

Evolution viewed from physics, physiology and medicine

Denis Noble

Department of Physiology, Anatomy & Genetics

University of Oxford

Parks Road, Oxford OX1 3PT

Denis.noble@dpag.ox.ac.uk

Abstract

Stochasticity is harnessed by organisms to generate functionality. Randomness does not, therefore, necessarily imply lack of function or ‘blind chance’ at higher levels. In this respect, biology must resemble physics in generating order from disorder. This fact is contrary to Schrödinger’s idea of biology generating phenotypic order from *molecular* level order, which inspired the Central Dogma of Molecular Biology. The order originates at higher levels, which constrain the components at lower levels. We now know that this includes the genome, which is controlled by patterns of transcription factors and various epigenetic and reorganisation mechanisms. These processes can occur in response to environmental stress so that the genome becomes “a highly sensitive organ of the cell” (McClintock). Organisms have evolved to be able to cope with many variations at the molecular level. Organisms also make use of physical processes in evolution and development when it is possible to arrive at functional development without the necessity to store all information in DNA sequences. This view of development and evolution differs radically from that of Neo-Darwinism with its emphasis on blind chance as the origin of variation. Blind chance is necessary but the origin of functional variation is not at the molecular level. These observations derive from and reinforce the principle of biological relativity, which holds that there is no privileged level of causation. They also have important implications for medical science.

Keywords: Evolution and physiology, Schrödinger’s error, biological relativity, stochasticity, Neo-Darwinism, Modern Synthesis.

Introduction: the original formulation of the Neo-Darwinist Modern Synthesis

The theory of evolution by natural selection was formulated by Charles Darwin and Alfred Russel Wallace who presented their ideas to the Linnean Society of London in 1858, followed by Darwin's book *On the Origin of Species* in 1859. Darwin was cautious in the presentation of his ideas. He wrote "Natural Selection has been the main, but not the exclusive means of modification." He was concerned that he did not know the origin of variation and he acknowledged the existence of other mechanisms, including the inheritance of acquired characteristics. Ernst Mayr wrote in 1962: "Curiously few evolutionists have noted that, in addition to natural selection, Darwin admits use and disuse as an important evolutionary mechanism. In this he is perfectly clear." (1) Although Darwin disagreed with Lamarck on whether evolution had a direction (what Lamarck called *le pouvoir de la vie* (2, 3)), he nevertheless acknowledged "this justly celebrated naturalist . . . who upholds the doctrine that all species, including man, are descended from other species" (4). However, Darwin's multi-mechanism approach to evolution became significantly narrowed with the rise of Neo-Darwinism.

Weismann's formulation of Neo-Darwinism involved three major assumptions. First, that all genetic variation is random. Second, that the germline is isolated from variations in the soma. This is the Weismann Barrier. Third, together with these two assumptions, that natural selection is then all-sufficient (*allmacht*) to explain evolution (5) The subsequent integration of Mendelian genetics into this scheme led to the formulation of the Modern Synthesis (6).

Several important consequences followed. First, genetic variation is not itself viewed as functional. It becomes so only through the operation of natural selection to weed out harmful variations and promote helpful ones. The origin of variation is therefore completely blind. If this view is correct, we should not explain genetic variation in terms of existing or anticipated functionality. Since physiology is the study of functional processes in organisms, physiology is thereby excluded from any direct role in the source of variation. Second, the inheritance of acquired characteristics, often called Lamarckism, cannot occur since it would require either that the germ line is not isolated from influences of somatic variations and/or that some forms of functional genetic re-organisation can be triggered as a response to environmental stress. In an 1896 publication (7), Weismann added his theory of germinal selection, involving competition and selection among the hereditary units *within* the germplasm but, as Charlotte Weissman shows, this change in Weismann's view did not make any real concessions to the Lamarckians (8).

The Neo-Darwinist Modern Synthesis was therefore both an extension and a simplification of Darwin's ideas. It was an extension through the incorporation of Mendelian genetics, about which Darwin unfortunately knew nothing. It was a simplification because it excluded the inheritance of acquired characteristics, whereas Darwin not only included this form of inheritance, he even proposed a theory for how it could happen, his pangenesis theory of gemmules (9), which resembles some forms of such inheritance discovered recently (see section 4 below).

Purpose of this article

A central thesis of this paper is that blind stochasticity is a misconceived idea as it has been used in evolutionary biology. Stochasticity is used by organisms to generate new functional responses to environmental challenges. Far from proving that evolution is necessarily blind, randomness is the clay from which higher level order can be crafted.

I will present the case for the following theses, which run counter to Neo-Darwinism and the Modern Synthesis. With respect to Neo-Darwinism, the view in this paper is a replacement more than an extension.

1. Randomness (stochasticity) is what one should generally expect at the molecular level even if determinate functionality rules at higher (cellular, tissue, organ, systems, organisms, sociological) levels. Randomness and functionality necessarily co-exist at different levels.
2. Organisms can and do harness stochasticity in generating function.
3. Functional genome reorganisation can occur in response to environmental stress.
4. Non-DNA information can be transmitted across generations.
5. By using diverse higher-level processes organisms can resist potentially harmful effects of many random genetic variations, at lower levels of function .
6. Physical constraints can and must influence both development and evolution.
7. The gene-centric view has so far been very disappointing from the viewpoint of medicine.

1. Stochasticity and order coexist at different levels

Physics teaches us that at a molecular level there must be stochasticity. At any temperature above a value near absolute zero, below which a Bose-Einstein condensate becomes possible (10), molecules have kinetic energy which generates random movement. But physics also teaches us that, once there is a constraint at a higher level, e.g. a gas in a container, thermodynamics can describe determinate behaviour arising from the averaged behaviour within the constraint. This is the reason why Schrödinger argued correctly in *What is Life?* that physics generates order from disorder (11).

Yet he contrasted this with biology, which he described as generating order at a high level from *order* at a molecular level, i.e that the functional order at a high level actually results directly from order at the molecular level. But this is highly problematic from a physical viewpoint. Why then did he propose a theory that even he initially characterised as ridiculous? The reason is that following Delbrück (12), he predicted that the genetic material would be found to be an aperiodic crystal, which is a good description of DNA sequences if one thinks of a polymer as a kind of crystal. Crystal structure can be investigated accurately using diffraction. I believe he saw the ‘read-out’ of genetic sequences as determinate in the same kind of way. In this respect, he anticipated the formulation of Crick’s Central Dogma of molecular biology (13). Francis Crick and James Watson both acknowledged Schrödinger’s influence in their thinking about the Central Dogma.

There are two fatal problems with this approach, as noted by Jean-Jacques Kupiec (14, 15) The first is that, as is clear from Crick's original statement, the Central Dogma refers only to the fact that *sequence* information passes one way, from DNA to proteins:

“The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that *such information* cannot be transferred back *from protein* to either protein or nucleic acid.” (16)

I have italicised ‘*such information*’ and ‘*from protein*’ since it is evident that the statement does not say that *no* information can pass from the *organism* to the genome. In fact, it is obvious that it must do so to produce many different patterns of gene expression, which enable many different phenotypes (e.g many different cell types in the same body) to be generated from the same genome. In addition to controlling relative expression levels, the organism also makes use of protein-mediated protein processing to add yet another layer of control following transcription.

This information from organisms is conveyed to their genomes by patterns of transcription factors, genome marking, histone marking, and many RNAs, which in turn control the patterns of gene expression. These controls are exerted through preferential targeted binding to the genome or histone proteins. For example, methylation of cytosines preferentially occurs at CpG sites. Binding to histones preferentially occurs at the histone tails. Even though these are the targeted molecular mechanisms by which the functional control is exerted, there is no guarantee that the functionality will be evident at the molecular level. It would require many correlations between the *patterns* of binding and the functional processes at a higher level to identify the functionality involved. Without that correlation the binding patterns will appear random. There are simply far too many sites. There are millions of CpG sites in the whole genome and tens of thousands of CpG clusters, which significantly are located near gene regulatory sites (17).

The second problem is that, as Schrödinger must have understood as a physicist, there is no way in which the molecules in an organism can avoid stochasticity. He wrote:

‘We seem to arrive at the ridiculous conclusion that the clue to understanding of life is that it is based on a pure mechanism, a ‘clock-work’ in the sense of Planck’s paper’ (18)

But he confuses the logic by continuing: ‘The conclusion is not ridiculous and is, in my opinion, not entirely wrong, but it has to be taken “with a very big grain of salt”’. He then explains the ‘big grain of salt’ by showing that even clock-work is, ‘after all statistical’ (p. 103).

Schrödinger realises that something is far from right but is struggling to identify what it might be. We would now say that the molecules involved (DNA) *are* subject to frequent statistical variations (copying errors, chemical and radiation damage, etc.), which are then corrected by the cell’s protein and lipid machinery that enables DNA to become a highly reproducible molecule.(19) This is a three-stage process that reduces the copy error rate from 1 in 10^4 to around 1 in 10^{10} , which is an astonishing degree of accuracy. In a genome of 3 billion base pairs this works out as less than 1 error in copying a complete genome, compared to millions of errors without error correction. The order at the molecular scale is therefore

actually created by the system as a whole, including lipid components that are not encoded by DNA sequences (20). This requires energy, of course, which Schrödinger called negative entropy. Perhaps therefore this is what Schrödinger was struggling towards, but we can only see this clearly in retrospect. He could not have known how much the genetic molecular material experiences stochasticity and is constrained to be highly reproducible *by the organism itself*. The order at the molecular (DNA) level is actually imposed by higher level constraints.

2. Organisms can and do harness stochasticity in generating function

Stochasticity is a population level attractor

Experiments on the stochasticity of gene expression in cell populations show that, at least in some cases, it is the population as a whole that controls the stochasticity. Figure 1 is taken from (21).

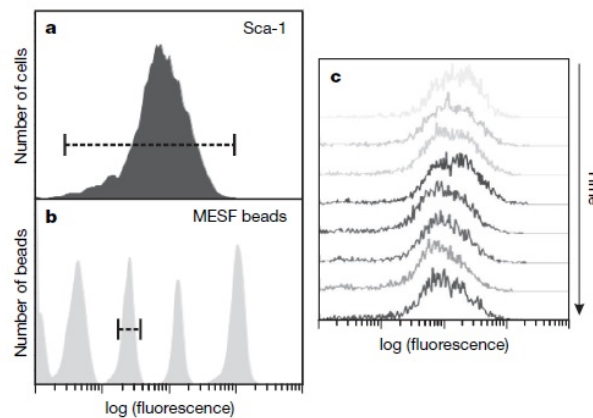


Figure 1. The robustness of heterogeneity of expression of Sca-1 protein expression in a cloned cell population. Heterogeneity detected by immunofluorescence flow cytometry (a) was significantly larger than the resolution limit of the method (b). (c) shows the stability of the clonal heterogeneity over a period of three weeks. Note that the spread of gene expression levels is three orders of magnitude (21).

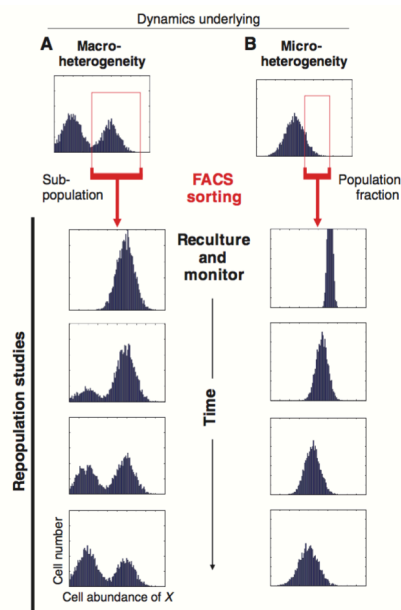


Figure 2. Two examples illustrating experiments in which populations were produced by cloning either from one of the peaks in a bimodal distribution (left) or from outliers in a monomodal distribution (right). In both cases the new population initially exhibits the range of expression of the parent subpopulation. Over time (several days) however the heterogeneity reverts to the original distribution (22).

The results show that in this case the range of gene expression is 1000 fold and it follows a simple bell-shaped curve. The range is a population level attractor, which is stable over long periods of time. That the population controls the heterogeneity is shown by experiments of the kind illustrated in Figure 2. In a cell population showing a bimodal distribution, new populations of cells were cloned from one of the peaks (left), while in a monomodal distribution cells were cloned from outliers. In both cases, after a few days the original heterogeneity became re-established.

Cell populations can therefore control stochasticity.

Cells can harness stochasticity to generate function

That cells can also harness stochasticity to generate specific function is known from experiments on the cells of the immune system that show the phenomenon of somatic hyper-mutation. Figure 3 summarizes what we know. Faced with a new antigen challenge, the mutation rate in the variable part of the genome can be accelerated by as much as 1 million times. So far as we know, the mutations occur randomly. But the location in the genome is certainly not random. The functionality in this case lies precisely in the targeting of the relevant part of the genome. The mechanism is directed, because the binding of the antigen to the antibody itself activates the proliferation process.

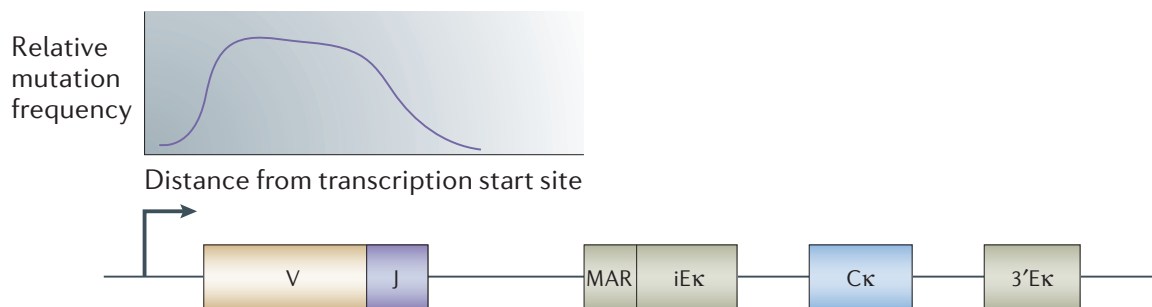


Figure 3. Schematic diagram of gene-specific targeted hyper-mutation in immunoglobulin gene loci. The mutation rate is greatly increased only in the variable part of the genome, which is a ~1.5 kilobase region in each of the three immunoglobulin loci. In this figure, the graph above the rearranged variable (V) and joining (J) gene segments that form the variable region of Igκ depicts the mutation domain in the κ-light chain (Igκ) locus. 3'Eκ, Igκ 3' enhancer; Cκ, Igκ constant; iEκ, Igκ intronic enhancer; MAR, matrix attachment region (23).

This example from the immune system shows that functionally significant targeted hyper-mutation can occur in the lifetime of an individual organism. There is no reason why this kind of mechanism should not be used in evolutionary change, as shown in the next example.

A well-known functionally-driven form of genome change is the response to starvation in bacteria. Starvation can increase the targeted reorganizations of the genome by five orders of magnitude, i.e. by a factor of over 100,000.(24, 25) This is one of the mechanisms by which bacteria can evolve very rapidly and in a functional way in response to environmental stress.

A similar targeting of location where genomic change can occur has been found in experiments on genetically modified fruit flies. One of the common ways in which genetic modification is achieved is to use a particular kind of mobile genetic element that can move around the genome using a cut-and-paste mechanism that does not require an RNA intermediate. Most often the insertions occur in a random way. But when DNA sequences from certain regulatory regions are used, they get inserted preferentially near the gene from which the sequence was derived (26). This process targets the changes in a way that is clearly not random with respect to possible function.

3. Functional genome reorganisation can occur in response to stress.

Barbara McClintock and the genome as an organ of the cell

Barbara McClintock first observed that whole domains of genetic material move around the genome, even from one chromosome to another. She was working on Indian corn in the 1930s and 1940s, but it was much later, in 1983, that she was recognised with the award of a Nobel Prize. In her Prize lecture she was very clear about the functional significance of her discovery. She described the genome “as a highly sensitive organ of the cell, monitoring genomic activities and correcting common errors, sensing the unusual and unexpected events, and responding to them, often by restructuring the genome.” (27).

She could not have anticipated the extent to which her idea would be confirmed by the sequencing of whole genomes. From the 2001 *Nature* paper on the first draft sequence of the human genome we have comparisons between sequences in completely different species of eukaryotes for two classes of proteins, transcription factor proteins and chromatin binding proteins (28). These show that the evolution of these proteins must have involved the movement of whole functional domains. This is far from the idea of slow progressive accumulation of point mutations. And it has much greater evolutionary significance since the rearrangement of whole domains including the functionality of those domains in response to stress could have been the origin of creativity in the evolutionary process. It is obvious that combining two or more domains each of which already has functionality is much more likely to produce a viable solution to a problem than waiting for random sorting of point mutations. This is why McClintock characterised the genome as a *highly sensitive organ of the cell*.

Can we observe genome reorganisation happening in evolutionary experiments?

We can now observe organisms making use of this ability to reorganise their genomes. Bos et al have observed the emergence of antibiotic resistance from multinucleated bacterial filaments. They write:

“the strategy of generating multiple mutant chromosomes within a single cell may represent a widespread and conserved mechanism for the rapid evolution of genome change in response to unfavorable environments (i.e., chemo-therapy drugs and antibiotics)” (29)

Jack et al (2015) have shown that

“signaling pathways that sense environmental nutrients control genome change at the ribosomal DNA. This demonstrates that not all genome changes occur at random and that *cells possess specific mechanisms to optimize their genome in response to the environment.*” (my italics). (30, 31)

How can genomes know about what is happening at the cell surface? The physiological mechanisms by which events in tiny micro-domains near the cell surface signal to the nucleus to control specific gene expression levels have now been studied in fine detail, see e.g. (31, 32). There is no longer any mystery in understanding the highly specific transmission of information to the nucleus that can control gene expression. There is no reason why genomes should not use similar communication pathways in response to stress signals received by cells and organisms.

4. Non-DNA information can be transmitted across generations.

Recent experiments have demonstrated that non-DNA information can be transmitted between generations (33), and this rapidly growing field has been reviewed in an important paper in *Science* (34). Two quotations from that review are relevant:

“Many phenomena and mechanisms of nongenetic and/or non–DNA sequence–based inheritance have been described in a range of model organisms, challenging our perception of the well-established relationship between transmitted genotype and phenotype.”

They conclude:

“The idea of certain sequences that might be refractory to germline epigenetic reprogramming provides a compelling mechanism for the inheritance of modulated epigenetic states.”

To illustrate the range of processes that can be involved, I will briefly describe three examples.

Rechavi et al investigated the inheritance of resistance to viral infection in the nematode worm, *C elegans* (35). The resistance is acquired when infected worms have the DNA required to make a viral-silencing RNAi, which is triggered by viral replication. They cross-bred these worms with a wild type population, including worms that do not have the required DNA. Some of the later generations have the required DNA, others do not. Yet subsequent generations inherited the acquired silencing response irrespective of whether they had the required DNA. The RNAi is inherited through the germline, and is then amplified by RNA polymerase in each generation. This non-DNA inheritance was followed successfully for 100 generations. It resembles Darwin’s gemmule theory (see Introduction).

Nelson et al found robust inheritance of epigenetic marking in mice with Apobec1 deficiency (36). They found that “these [epigenetic] effects persist for many generations and are as strong as conventional genetic inheritance.” The journal, PNAS, published a commentary

article in the same issue, which concludes: “the belief that the soma and germline do not communicate is patently incorrect.” (37).

The question whether epigenetic transmission of acquired characteristics could have been responsible for the evolution of separate species has been answered by Skinner et al who investigated the DNA mutations and non-DNA epigenetic changes in one of the icons of Darwinian speciation, the Galapagos finches (38). Five species were studied with different phylogenetic distances between them. Figure 4 shows the results. Both DNA mutations and epigenetic variations increase with the phylogenetic distance, with the epigenetic changes correlating better with distance. The authors conclude that both changes were involved in speciation and that they must have interacted.

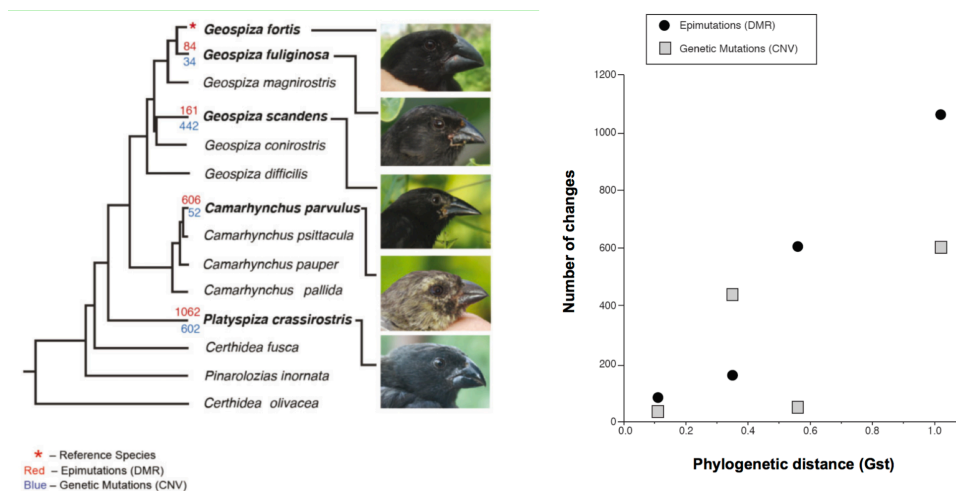


Figure 4. Left: Five of the Galapagos finch species were studied, the reference species *Geospiza fortis* and four others. The graph at the right shows the number of genetic and epigenetic changes plotted as a function of phylogenetic distance. The epigenetic changes correlate well with phylogenetic distance, the genetic mutations do not correlate as strongly. (38)

5. Organisms can resist the harmful effects of many molecular level variations.

One of my own fields of research is cardiac rhythm and arrhythmias. The main pacemaker in the heart, the sinus node, is an example of a robust functional process. Several different ionic transporter circuits are involved, any one of which could generate rhythm. The evolutionary advantage of this situation is obvious: if one mechanism fails another can take over the function. In 1992 we investigated this robustness by reverse engineering an experimentally based computer model. We found that removing a transporter that could carry as much as 80% of the ionic current necessary for generating the rhythm would change the overall frequency by only around 10-15% (39). Reverse engineering studies using a physiological model reveals the mechanism of the substitution. The small voltage changes that occur when one component is knocked out are sufficient to activate the substituting mechanism. This discovery formed the basis of the development of a safe heart slowing medication, ivabradine (40).

This kind of ‘back up’ of important physiological functions is ubiquitous. A systematic study of gene knockouts in yeast showed that 80% of knockouts have little or no effect on physiological functions under normal physiological conditions (41). Metabolic stress was needed to reveal the functional roles of most of the genes involved.

These studies pose a serious problem for bottom-up gene-centric theories of biology. The functionality will simply not be seen at that level or may be far from quantitatively accurate. Organisms seem to be very resourceful when challenged with knockouts, blockers or absence of nutrients. If we look for that ingenuity at the molecular level we may not find it.

Again, we can ask the question whether such processes can be demonstrated in actual evolutionary time. This was done recently by Taylor et al (42) who have shown that

“bacteria that have lost their flagella through deletion of the relevant DNA sequence can evolve the regulatory networks required to restore flagella and so restore motility in response to a stressful environment within just four days.”

That ability is a property of the bacterium regulatory networks and of the ability of the organism to signal the environmental pressure to the genome to activate mutation.

It is important to note that such examples, and the earlier ones I quoted above in section 3, involve what, so far as we know, are random mutations. At each location on the DNA sequence level this will therefore appear as ‘blind’ variation. At that level there will also be a form of Darwinian selection operating (14). But the targeting of particular locations, which is what enables the response to the environmental challenge to be effective, is not blind.

This mechanism has been extensively investigated by Richard Moxon and his colleagues who use the term ‘contingency locus’ to characterise the targeted loci of hypermutable DNA (43). In bacteria, these loci are simple sequence repeats in which the repeating unit is one to several nucleotides. In eukaryotes these loci are called microsatellites and often consist of hundreds of repeats. Since “mutation rates vary significantly at different locations within the genome” they propose that “it is precisely in the details of these differences and how they are distributed that major contributions to fitness are determined.” In an earlier article, Moxon and Thaler write “This phenotypic variation, which is stochastic with respect to the timing of switching but has a programmed genomic location, allows a large repertoire of phenotypic solutions to be explored, while minimizing deleterious effects on fitness.” (44).

6. Physical constraints can and must influence both development and evolution.

Natarajan et al, in a paper significantly entitled “predictable convergence in hemoglobin function has unpredictable molecular underpinnings”, have examined the molecular basis of convergence in hemoglobin function involving 56 avian taxa that have contrasting altitudinal range limits (45). They found that “Convergent increases in hemoglobin-oxygen affinity were pervasive among high-altitude taxa, but few such changes were attributable to parallel amino acid substitutions at key residues. Thus, predictable changes in biochemical phenotype do not have a predictable molecular basis.” This article beautifully illustrates the main point I am making in this paper, which is that unpredictability at the molecular level, which would lead one to think the changes are random, can be perfectly compatible with predictability and functionality at a higher level. This is biology’s equivalent of the physical principle that

determinate thermodynamics can co-exist with unpredictable stochastic behaviour at a molecular level. The difference is that, in biological systems, through the process of evolution, the higher level becomes functional. That is the level at which the functionality can be seen. It is then the level from which the lower level stochasticity can be understood, including the functional constraints.

If physics can be so important by using stochasticity in convergent evolution, can it also be important in a similar way in constraining development? It is tempting to think so since early embryonic development is similar in all multicellular eukaryotes, despite many differences in genome sequences. Edelman et al have explored this question by showing graphically how some simple physical constraints might be sufficient to explain certain aspects of embryonic development without having to assume that there must always be a specific DNA basis for all such processes (46). Their images are speculative and would require computational modelling to develop and test the ideas. Stuart Newman, Santiago Schnell and Philip Maini have led the way on this approach (47, 48). There must be interaction between overall physical constraints and molecular-level specifications. Erlich et al (49) show how modelling such physical constraints can account for the evolution of shell form in ammonites.

These examples illustrate a general point. Nature does not need to write to the ‘hard disc’ of the organism, its DNA, when it can get functions automatically from physical ‘free rides’, i.e. by letting physics do what it will do naturally. There is no need for DNA to be involved, for example, in ensuring that lipid membranes naturally fuse and form vesicles and many of the other properties of thin oily bilayers. And of course there is no DNA forming templates for the wide variety of lipids in organisms.

7. The gene-centric view has so far been very disappointing from the viewpoint of medicine.

There is another field of science where focusing on the molecular level has blinded us to functional processes at higher levels. That is the field of medicine. But before I explain why that is the case, I want to make it quite clear that I fully recognize the great scientific value of genome sequencing.

Sequencing whole genomes has been of immense value in evolutionary biological studies. The benefits for phylogeny and in discovering new parts of the ‘trees’ or ‘networks’ of life are obvious. It was sequencing that enabled Carl Woese to make his fundamental discovery of the archaea and how they differ from bacteria and eukaryotes.(50) Sequencing also enabled us to identify the extent to which mobile genetic elements must have been involved in the evolution of many proteins. In this sense, describing the genome as the ‘book of life’ has been a useful metaphor. But, as a metaphor used to publicise the health benefits that would accrue from genome sequencing it has been misinterpreted. The promise was that by a decade or so following sequencing of the human genome the ‘book of life’ would reveal how to treat cancer, heart disease, nervous diseases, diabetes, and many others through the discovery of many new pharmaceutical targets. This did not happen. An editorial in *Nature* in 2010 spelt this out:

“But for all the intellectual ferment of the past decade, has human health truly benefited from the sequencing of the human genome? A startlingly honest

response can be found on pages 674 and 676, where the leaders of the public and private efforts, Francis Collins and Craig Venter, both say 'not much'." (51)

The targets were identified all right. At least 200 new possible pharmaceutical targets are now known and there may be more to come, but we simply do not understand how to use them. The problem does not therefore lie in the absence of knowledge about the sequences. The problem is that we neglected to do the relevant physiology at the higher levels. A valuable critique of genotype-phenotype relations as a basis for the common-disease-common-variant hypothesis has been published by Joyner and Prendergast (52).

Before the shift towards genomic approaches to pharmacology, we did in fact have reasonably adequate methods for developing new drugs against specific diseases. The method was to work initially at a phenotype level to identify possible active compounds, and then to drill down towards individual protein or other molecular targets. This was the approach used so successfully by Sir James Black, the Nobel laureate discoverer of beta-blockers and H2 receptor blockers (53). It is the method by which the work of collaborators in my laboratory eventually led to the successful heart drug, ivabradine, to which I have already referred.

But the consequence of diverting large-scale funding towards the search for new drugs via genomics has been that the Black approach is now much less common and that the pharmaceutical industry is producing fewer new medications at vastly greater cost. Of course, the Black approach could and should be complemented by genomics, and there are successful cases where protein targets found by classical methods were later also identified as coding templates formed by particular genes. A good example is Duchenne muscular dystrophy, where the gene for the protein utrophin that can substitute, in mice at least, to cure the disease was discovered before the DNA sequence was identified (54).

Conclusions

There has been much debate about whether the Neo-Darwinist Modern Synthesis needs extending or replacing. Both views are correct. It depends on the context in which they are assessed. Theories in biology, as in any branch of science, can be judged by several criteria.

Falsifiability. The original Neo-Darwinist assumptions of the Modern Synthesis have been clearly falsified. I will consider the three basic assumptions outlined in the introduction.

The Weismann Barrier

The Weismann Barrier should be seen as a relative not an absolute barrier. Strict isolation of the genome was required in order to exclude the inheritance of acquired characteristics. Since we now know that acquired characteristics can be inherited, I believe it is more honest to admit that this reason for departing from Darwinism is no longer valid. In any case, the Barrier could only apply in those organisms that have a separate germ line. For the great majority of the duration of life on earth, there was no separate germ line. And plants can reproduce separately from their germ line. Quite simply, then, two of the original basic assumptions, isolation of the germ line and the impossibility of inheritance of acquired characteristics, can be seen to be incorrect.

Some criticisms of this conclusion refer to the rarity of experiments showing intergenerational transmission of epigenetic mutations. Originally, this was based on the idea that the genome was always wiped clean of epigenetic marking, so that it was thought that the idea was misconceived and impossible. As I have shown, this is simply not correct.

Another criticism was that it would not be robust. It has been demonstrated to persist for as much as 100 generations, and that it can, in some cases, be as robust as DNA transmission. Moreover, it does not need to be robust in all cases. As the review by Burggren shows, the softness and therefore reversibility of epigenetic inheritance is one of its evolutionary virtues (55). Sonia Sultan and her colleagues have also identified the factors that may determine the transience or persistence of epigenetic variation (56).

The third criticism is that it is observed in only rare cases. My reply is that so is speciation. Speciation is such a rare event that in thousands of years of selective breeding of cats, dogs, fish etc., we have not succeeded in producing new species, as defined by reproductive isolation.

Note also that these criticisms obviously do not apply to functionally significant reorganisation or hypermutation of genomes.

Blind stochasticity

The other basic assumption is blind stochasticity, meaning that what are seen as random genetic variations are not functionally directed. The concept of randomness is a major topic of research in philosophy, mathematics and physics. One way to by-pass these highly technical issues is to ask the question “random with respect to what”? The key in relation to evolutionary biology is whether variations are random with respect to function and whether they can be seen to be so. Even if the molecular level variations do in fact represent functional order at a higher level we will almost certainly require insight from the functional level to appreciate the functional nature of the molecular variations. The randomness I am referring to is therefore epistemological: without knowing the constraints by higher levels the variations will appear to be random and unpredictable. Once we know those constraints the possibility of prediction at the molecular level begins to exist. Whether it is computable is a very different question. Given the huge differences of scale, e.g. between molecular and cellular, it is implausible to expect molecular level computation alone to reveal the functionality.

Even before we consider whether a theory based on blind stochasticity has been falsified, we have to examine its conceptual status. A very basic lesson from physics is that stochasticity at lower, such as molecular, levels is not only inevitable as a consequence of molecular kinetic energy, it is also perfectly compatible with regular law-like behaviour at higher levels, a fact that was appreciated long ago by one of the founders of population genetics, R.A.Fisher (57). Even if behaviour at a high level is directed, stochasticity is what we can expect at lower levels. The example in this paper concerning the evolution in different species of hemoglobins at high altitude illustrates that point perfectly. As the authors of that paper say “predictable changes in biochemical phenotype do not have a predictable molecular basis.”(45). It is the physics of oxygen transport in organisms living at low partial pressures of oxygen that dictates the changes that occur to adapt to such environments, not specific changes in the genome.

From a gene-centric viewpoint, it could be objected that the genome changes are nevertheless those that enable the beneficial changes in oxygen transport to happen. That is certainly true. But it is precisely the higher level perspective that enables us to show that fact. What we can see here is that a conceptual issue, which is the question of the level at which functionality occurs, interacts with an empirical issue, which is whether the changes at the molecular level are predictable, *from that level alone*. Another way to put the conceptual issue is to say that, in any information transmission system, whether languages or genomes, sequences by themselves do not have meaning. They acquire meaning through their context, which can only be understood at a much higher level. As a linguistic example, the three letter alphabetic sequence 'but' has two totally different meanings and pronunciations in English and French. Similarly, genome sequences acquire meaning in their context. Sequences enabling arms, legs and eyes derive from organisms that had none of these.

Unravelling the problem

My paper unravels this problem by showing where some aspects of biological thought went wrong in the 20th century. Schrödinger's book, *What is Life?*, was a landmark in predicting correctly that the genetic material would be found to be an aperiodic crystal. But it contained the seeds of a major misunderstanding, leading Schrödinger, and then Crick and Watson, to maintain that, like a crystal, the genetic material could be read in a determinate way. That could be true only if the 'crystal', that is the linear polymer DNA, could be read and copied faithfully, with few or no copy errors. As we can now see, that is not an inherent property of DNA alone. On the contrary it is a property of the complex system by which the copy error rate can be reduced from an unacceptable frequency of millions per genome to less than 1. That is a higher level systems property of cells, including an army of proteins and lipids, not of DNA alone. In life as we know it on earth, this process occurs only in the context of living cells.

A possible objection to this conclusion is that all proteins have DNA templates that determine their amino acid sequences. That includes the proteins that contribute to the error-correcting systems for DNA. That is true, but it is usually taken a step further to mean that therefore the genome determines everything. That is not true. The error-correcting systems operate within cellular structures that contain molecular elements, such as lipid membranes, that do not require DNA templates in order to exist. Elsewhere, I have shown that the structural information in cells can be represented as comparable to that in the genome (58). Organisms always inherit both. In one of the rare examples of a successful clone from the nucleus of one species inserted into the enucleated but fertilized egg cell from another species, both the cell and the nucleus contribute to the final structure of the adult. Reproductive hybridization between species has also been shown to produce intermediate forms which can generate speciation (59).

Experimentally, we need to re-examine the way in which functional change in organisms can harness stochasticity at lower levels to create new functionality. Huang and his colleagues have shown the way forward here by demonstrating that stochasticity in gene expression is an attractor produced by a cell population. The many studies of targeted hyper-mutation, e.g by Richard Moxon's group, also show the way forward. Organisms in their evolution had to harness stochasticity since at a low enough level this is an inevitable property of the physics of molecular level systems that have kinetic energy.

We can now return to the question whether the assumption of blind stochasticity has been falsified. If the case presented in this paper is correct then one answer would be that it is very difficult for it to be falsified since stochasticity necessarily reigns at a low enough level, even if functionality reigns at higher levels. The constraints may have too subtle an effect at the molecular level. The falsifiability then depends on a prior conceptual question, which is whether one accepts multilevel functionality. A purely gene-centric theory does not accept multi-level functionality and can therefore maintain its view of everything being ‘blind chance followed by natural selection’.

To a physiologist or a medical scientist this is not a useful viewpoint. Functionality arises in organisms at many different levels. This is one of the bases of my formulation of the principle of Biological Relativity, first proposed in a previous article in this journal, and developed more completely in a book, *Dance to the Tune of Life. Biological Relativity* (60).

Utility. These points naturally lead to the other main criterion for judging a theory, which is its utility. Theories can be useful, even if they are false. Indeed, on a Popperian view of the logic of science that must always be true. We can only ever falsify theories about the natural world, never conclusively prove them. I want therefore to acknowledge the fact that the Neo-Darwinist Modern Synthesis was very useful. Whole fields of mathematical biology, such as population genetics, would not have flourished in the 20th century without the Modern Synthesis as a framework.

But, I also think that we have reached a watershed in relation to the issue of the utility of the Neo-Darwinist Modern Synthesis. As I have argued in detail elsewhere, there are too many experimental breaks with the original theory as formulated by Weismann and Wallace (61). Moreover the metaphorical language of Neo-Darwinism is a problem. The metaphors used strongly reinforce a simplistic gene-centric view. The time has come to see that evolutionary biology would progress faster if we used a different framework to develop a more inclusive theory, as illustrated in Figures 5 and 6.

Figure 5 shows the Extended Evolutionary Synthesis, which is represented as a development from the Neo-Darwinist Modern Synthesis, in turn developed from Darwinism.

Figure 6 shows the version of this diagram that better represents the conclusions of this paper. There are several important differences. First, it represents the fact that Darwin’s view of inheritance included the inheritance of acquired characteristics, which was excluded by Neo-Darwinism. Darwin’s concept of inheritance is therefore shown as being partly outside the Neo-Darwinist Modern Synthesis. Second, it represents the features of the extended synthesis (highlighted in bold in both Figure 5 and Figure 6) that lie outside the range of Neo-Darwinism as defined by Weismann and Wallace. The features of that theory that were excluded are shown as corresponding bold-face items. The highlighted items on the far left correspond with the highlighted items at the far right. Also included as a bold-face item is the principle of biological relativity. Although beyond the scope of this paper, I have included sexual selection.

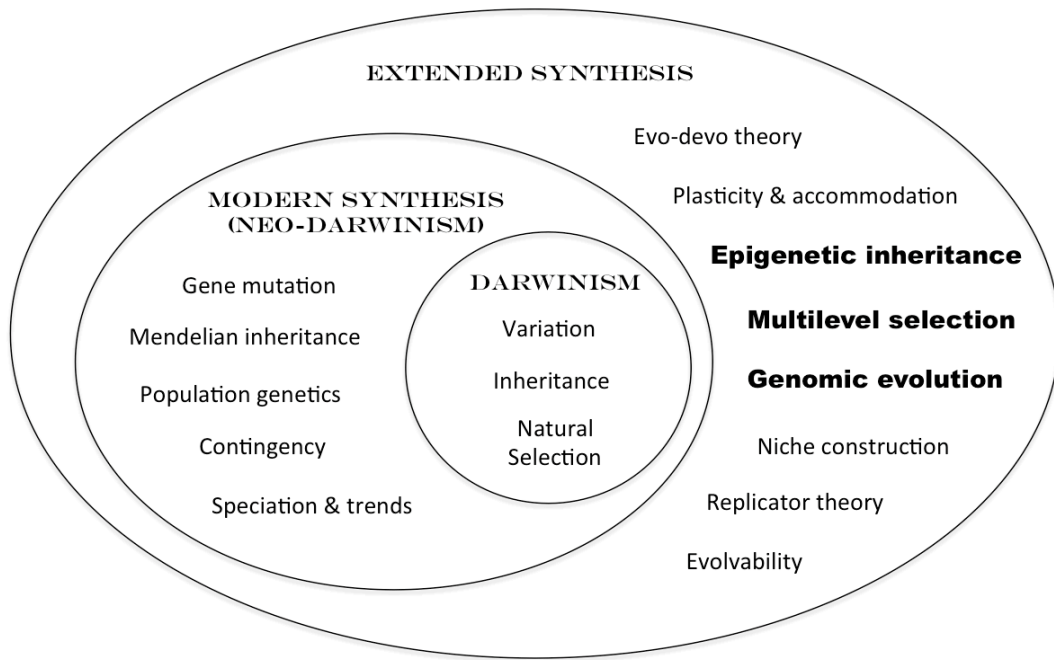


Figure 5. The Extended Evolutionary Synthesis representing the extension as extensions of Darwinism and then of the Neo-Darwinist Modern Synthesis (From (62)).

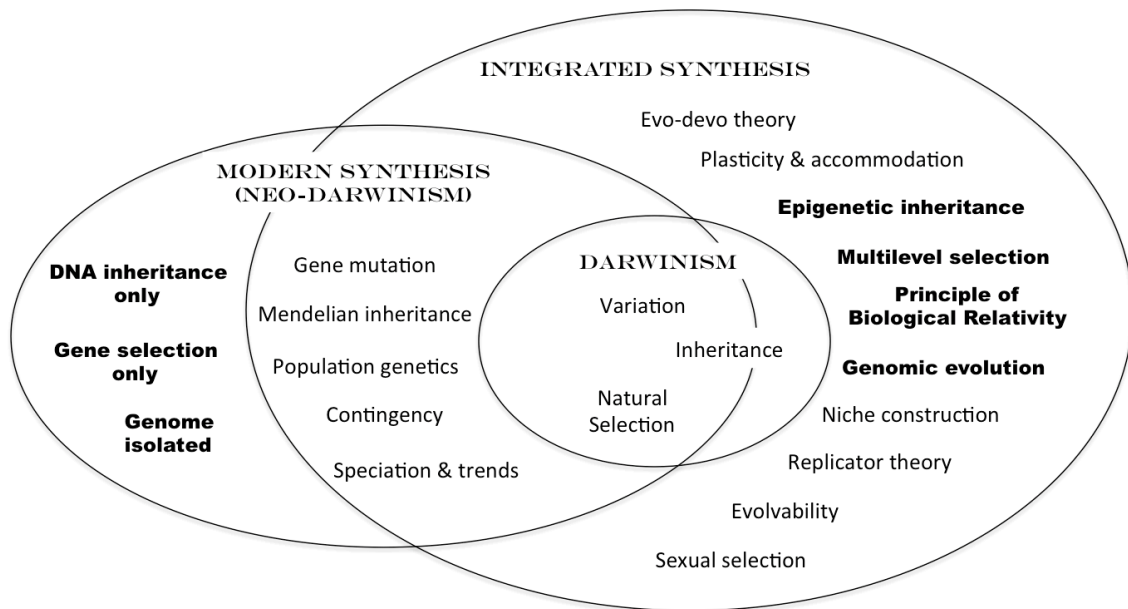


Figure 6. The Integrated Synthesis representing the extensions as extensions of Darwinism but only partially from Neo-Darwinism. Darwin's view of inheritance is also represented as extending outside the boundary of Neo-Darwinism/ (Developed for this article from (61), based on (62))

In spirit, this approach inherits an important part of Darwin's more nuanced philosophical approach. I emphasise *philosophical* here because it is obvious that we have moved way beyond what Darwin knew experimentally, as Figures 5 and 6 also show. But we can learn from his approach. Darwin was cautious in acknowledging the limits of what he knew. He was even unsure whether he had discovered the title of his book, since he did not know what produced variations in organisms, and he did not exclude the inheritance of acquired characteristics. Unjustified certainty is not the best way forward in scientific research. It remains open to further experimentation to clarify the extent of the many mechanisms now known to be available to Nature, and to determine how she used them, alone or more probably in various combinations, to evolve life as we now know it.

Acknowledgements

I wish to thank: Raymond Noble for discussion of evolutionary biology over many years; Sir Anthony Kenny for introducing me to the argument that a symbolic sequence in itself is meaningless out of its context, in a debate with Richard Dawkins in 1976; Jean-Jacques Kupiec for first pointing out the error in Schrödinger's *What is Life?*; Charles Taylor for his insights into the conceptual nature of teleology in a debate with me published in *Analysis* in 1967; James Shapiro for introducing me to the significance of mobile genetic elements; Lynn Margulis for introducing me to the importance of symbiogenesis in her debate with Richard Dawkins in 2009; Eva Jablonka for discussions on Lamarckism; Michael Joyner for his insights on the deficiencies of the gene-centric view in medicine; David Paterson for arranging and chairing the EB2015 Boston Conversation (URL); the co-organisers of a Balliol Interdisciplinary Institute (BII) project on the conceptual foundations of System Biology, Jonathan Bard, Tom Melham and Eric Werner; Nicholas Beale for valuable comments on a draft of this paper; Susan Noble for great support and criticism of my articles and books during the last years of her life – with Hilary Brown and Dario DiFrancesco she was responsible for the discovery that led to the life-saving drug ivabradine which depends precisely on the ability of the heart's pacemaker function to adapt to molecular level changes.

References

1. Mayr E. Introduction. *On the Origin of Species*. Cambridge, MA: Harvard; 1964. p. xxv–xxvi.
2. Lamarck J-B. *Histoire Naturelle des animaux sans vertebres*. Paris: Verdiere; 1815-1822.
3. Lamarck J-B. *Philosophie Zoologique*, original edition of 1809 with introduction by Andre Pichot. Paris: Flammarion; 1994.
4. Darwin C. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray; 1869.
5. Weismann A. *Die Allmacht der Naturzüchtung; eine Erwiderung an Herbert Spencer*. Jena: Fischer; 1893.
6. Huxley JS. *Evolution: the modern synthesis*. London: Allen & Unwin; 1942.
7. Weismann A. *On germinal selection as a source of definite variation*. Chicago: Open Court; 1896.
8. Weissman C. Germinal Selection: a Weismannian solution to Lamarckian problematics. In: Gissis SB, Jablonka E, editors. *Transformations of Lamarckism From subtle fluids to molecular biology*. Cambridge, MA: MIT; 2011. p. 57-66.
9. Darwin C. *The Variation of Animals and Plants under Domestication* London: John Murray; 1868.

10. Whitfield J. Molecules form new state of matter. *Nature*. 2003;doi:10.1038/news031110-16.
11. Schrödinger E. *What is Life?* Cambridge: Cambridge University Press; 1944.
12. Dronamraju KR. Erwin Schrödinger and the origins of molecular biology. *Genetics*. 1999;153:1071-6.
13. Crick FHC. On protein synthesis Symposia of the Society for Experimental Biology 1958;12:138-63.
14. Kupiec J-J. *The Origin of Individuals: A Darwinian Approach to Developmental Biology*. London: World Scientific Publishing Company; 2009.
15. Kupiec J-J. Cell differentiation is a stochastic process subjected to natural selection. In: Minelli A, Pradeu T, editors. *Towards a Theory of Development*. Oxford: OUP; 2014. p. 155-73.
16. Crick FHC. Central Dogma of Molecular Biology. *Nature*. 1970;227:561-3.
17. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes and Development*. 2011;25:1010-22.
18. Planck M. Dynamische und statistische Gesetzmäßigkeit. *Zeitschrift für Elektrochemie und angewandte physikalische Chemie* 1917;23:63.
19. McElhinny SAN, Pursell ZF, Kunkel TA. Mechanisms for High Fidelity DNA replication. In: Cox LS, editor. *Molecular themes in DNA replication*. London: RSC Publishing; 2009. p. 86-111.
20. Pani A, Dessi S. *Cell Growth and Cholesterol Esters*: Kluwer/Plenum; 2004.
21. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature*. 2008;453:doi:10.1038/nature06965.
22. Huang S. Non-genetic heterogeneity of cells in development: more than just noise. *Development*. 2009;136:3853-62.
23. Odegard VH, Schatz DG. Targeting of somatic hypermutation. *Nature Reviews Immunology*. 2006;8:573-83.
24. Jablonka E, Lamb M. *Evolution in Four Dimensions*. Boston: MIT Press; 2005/2014.
25. Bridges BA. Hypermutation under stress. *Nature*. 1997;387:557-8.
26. Bender. W, Hudson A. P element homing to the *Drosophila* bithorax complex. *Development*. 2000;127:3981-92.
27. McClintock B. The significance of responses of the genome to challenge *Science*. 1984;226:792-801.
28. Landers et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(860-921).
29. Bos J, Zhang Q, Vyawahare S, Rogers E, Rosenberg SM, Austin R. Emergence of antibiotic resistance from multinucleated bacterial filaments. *Proceedings of the National Academy of Sciences*. 2015;112:178-83.
30. Jack CV, Cruz C, Hull RM, Ralser M, Houseley J. Regulation of ribosomal DNA amplification by the TOR pathway. *Proceedings of the National Academy of Sciences*. 2015;112:9674-9.
31. Kar P, Mirams GR, Christian HC, Parekh AB. Control of NFAT Isoform Activation and NFAT-Dependent Gene Expression through Two Coincident and Spatially Segregated Intracellular Ca²⁺ Signals. *Molecular Cell*. 2016;64:746-59.
32. Ma H, Groth RD, Cohen SM, Emery JF, Li B, Hoedt E, et al. γ CaMKII shuttles Ca²⁺/CaM to the nucleus to trigger CREB phosphorylation and gene expression. *Cell*. 2014;159:281-94.

33. Tollefsbol T, editor. Transgenerational Epigenetics: Evidence and Debate. Waltham, MA: Academic Press; 2014.
34. Miska EA, Ferguson-Smith AC. Transgenerational inheritance: models and mechanisms of non-DNA sequence-based inheritance. *Science*. 2016;354:59-63.
35. Rechavi O, Minevish G, Hobert O. Transgenerational Inheritance of an Acquired Small RNA-Based Antiviral Response in *C. elegans* *Cell*. 2011;147:1248-56.
36. Nelson VR, Heaney JD, Tesar PJ, Davidson NO, Nadeau JH. Transgenerational epigenetic effects of Apobec1 deficiency on testicular germ cell tumor susceptibility and embryonic viability. *Proceedings of the National Academy of Sciences*. 2012;109: E2766–E73.
37. Mattick JS. Rocking the foundations of molecular genetics. *Proceedings of the National Academy of Sciences*. 2012;109:16400-1.
38. Skinner MK, Gurerrero-Bosagna C, Haque MM, Nilsson EE, Koops JAH, Knutie SA, et al. Epigenetics and the evolution of Darwin’s finches. *Genome Biology and Evolution*. 2014;6:1972-89.
39. Noble D, Denyer JC, Brown HF, DiFrancesco D. Reciprocal role of the inward currents $i_{b,Na}$ and i_f in controlling and stabilizing pacemaker frequency of rabbit sino-atrial node cells. *Proceedings of the Royal Society*. 1992;B 250:199-207.
40. DiFrancesco D, Camm JA. Heart rate lowering by specific and selective I(f) current inhibition with ivabradine: a new therapeutic perspective in cardiovascular disease. *Drugs*. 2004;64: 1757-65.
41. Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, et al. The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science*. 2008;320:362-5.
42. Taylor TB, Mulley G, Dills AH, Alsohim AS, McGuffin LJ, Studholme DJ, et al. Evolutionary resurrection of flagellar motility via rewiring of the nitrogen regulation system. *Science*. 2015;347:1014-7.
43. Moxon R, Bayliss C, Hood D. Bacterial Contingency Loci: The Role of Simple Sequence DNA Repeats in Bacterial Adaptation. *Annual Review of Genetics*. 2006;40:307-33.
44. Moxon ER, Thaler DS. The tinkerer’s evolving tool-box. *Nature*. 1997;387:659-62.
45. Natarajan C, Hoffmann FG, Weber RE, Fago A, Witt CC, Storz JF. Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science*. 2016;354:336-9.
46. Edelman DB, McMenamin M, Sheesley P, Pivar S. Origin of the vertebrate body plan via mechanically biased conservation of regular geometrical patterns in the structure of the blastula. *Progress in Biophysics and Molecular Biology*. 2016;121:212-44.
47. Müller G, Newman SA, editors. *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology* Bradford; 2003.
48. Schnell S, Maini PK, Newman SA, Newman T, Schatten GP, editors. *Multiscale Modeling of Developmental Systems*: Academic Press; 2007.
49. Ehrlich A, Moulton DE, Goriely A, Chirat R. Morphomechanics and Developmental Constraints in the Evolution of Ammonites Shell Form. *Journal of Experimental Zoology (Mol Dev Evol)*. 2016;00B:1-14.
50. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*. 1977;74:5088-90.
51. Editorial. The Human Genome at Ten. *Nature*. 2010;464:649-50.
52. Joyner MJ, Prendergast FG. Chasing Mendel: Five Questions for Personalized Medicine. *Journal of Physiology*. 2014;592:2381-8.

53. Black JW, Crowther AF, Shanks RG, Smith LH, Dornhorst AC. A new adrenergic betareceptor antagonist. *The Lancet*. 1964;283:1080-1.
54. Fairclough RJ, Wood MJ, Davies KE. Therapy for Duchenne muscular dystrophy: renewed optimism from genetic approaches. *Nature Reviews Genetics*. 2013;14:373-8.
55. Burggren W. Epigenetic Inheritance and its Role in Evolutionary Biology: Re-Evaluation and New Perspectives. *Biology*. 2016;5:doi: 10.3390/biology5020024.
56. Herman, J.J., Spencer, H.G., Donohue K, Sultan SE. How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. *Evolution*. 2013;68:632-43.
57. Fisher RA. Indeterminism and natural selection. *Philosophy of Science*. 1934;1:99-117.
58. Noble D. Differential and integral views of genetics in computational systems biology. *Interface Focus*. 2011;1:7-15.
59. Sun Y-H, Zhu Z-Y. Cross-species cloning: influence of cytoplasmic factors on development. *Journal of Physiology*. 2014;592:2375-9.
60. Noble D. *Dance to the Tune of Life. Biological Relativity*. Cambridge: Cambridge University Press; 2016.
61. Noble D. Evolution beyond neo-darwinism: a new conceptual framework. *Journal of Experimental Biology*. 2015;218:7-13.
62. Pigliucci M, Müller GB. *Evolution - The extended synthesis*. Cambridge, Mass: MIT Press; 2010.