-life. Life appeared when these proto-life forms recruited nucleic acids in the form of DNA to act as a template for replication and to code for essential peptides (Annala & Baverstock, 2014) through the process of reverse translation making it possible for true replication to occur. On this basis the first bacterial life was far from short of genetic variation and as coding sequences can be exchanged between cells (Shapiro, 2011), including between bacteria and eukaryotes, eukaryotes were, and are, not short of genetic variation. In other words mutation of existing coding sequences is unnecessary for evolution to have taken place – that is not to say that evolution has not taken advantage of mutational events, but that genetic variation is not rate limiting.

The thermodynamic tenet states that the essence of a living organism is that it is a natural process extracting energy from its environment in order to grow and replicate; replication being primarily a means of consuming more of the available energy (Annala & Annala, 2013). In the process governed by the second law, which entails the principle of least action, matter, energy and information are inter-converted. Most notably, energy is converted to information as peptides fold into proteins. This was demonstrated by Anfinsen (1973) when he showed that the enzymatic activity of ribonuclease was only present when the peptide was folded. His interpretation that the folding process was contingent on the amino acid sequence of the peptide is not generally true, but he did demonstrate the general principle that the

folding process generates information and since folding is a dissipative process, the information is the entropy of the process (see Fig. 1). However, although there are exceptions, phenotypic traits generally derive from several proteins interacting one with another as specified by the rules of engagement (Baverstock & Rönkkö, 2008). This interactivity gives rise to a near infinite potential for form incorporating a range of functions, which environmental conditioning can optimise (Baverstock & Rönkkö, 2008) and natural selection can act upon, in the context of the ecosystem, to give rise to stably replicating species (see Fig. 2). Thus, for example, mouse and man are phenotypically distinct organisms with

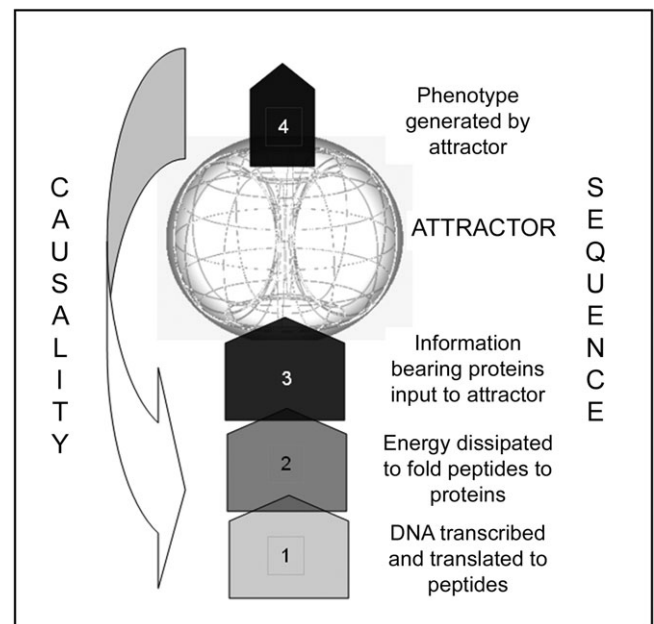


Figure 1. Genotype to phenotype: sequence and causality

Cellular phenotype (4) is generated from peptides derived from the genomic DNA coding sequences (1) through the dissipative folding process (2) to produce information bearing proteins (3) which interact dissipatively within the attractor (a minimum free energy state of the system) represented above by a torus. The dissipative processes are symmetry breaking so their products, particularly the proteins (3) and the cellular phenotype (4), are 'emergent', that is, have new properties that are not derivable from the processes that yield them. Specifically, the information acquired in the peptide folding process is unrelated to the sequence coding information in the DNA, which specifies the amino acid sequence of the peptide, and thus, the cellular phenotype is also unrelated to the DNA coding sequence. At each dissipative stage of the process the product, particularly the information generated, is the entropy of the system as stipulated by the second law. The sequence of events, therefore, should not be confused with causality, which acts downwardly from the cellular phenotype directing such processes as transcription and mRNA splicing. The cellular phenotype contains the information to provide, through a self-similar attractor state, the form and function of the organism it derives from (see Fig. 2). (The image of the torus is reproduced by kind permission of BioResilience.com/Galen Guerrero-Murphy.)

closely similar genotypes (Baverstock, 2011), that is, a near identical complement of peptides, which give rise through dissipative information generating processes within the cell, to two distinct information outputs (phenotypes). It is this common genotypic origin of mammals that gives rise to the scaling observed between, for example, body size and metabolic rate (Savage *et al.* 2007) over a wide range of biota and the mouse heart beating the same number of times in its lifetime as that of the human in its lifetime.

To some it maybe surprising that the conundrum of how the vast diversity of life (and its development from the zygote to adult organism) could boil down to the intangible and non-determinate process of converting energy into information in the peptide folding process.

Proteins have in recent times been regarded as of minor significance as compared to the ribonucleic acids in the attainment of the animate state, but in fact their interdependency is asymmetrical in the reverse direction: proteins can provide functionality to cells in the absence of DNA whereas the reverse is not true (Cox *et al.* 1976). The debate over whether the origins of development and inheritance lay in the nucleus or the cytoplasm smouldered over several decades (Sapp, 1987). It was effectively resolved in favour of the nucleus by the US geneticist T. H. Morgan in 1926 (Morgan, 1926), although interest was strongly maintained in the cytoplasmic origins in, for example, France and Germany until the discovery of the structure of DNA in 1953 appeared to resolve the issue

finally. Even so, scepticism was not totally extinguished and particularly in terms of development, the inheritance of cytoplasmic features is recognised (Griffiths & Gray, 2001).

It can be argued that, in the context of the evolution of bacterial antibiotic resistance, genes do matter, since either mutations or the horizontal transfer of specific DNA sequences, seem to confer such resistance. However, empirical evidence suggests that this is not as simple a process as just the acquisition of gene coding sequences. Bacteria in long term culture (20,000 generations) challenged with a reduced lactose environment, adapt far more rapidly than they acquire the mutations that might account for the adaptation (Barrick *et al.* 2009). Furthermore, bacteria engineered with an artificial network of mutually inhibitory gene coding sequences that permit adaptation to two different nutrient environments are able to adapt to a changed nutrient environment in less than an hour, in spite of the absence of pre-existing pathways (Kashiwagi *et al.* 2006). Finally, sequencing showed that the bacterium, *M. pneumoniae*, with a 'reduced genome', which has adapted to living in nutrient rich lungs, had lost many of its transcription factors that regulate metabolism, but in spite of this both environmental stresses and metabolic insults induced complex and specific transcriptional responses similar to more complex bacteria (Yus *et al.* 2009). These examples indicate empirically that the adaptation to environmental stress in bacteria is only at best very loosely coupled to the genome and involves features that are not resolvable in terms of the inheritance of DNA sequence alone.

In practical terms, evidence for the thermodynamic tenet can be seen in the dependence on latitude of the size of snakes (Head *et al.* 2009); the change in human height in the transition from foraging to farming due to a more restricted nutritional environment (http://en.wikipedia.org/wiki/Neolithic_Revolution) and in the relationship between the proximity of fast food outlet location and childhood obesity in England (Cetateanu & Jones, 2014). The reformulated foundation for biology upon which the above arguments are based throws light on the processes responsible for the phenomenon of genomic instability, on the issue of the 'missing heritability' and helps to explain how the overall process of evolution has been towards greater complexity (Annala & Baverstock, 2014).

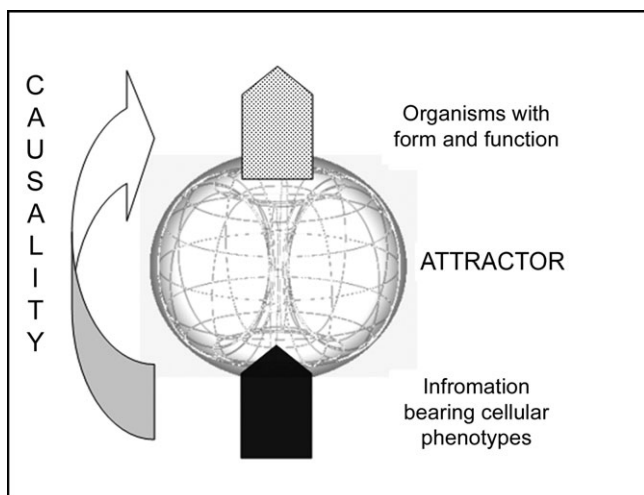


Figure 2. Cells to organisms

Information bearing cells interact dissipatively in the attractor (minimum free energy state of interacting cells) to provide the form and function (entropy in the form of information and bound energy) of the organism. Organisms, subject to natural selection for their ability to transduce energy from their ecosystem, grow and reproduce utilising the available energy as efficiently as possible according to the second law. Organisms are the emergent properties of interacting cells. (The image of the torus is reproduced by kind permission of BioResilience.com/Galen Guerrero-Murphy.)

Conclusions

The evolution of multicellular organisms with complex forms and functional abilities can be accounted for based on a fundamental tenet underpinned by the second law of thermodynamics, with natural selection acting on the ability of the organism to transduce energy (nutrient) most efficiently from its ecosystem by deploying that form and those functions. The information that gives rise to form

and function is dispersed throughout the organism in the constituent cellular phenotypes and derives mainly from the interactions between information bearing proteins. The concept of a gene, beyond a means of specifying the amino acid sequences of the peptides from which the proteins are formed, is both mostly unnecessary and possibly misleading.

References

- Anderson PW (1972). More is different. *Science* **177**, 393–396.
- Anfinsen CB (1973). Principles that govern the folding of protein chains. *Science* **181**, 223–230.
- Annala A & Annala E (2013). The significance of sex. *Biosystems* **110**, 156–161.
- Annala A & Baverstock K (2014). Genes without prominence: a reappraisal of the foundations of biology. *J R Soc Interface* **11**, 20131017.
- Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE & Kim JF (2009). Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247.
- Baverstock K (2011). A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype. *Prog Biophys Mol Biol* **106**, 443–449.
- Baverstock K (2013). Life as physics and chemistry: A system view of biology. *Prog Biophys Mol Biol* **111**, 108–115.
- Baverstock K & Rönkkö M (2008). Epigenetic regulation of the mammalian cell. *PLoS One* **3**, e2290.
- Blyth E (1835). An attempt to classify the ‘varieties’ of animals, with observations on the marked seasonal and other changes which naturally take place in various British species, and which do not constitute varieties. *The Magazine of Natural History* **8**, 40–53.
- Cetateanu A & Jones A (2014). Understanding the relationship between food environments, deprivation and childhood overweight and obesity: Evidence from a cross sectional England-wide study. *Health Place* **27C**, 68–76.
- Cox RP, Krauss MR, Balis ME & Dancis J (1976). Studies on cell communication with enucleated human fibroblasts. *J Cell Biol* **71**, 693–703.
- Gould SJ (2002). *The Structure of Evolutionary Theory*. Belknap, Cambridge, MA, USA; London.
- Griffiths PE & Gray RD (2001). Darwinism and developmental systems. In *Cycles of Contingency: Developmental Systems and Evolution*, eds Oyama S, Griffiths PE & Gray RD, pp. 195–218. MIT Press, Cambridge, MA, USA.
- Hall-Stoodley L, Costerton JW & Stoodley P (2004). Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol* **2**, 95–108.
- Head JJ, Bloch JJ, Hastings AK, Bourque JR, Cadena EA, Herrera FA, Polly PD & Jaramillo CA (2009). Giant boid snake from the Palaeocene neotropics reveals hotter past equatorial temperatures. *Nature* **457**, 715–717.
- Kadhim MA, Macdonald DA, Goodhead DT, Lorimore SA, Marsden SJ & Wright EG (1992). Transmission of chromosomal instability after plutonium alpha-particle irradiation. *Nature* **355**, 738–740.
- Kashiwagi A, Urabe I, Kaneko K & Yomo T (2006). Adaptive response of a gene network to environmental changes by fitness-induced attractor selection. *PLoS One* **1**, e49.
- Morgan TH (1926). Genetics and physiology of development. *American Naturalist* **60**, 489–515.
- Noble D (2013). Physiology is rocking the foundations of evolutionary biology. *Exp Physiol* **98**, 1235–1243.
- Ronkko M (2007). An artificial ecosystem: emergent dynamics and lifelike properties. *Artif Life* **13**, 159–187.
- Sapp J (1987). *Beyond the Gene: Cytoplasmic Inheritance and the Struggle for Authority in Genetics*. Oxford University Press, New York.
- Savage VM, Allen AP, Brown JH, Gillooly JF, Herman AB, Woodruff WH & West GB (2007). Scaling of number, size, and metabolic rate of cells with body size in mammals. *Proc Natl Acad Sci U S A* **104**, 4718–4723.
- Shapiro JA (2011). *Evolution: A View from the 21st Century*. FT Press Science, Upper Saddle River, NJ, USA
- Sharma V & Annala A (2007). Natural process – natural selection. *Biophys Chem* **127**, 123–128.
- Turkheimer E (2011). Still missing. *Res Hum Dev* **8**, 227–241.
- Ville R, Kaila I & Annala A (2008). Natural selection for least action. *Proc Roy Soc A* **464**, 3055–3070.
- Yang X, Xie L, Li Y & Wei C (2009). More than 9,000,000 unique genes in human gut bacterial community: estimating gene numbers inside a human body. *PLoS One* **4**, e6074.
- Yus E, Maier T, Michalodimitrakis K, van Noort V, Yamada T, Chen WH, Wodke JA, Guell M, Martinez S, Bourgeois R, Kuhner S, Raineri E, Letunic I, Kalinina OV, Rode M, Herrmann R, Gutierrez-Gallego R, Russell RB, Gavin AC, Bork P & Serrano L (2009). Impact of genome reduction on bacterial metabolism and its regulation. *Science* **326**, 1263–1268.

Additional information

Competing interests

None declared.

Funding

None declared.

Acknowledgements

We thank our reviewer for his challenges as devil’s advocate.