



HAL
open science

From Information to Physics to Biology

Giuseppe Longo

► **To cite this version:**

Giuseppe Longo. From Information to Physics to Biology. Progress in Biophysics and Molecular Biology, 2022, 164 (September 2021), pp.46-62. hal-03935955

HAL Id: hal-03935955

<https://ens.hal.science/hal-03935955v1>

Submitted on 12 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cover Letter:

the following paper is an invited commentary, revised following the referees' advise

From Information to Physics to Biology

Giuseppe Longo

Commentary to "The gene: An appraisal" by Keith Baverstock,
PBMB, [Volume 164](#), September 2021, Pages 46-62.

Best regards

Giuseppe Longo

Centre Cavallès, CNRS et Ens, Paris

<http://www.di.ens.fr/users/longo>

Progress in Biophysics and Molecular Biology

From Information to Physics to Biology

--Manuscript Draft-- October, 2022

Abstract:

Commentary to "The gene: An appraisal" by Keith Baverstock

Note : this short commentary constructively criticizes the very interesting approach in the paper by an informal survey of the work that a few of us develop since several years. I will first survey the very pertinent critique of the Modern Synthesis and the genocentric approach, then hint to a methodological (and theoretical) critique of the approach by Baverstock and others and hint to alternatives paths that are compatible, but "extend" the physics for biology presented by the author. The purposes and the space allowed force a limited number of references and technical details. These may be found in the references contained in the few papers quoted below that are not the most nor the only representative contributions to that work, but are inserted as a source of references or as synthetic presentations of our views.

From Information to Physics to Biology

Giuseppe Longo

Centre Cavallès, CNRS et Ens, Paris

<http://www.di.ens.fr/users/longo>

Commentary to "The gene: An appraisal" by Keith Baverstock

Note: this short commentary constructively criticizes the very interesting approach in the paper by an informal survey of the work that a few of us develop since several years. I will first survey the very pertinent critique of the Modern Synthesis and the genocentric approach, then hint to a methodological (and theoretical) critique of the approach by Baverstock and others and hint to alternatives paths that are compatible, but "extend" the physics for biology presented by the author. The purposes and the space allowed force a limited number of references and technical details. These may be found in the references contained in the few papers quoted below that are not the most nor the only representative contributions to the that work, but are inserted as a source of references or as synthetic presentations of our views.

DNA, the organism and Evolution in an "informational" context

Baverstock's article conducts first a careful examination of the conceptual and theoretical errors that have accompanied the various projects aimed at the knowledge of DNA, human in particular. It recalls the persistence and for a long time, until year 2000, of the myths of the bijective determination by the genomic coding of polypeptides (primary sequences that, if sufficiently long and folded in three dimensions, form the active proteins in cells). So, the top managers of the Human Genome Project, until the eve of the announcement of the decoding of the human genome, estimated at 80,000 or more the human "genes", promising at the same time miracles of knowledge and therapeutics, once the chemical structure of DNA was going to be known. The hype on the potentialities of this new, and in fact very relevant knowledge, did not disappear once it was shown that human "genes" are in fact 25,000, indeed less What do they "determine" then, what sequences they "program" and, from there, what phenotypes? With ideological and non-scientific arrogance, too many continued to say that humanity had at last "decoded" the "book of life written by God" and that we could thus understand all or almost all of the biological dynamics, and ... cure cancer and definitively, within 15 years (!), wipe out all monogenetic diseases... diagnose, immediately, and cure, very soon, almost all human diseases (see (Longo 2018) for references). Unfortunately, God seems to have a strong propensity to write books with overlapping words (or genes), to dynamically "transpose" fragments of words from one place to another under the influence of changing contexts, as observed by B. McClintock, among many other peculiarities that make their reading very difficult. And even more so, their "editing".

Baverstock then rigorously explains why the analyses of even the later projects dedicated to genome wide associations (GWA) have had little clinical utility, particularly in view of the abundant pleiotropy and polymorphism that characterize gene-protein matches (i.e., they are of the "many-to-many" type, with peaks of 38,016 different peptides for the drosophila DSCAM gene and, conversely, the possibility that the same primary structure is produced by different DNA segments). So, the author very clearly explains the abuse that led to identify the "Mendelian genes" with a segment of DNA and how this has distorted research, clinical in particular, since very rarely diseases can be associated with molecular dynamics entirely intelligible in terms of gene expression, even as large networks of genome-wide associations (as claimed by the GWA project).

One would have also to wonder what was meant, and always has been, by "decoding" the genome. In general, if you have an "encoded message", to use a fashionable molecular terminology, as a sequence of signs, "decoding" means its translation into a language and context that is *completely meaningful* to the intelligence agent or the (biological) structure using it. Now, the "meaning", in a cell, in an organism, of a segment of DNA should at least be the *function* of its chemical structure in said context – that is what the DNA *does*. Baverstock illustrates how far we are from this, that is, from associating, in general and not in a few special cases, "DNA sequence information into the functional information that informs the phenotype."

Unfortunately, but perhaps just to conform to the dominant fashion, Baverstock continues to use "informational" language. I have criticized in several articles, including, most recently, (Longo 2019), the consequences of a terminology borrowed from other sciences. With an implicit reference to the two main theories of elaboration (Turing) and transmission (Shannon) of digital data, one imports a Laplacian "structure of determination", as Turing and Schroedinger explicitly acknowledge. That is, one thinks, first

and foremost, that determination implies predictability, since this is proper to both theories, and, again, Turing and Schrödinger say this with great lucidity – Shannon assumes it, since one of the purposes of his theory is exactly to reduce noise, that is to allow to predict the result of a well determined transmission. The first, Turing, stresses the Laplacian nature of his 1936 Logic Machine to process information, at the origin of computer science, in 1952; Schrödinger observes, in 1944, that " ...In calling the structure of the chromosome fibers a code-script we mean that the all penetrating mind, once conceived by Laplace, ... could tell from their structure whether the egg would develop, under suitable conditions, into a black cock or into a speckled hen...". Not only that, but the information theoretic language suggests the idea that dimensionality and materiality (hardware) do not matter. In fact, both digital theories allow to encode any finite dimension in one, a fundamental property of processing and transmission of discrete data types (digital data) that allows to process and transmit as sequences of digits even three-dimensional pixelated images, in perfect fidelity: in the mathematical discrete, there are no dimensional differences, everything is completely encoded in a linear sequence, an essential property of transmission and of computers' elaboration of data.

In addition, the independence of software from hardware makes computer science possible, as a general theory of processing sequences of signs, independently of the hardware. The same can be said of the transmission of information that can be done with smoke signals, frequency modulations, drums, sequences of bits The living is instead of a "radical materiality", that is, it is made only of these molecules, those that make up DNA, RNA and proteins, these membranes, and not others, in short, of this only "flesh" that makes us alive. At most, the non-specificity of macromolecular interactions makes it possible to replace a molecule (a hormone, for example) with another similar one, which may interact, with more or less probabilities, with the same cellular receptor (an endocrine disruptor, typically) – yet the non-specificity of macromolecular interactions should not be confused with the independence of information from the material support, it simply stresses the prevailing stochasticity of these interactions, that must be given in probabilities whose value depend on the context.

This idealistic obliteration of the dimensionality and materiality of the living, at all levels of organization, is one of the most serious distortions of knowledge that we owe to dominant molecular biology, along with the myth of the "exact stereospecificity" of molecular interactions, which has expelled randomness from molecular analyses, an ideological absurdity among the most serious, see (Paldi 2020) for a recent introduction. Again, informational language has contributed to this expulsion, since digital information is not processed nor transmitted at random: at most, probabilities, as a frequentist analysis, can help give syntactic relevance to sequences of signs, in Shannon's approach.

Let us therefore avoid to treat material flows and their gradients as "information", while stressing the contribution that this approach has given to the most deleterious of modern mechanisticisms, that of the cell as a "Cartesian mechanism", today a digital computer, a "boolean algebra", of which the Central Dogma is at the heart, although accompanied by the feedbacks described by Crick and Monod. The latter, if expressed in a digital language as they are, are forms of *recursion*, therefore perfectly described by Turing's calculus or, even better, by Church's lambda-calculus, another wonderful theory of computability with its usual structure of Laplacian determination (Longo 2018, 2019), without bothering the differential theory of control in the continuum by Wiener. The computational language is perfectly closed on discrete data types and it is fantastic for its purposes, which need to by-pass dimensions, materiality ... historicity (see below), that is all what matters in the analysis life.

In short, Baverstock falls into the linguistic trap that has done so much to structure the approach from which he wants to move away. Fortunately, this does not prevent him from telling very well a famous experience that falsifies the best known informational theory of Evolution, the Modern Synthesis (MS). It is in fact very interesting the analysis conducted in the paper concerning the Long-term Evolution Experiment (LTEE), a laboratory experience that followed 66,000 generations of a population of E. coli bacteria. The choice of E. Coli is part of the experimental practices, as this bacterium is "the standard model" of many laboratories. The standardization of the model is a good practice, but has also other motivations. Very few unicellular organisms survive and reproduce in solitude, as E. coli can do: the vast majority of them in fact needs a diverse ecosystem, populated by thousands or millions or more of different (micro-)organisms. These organisms horizontally exchange fragments of DNA, RNA, proteome... increasing at the same time both the production of diversity, through hybridization and contagions, and the forms of channeling and mutual control of evolutionary dynamics. The excellent arguments by Baverstock that use the LTEE to falsify the Modern Synthesis are a fortiori valid, if followed one by one, in reference to natural contexts, far from the artificial, though very interesting, evolution in the LTEE flasks: on the one hand the production of diversity and novelty is much greater, on the other hand the natural ecosystems constrains Evolution by excluding the incompatible in an even more effective way. Both phenomena, in nature, induce the change that matters most

in an evolutionary dynamics, as "history": the change of the "space of the possible", or of the "phase space", we will return to this. In fact, this change, unlikely, but not impossible, in a monospecies' ecosystem, is the focus of a much more radical criticism, which we will mention, of the MS and its "mathematical" heart, the simplistic Fisher's law - sometimes even called "theorem", while it is only an abstract principle, expressed in mathematical writing, and that the LTEE is sufficient to demonstrate of no empirical foundations.

The history of physics and physicalism in biology

I will address a methodological critique, as an a priori of any discussion concerning an analysis dealing with the *historicity* of Evolution and the *specificity* of the pertinent objects, the organisms. A critique by "facts" may only follow a perspective from which one may read the facts.

The main richness of physics in the last four centuries, which has contributed to its hegemony among the sciences, is mainly due to its many "theoretical inventions". Faced with a new phenomenon or simply a change in scale or a new look at a common phenomenon, physics has been able to invent radically original theories. It was enough to look differently at a falling stone, a body in friction on an inclined plane, and a profoundly original theory of gravitation and movement (inertial) was born, which had though nothing to do with the movements of the planets, according to Galileo. Newton will *unify* the two very different scales, falling apples and moving planets, introducing a further theoretical revolution. Faced with new thermal machines, physicists then invented a new theory with different observables, such as entropy. Again, the *unity* with particles movements will be found in Boltzmann's revolution, thanks to a third theory, based on new principles and concepts, the ergodic principle, the thermodynamic integral; both hypotheses and notions are infinite constructions, that is new asymptotic principles that allow to "understand together" particles in motion and macroscopic thermodynamic quantities like temperature. So Maxwell and Einstein will unify different fields (optics, electricity and magnetism, the first; gravitation and inertia, the second), further revolutions. Few measurements in microphysics, and there are those who dare to propose principles and theories incompatible with classical physics, based on the indetermination and non commutativity of measurement, to invent new spaces for quantum dynamics, the Hilbert spaces. The unity, in this great theoretical diversity, even a physical and mathematical incompatibility, is given by common principles of "conservation" (of energy, of momentum...), or rather by very general principles of "symmetry", as H. Weyl will explain (1929).

Many physicists, instead, turning to the living, locus of a rather original phenomenality, an entanglement of almost all physical scales, think that there is nothing else to say, that physical theories are enough to derive its fundamental properties. Suffice it instead to think that in a cell ... there is a lot of water and that the hydrodynamics of incompressible fluids is not describable in terms of either thermodynamics or quantum physics, that is, in terms of known theories of "particles". Partial results give bridges between the equations of Boltzmann and Schrödinger and those of Navier-Stokes (see Chibbaro et al book, reviewed in Longo (2016)), but we are far from a unification, which, as always, would presuppose the invention of a new theory. Yet, Bavestock tells us that the properties of the living are *derivable* from a thermodynamic principle.

The starting point is certainly valid: an organism is at *least* a thermodynamically open system, to which certainly apply principles of thermodynamics, Prigogine style - systems far from equilibrium (Nicolis, Prigogine 1977). However, even if an animal falls with the acceleration proposed by Galileo, very little is *derived* of its biological properties from Galileo's theory and just some more from Prigogine's thermodynamics (although both are very important sciences and do affect organisms). Beginning with the properties of water in a cell, observe that these properties need, in addition to hydrodynamics, also Quantum Electrodynamics, see (del Giudice 2007), all theories intractable from thermodynamics and long from being unified. Invoking "emergence" does not help much: the properties of incompressible fluids do not emerge from those of quanta, nor from statistical physics, just as Newton's laws or the relativistic field do not emerge from the properties of falling stones nor from the quantum field, respectively. To understand them, a new theory had or must be invented, which correlates these theories, or, better, unifies them - the reduction to an underlying level from which the "laws" of the one above would emerge does not exist in physics and the best approximation of such a practice, statistical physics, requires, as we said, asymptotic principles, that is, a new theoretical invention. In other words, when we pass from the theories of particles, be it statistical physics or quanta, to the hydrodynamics of incompressible fluids, we change theory by a conceptual transition to different pertinent observables that requires a new theoretical frame. The unity between phenomenal levels or scales, the only way we know to move among them, as between apples and planets, is to be invented - and much is being done. The notion of "emergence" masks an impossible reduction, and hides the need for a new unified frame, making the different phenomena and their physical transition intelligible.

Physics, or better, according to Baverstock, only one of its theories in the very rich theoretical scattering of often incompatible theories, should instead allow to derive the properties of the living on the basis of a single, in itself very interesting, principle of optimality, the principle of "least action".

Changing phase spaces, or the historicity of the living

Before further criticizing the thermodynamic perspective in biology, I would first like to recognize in its promoters important allies in the battle against the vision of the living as a "Cartesian mechanism", driven by the genetic program, a molecular machine for processing and transmitting information, fashionable distortion of biological knowledge. The understanding of the importance, in biology, of the thermodynamics of systems far from equilibrium is an important first step, an essential component of the intelligibility of life. But we must go further.

First, an organism is not a self-organizing system. It does not emerge *spontaneously* and *necessarily* under certain boundary conditions, like a hurricane or a flame. It is not a self-organizing system of flows, but it *uses* flows of energy and matter, starting from the most ancient cell capable of metabolism and reproduction. To say that this spontaneous emergence must have happened, once (!), at the origin of life, is like moving the problem of unification of statistical physics and hydrodynamics at the moment of the Big Bang or seven seconds later and call it "emergence": it does not help much and it even prevents to give an autonomous theory that unifies or at least relates the two theories, as physics has been able to do in the many examples described above, a possible and preliminary way also to understand what happened in the seven seconds after the Big Bang. So it is a good practice, closer to the method of physics, to give oneself first an adequate theory of the pertinent phenomenal level, the organism, as some attempt in many writings, see for example (Soto et al. 2016), and to unify, then, or better, to extend by it the pertinent physical theories (Longo 2020). Darwin followed this path and proposed a robust theory of Evolution, while avoiding to discuss the origin of life. He founded it on a "non-conservation" principle (heredity is "descent with modification"), while Hamilton was founding physics on conservation principles and deriving from them Maupertuis principle for classical dynamics.

In a Darwinian context, time has a novel and crucial role, both in phylogenesis and ontogenesis. Every organism has a history: instead, hurricanes and flames, on Earth, are of the same type, they are thermodynamically and mathematically identical, since four billion years, they have no history and can be described in the same "phase space". The living has changed quite a bit in this time – changing or new pertinent observables and parameters must be part of the analysis. And this is a crucial point: thermodynamics offers us a theory of irreversible time, which is necessary to understand the living, well beyond the time of classical and relativistic theories (Longo 2021), but is insufficient to grasp the additional *dimension* of historical time. This further form of time, historicity, is given by the change in phase space, a change not addressed by existing physical-mathematical theories, we will return to this, and by the role of rare events in determining phylogenetic trajectories, a role unknown to physical theories, even those dealing with "large fluctuations" (Longo 2018a). I am aware that this language is foreign both to the biologists of the "genetic program" and to the best physicalist reductionism such as Baverstock's: we are a small community trying to advance in a theory *proper to* the organism and for us nothing is understandable in biology without an analysis of its historical, that is phylogenetic and ontogenetic timing. Main stream, genocentric, biology goes elsewhere, like the fantastic main stream astronomy, trigonometry and geometry of epicycles in the Arabic world in year 1000 (Longo, Mossio 2020; Longo 2021a).

Second, let us return to the issue of "phase space". Kant beautifully framed the Newtonian practice in mathematical-physics by proposing an epistemic pre-existence of space and time to knowledge construction, as the locus of all possible trajectories, all already in potentia in space-time, or, also, he will see space and time as a condition of possibility of theorizing in physics – that is for writing equations and describing actual trajectories. The nineteenth century will extend this pre-condition of any theoretical description in physics to the "a priori" of the "space of phases", that is the choice of observables and parameters relevant to the theory intended – including space or time. A very productive theoretical boldness will lead physics to invent new spaces of phases for new theories, that is new observables, such as entropy, new pertinent parameters or to make variable the number of parameters (statistical physics), to propose spaces of infinite dimensions (Hilbert spaces in quantum mechanics), but each theory will have, and in principle, a pre-given space of phases. The symmetries proper to the description of these spaces make them always mathematically describable in the finite and a priori, even Hilbert spaces.

This is no longer valid in a historical science, such as biology. In biology, the historicity of the processes is due precisely to the dynamics of the space of the possible, a phase space, which changes along with the co-construction of phenotypes and organisms, their niches and ecosystems: for example, the observable "the

placenta" was not already there, albeit in *potentia*, before a retroviral infection 120 mln years ago. All forms of exaptation à la Gould propose new observables, if we reason at the relevant scale and biological observables, that of phenotypes and organisms. Their reduction to the level of colliding molecules stops, we said, at least in front of the theoretical irreducibility of incompressible fluids to classical or quantum molecular dynamics. But there is much more. Biological relativity, (Noble et al. 2019), does not allow to privilege one causal level over others: the heart and the vascular system, in their example, are formed at a critical transition in which the interactions between different scales and levels of organization allow the formation of the organ and its function, at once. Indeed, the formation of an organ in embryogenesis, but even more so the establishment of an evolutionarily novel phenotype, feathers and, then, wings for flight in dinosaurs, requires an extended critical transition (Longo, Montévil 2014), i.e., a mathematically dense cascade of changes in symmetries, which depends on the entire ecosystem context. For feathers and functional wings, it ranges from muscular and lung structure, to... air density: the phenomenon depends on all these constraints of which the last one is a typical stable boundary condition, of physical type, the others are co-constituted at and by the evolutionary event and depend on the very formation of the possibility to fly, which, for vertebrates did not exist before – it is a new observable, depending on new observables and parameters, all of them. The new possible actually depends on the global interaction of all these constraints, some of which are due to the constitutive dynamics itself; mathematically, in my view, it depends on a *non-local variable*, that is on a parameter depending on the global network of interactions making the novelty possible, but not necessary.

This is what we would like to add to Noble's biological relativity, the non-locality of parameters or of causal dependence: at least one of the pertinent parameters that allows/govern the new observable (the heart in embryogenesis, wings in Evolution...) depends on the entire new global structure that did not exist before. As for Evolution, this is part of its historicity, which does not allow to give, a priori, a phase space. (By the way, it is possible that Cosmology, at least when dealing with the formation of novelty at the early stages after the Big Bang, such as the fundamental constants of physics, is facing a similar challenge, historicity, a major problem for physicists.)

To give a principle of optimum, instead, such as Maupertius, we know from Hamilton that it is necessary to give, a priori, a space with a partial order, in which we can speak of extremal values, minima or maxima. This is the pre-given phase space that the historicity of the living forbids. In our perspective, the analysis of the dynamics of this changing space is part of the theory of Evolution and, therefore, of the organism (similarly, no way to derive the values of the fundamental constants of physics from some "optimality principle", when these constants are not yet fixed – infinitely many values are possible, thus the "many Universes" hypothesis in Cosmology). Mathematically, some of us are trying to treat this aspect of historical dynamics in life sciences as "heterogenesis" (Sarti et al. 2019), well beyond the (very interesting) physical morphogenesis to which far from equilibrium thermodynamics belongs.

To put it differently, but still summarizing apodictically a lot of work, history, evolutionary history in particular, does not follow optimal trajectories, therefore specific, but possible or generic trajectories, thus not necessary ones, in spaces co-constituted with and by the trajectories themselves. This process produces specific objects, that is, historical, individuated organisms in the case of the living. The physical trajectories, instead, are specific, geodesics in the appropriate space of phases, a space for each relevant theory, of course, but always given a priori. Physical objects are then generic, definable by a finite and a-temporal list of properties. The electron is a solution of the Dirac equation, just as a weight is perfectly described by some Galilean properties, without history. In contrast, a mouse is only definable in phylogenetic terms, telling its evolutionary story (Montévil 2020). In (Bailly, Longo 2009), we dared to fix a counting of some measurable features of phenotypes independently of their actual realization. The aim was to model Gould's analysis of "increasing complexity" in Evolution as a-symmetric random diffusion of "complexity" over bio-mass (Gould 1996), a new notion that we called anti-entropy. Anti-entropy is a new observable, proper to life, to be added to entropy/negentropy; its analysis does not assume extremal principles. The production of anti-entropy, typically, requires the production of entropy (it does not oppose to it, like negentropy), it is related to historical changes in the phase space (organisms' phenotypes and their pertinent parameters).

Let's follow the method, not the already given theories of physics: let's propose a theory adequate to the phenomenality of living, then we may try to correlate it, possibly as a non-conservative extension to the relevant physical theories (Longo 2020) - not easy because these theories are not unified. Non-conservative means that the extension of physics by biology can also allow us to explain, for example, the physical-chemical properties of the huge networks of molecular interactions that take place only in a living cell. These networks, from the physical point of view, have almost zero probability to exist and are found only inside cells. In short, there are macromolecular, thus physical-chemical activities, made possible only by biological

contexts. That is, there is no plausible physical explanation of the origin and maintenance of huge networks of chemical interactions without considering the cellular structure in which they take place. These networks contribute to continually regenerate the cell, its membranes and its other organs, which, in turn, make the networks possible, *enable* them. This is a form of "constrained process" that is eminently biological and quite different from the ones given within boundary conditions treated by physics, since the constraints are produced by the dynamics itself, including the interaction networks they enable. In my opinion, the best treatment to date of this "closure of constraints" may be found in (Montévil, Mossio 2015). In that theoretical framework, the DNA itself is analyzed as an internal constraint of the cell, permanently reconstructed by the molecular activities that it contributes to structure and canalize, in the cell. Obviously, a change in this constraint, the DNA, will affect cellular and organismal dynamics, but does not "drive" them. No "book of life" but an amazingly important constraint, sitting within each cell, a core component of the process of protein formation, continually modified and repaired by the macromolecular dynamics within the cell – like vital tissues and organs, at their scale, are repaired in an organism. For example, double-strand breaks can be repaired by non-homologous or microhomological-mediated end joining, which may also introduce changes in the DNA. By this, the molecular networks in the cell, by their dimensions and materiality, their variations, affect the DNA itself, while pressures and torsions on the chromatin modify the opening sites and, thus, in addition to (de-)methylation, change what matters in these processes, the *function* of the DNA.

This role of the context in the functioning of the DNA, as a constraint, is outside the scope of the Central Dogma (CD) as well of its negation: it is not a matter of "information" going back from the proteins to the DNA, but of protein networks, of three-dimensional and material, physical-chemical and biological contexts that modify the DNA or what DNA *does* in the cell. In this, I disagree with both Baverstock and Noble on their critique of the validity of the CD. The CD, as stated, in the language of information, cannot be falsified. As René Thom writes, the limit of truth is not falsehood, but insignificance – possibly by the construction of a language-frame that forces to miss what is significant. So, that language, joint to the CD, has proposed a perspective that made insignificant for biology the many material processes that are essential components of how the DNA works, often referred to as "epigenetic". Conversely, the CD itself is insignificant w.r. to that meaningful, material frame (including epigenetic phenomena). By this, the CD has had a major impact on research: in a community using the language of information, it let too many disregard the physico-chemical (and biological) role of the proteome, the cell, the organism and the environment in determining biological dynamics. Evolutionary theory (Modern Synthesis), research on GMOs in agriculture and on cancer (Longo 2018), as for some major examples, have been heavily affected.

Acknowledgments These views have been constructed thanks to the collaborations with F. Bailly, P.A. Miquel, M. Montévil, M. Mossio, C. Sonnenschein, A. Soto, A. Vianelli.

Minimal references (Longo's papers can be found in <https://www.di.ens.fr/users/longo/download.html>)

- Bailly, F., Longo, G. (2009). Biological Organization and Anti-Entropy. *J. Biological Systems*, Vol. 17, No. 1, pp. 63-96.
- Del Giudice E. (2007). Old and new views on the structure of matter and the special case of living matter. *Journal of Physics: conference Series*, 67, 012006.n..
- Gould, S.J. 1996. *Full House*. Three Rivers Press, New York.
- Longo, G. (2016). A review-essay on reductionism: some reasons for reading "*Reductionism, Emergence and Levels of Reality. The Importance of Being Borderline*", a book by S. Chibbaro, L. Rondoni, A. Vulpiani. In *Urbanomic*, <https://www.urbanomic.com/document/on-the-borderline/> London, May 8.
- Longo, G. (2018). Information and Causality: Mathematical Reflections on Cancer Biology. In *Organisms. Journal of Biological Sciences*, vol 2, n.1.
- Longo, G. (2018a). How Future Depends on Past Histories and Rare Events in Systems of Life, "How future depends on past histories and rare events in systems of life", *Foundations of Science* 23 (3):443-474 (2018)
- Longo, G. (2019). Information at the Threshold of Interpretation, Science as Human Construction of Sense. In Bertolaso, M. & Sterpetti, F. (Eds.) *Will Science Remain Human?* pp. 67-100, Springer, New York.
- Longo G. (2020). Naturalizing Physics. Or, embedding physics in the historicity and materiality of the living. *Deleuziana*, n. 11, special issue on "Differential Heterogenesis: Deleuze, Mathematics And The Creation Of Forms".
- Longo G. (2021). Confusing biological rhythms and physical clocks. Today's ecological relevance of Bergson-Einstein debate on time. In A. Campo, S. Gozzano (eds.), *Einstein vs Bergson. An enduring quarrel of time*, De Gruyter.
- Longo G. (2021a). Programming Evolution: a Crack in Science. *A Review of the book by Nobel winner, Jennifer A. Doudna, and Samuel H. Sternberg "A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution"* 2017, in *Organisms. J. Bio Sci.*, Vol. 5, No. 1.

- Longo, G. and Montévil M. (2014). *Perspectives on Organisms: Biological Time, Symmetries and Singularities*. Dordrecht: Springer.
- Longo, G., Mossio, M., 2020. Geocentrism vs genocentrism: theories without metaphors, metaphors without theories. *Interdisciplinary Science Reviews* 45, 380–405.
- Montévil, M. (2020) Historicity at the heart of biology. *Theory in Biosciences*, Springer Verlag.
- Montévil, M. & Mossio, M. (2015). Closure of constraints in biological organisation. *Journal of Theoretical Biology*, vol. 372: 179-191.
- Nicolis G., Prigogine I. (1977). *Self-organization in non-equilibrium systems*. New York, Wiley.
- Noble R., Tasaki K., Noble P., Noble D. (2019) Biological Relativity requires circular causality but not symmetry of causation: so, where, what and when are the boundaries? *Frontiers in Physiology* 10: 827.
- Paldi A. (2020). Stochastic or Deterministic? That is the Question. *Organisms. J. Bio Sci.*, vol. 4, N.1.
- Sarti,A., Citti G., Piotrowski, D (2019) Differential heterogenesis and the emergence of semiotic function. *Semiotica*. Vol 2019, Issue 230
- Soto A., Longo G., Noble D. (eds.) (2016) *From the century of the genome to the century of the organism: New theoretical approaches*, a Special issue of *Progress in Biophysics and Mol. Biology*, Vol. 122, 1, Elsevier.

From Information to Physics to Biology

Giuseppe Longo

Centre Cavallès, CNRS et Ens, Paris

<http://www.di.ens.fr/users/longo>

Commentary to "The gene: An appraisal" by Keith Baverstock, PBMB, [Volume 164](#), September 2021, Pages 46-62.

Note: this short and informal commentary constructively criticizes the very interesting approach in the paper by a brief survey of the work that a few of us develop since several years. I will first recall the very pertinent critique of the Modern Synthesis and the genocentric approach presented in the paper, then suggest a methodological (and theoretical) critique of the approach by K. Baverstock and hint to alternatives paths that are compatible, but "extend" the physics for biology presented by the author. The purposes and the space allowed force a limited number of references and technical details. These may be found in the references contained in the few papers quoted below that are not the most nor the only representative contributions to the that work, but are inserted as a source of references or as synthetic presentations of our views.

DNA, the organism and Evolution in an "informational" context

Baverstock's article conducts first a careful examination of the conceptual and theoretical errors that have accompanied the various projects aimed at the knowledge of DNA, human in particular. It recalls the persistence and for a long time, until year 2000, of the myths of the bijective determination by the genomic coding of polypeptides (primary sequences that, if sufficiently long and folded in three dimensions, form the active proteins in cells). So, the top managers of the Human Genome Project, until the eve of the announcement of the decoding of the human genome, estimated at 80,000 or more the human "genes", promising at the same time miracles of knowledge and therapeutics, once the chemical structure of DNA was going to be known (Collins 1999), see also (Liang et al 2000) for further estimates at the eve of "decoding". The hype on the potentialities of this new, and in fact very relevant knowledge, did not disappear once it was shown that human "genes" are in fact 25,000, indeed less What do they "determine" then, what sequences they "program" and, from there, what phenotypes? With ideological and non-scientific arrogance, too many continued to say that humanity had at last "decoded" the "book of life written by God" and that we could thus understand all or almost all of the biological dynamics, and ... cure cancer and definitively, within 15 years (!), wipe out all monogenetic diseases... diagnose, immediately, and cure, very soon, almost all human diseases (see (Longo 2018) for references). Unfortunately, God seems to have a strong propensity to write books with overlapping words (or genes), to dynamically "transpose" fragments of words from one place to another under the influence of changing contexts, as observed by B. McClintock, among many other peculiarities that make their reading very difficult. And even more so, their "editing".

Baverstock then rigorously explains why the analyses of even the later projects dedicated to genome wide associations (GWA) have had little clinical utility, particularly in view of the abundant pleiotropy and polymorphism that characterize gene-protein matches (i.e., they are of the "many-to-many" type, with peaks of 38,016 different peptides for the drosophila DSCAM gene and, conversely, the possibility that the same primary structure is produced by different DNA segments). So, the author very clearly explains the abuse that led to identify the "Mendelian genes" with a segment of DNA and how this has distorted research, clinical in particular, since very rarely diseases can be associated with molecular dynamics entirely intelligible in terms of gene expression, even as large networks of genome-wide associations (as claimed by the GWA project).

One would have also to wonder what was meant, and always has been, by "decoding" the genome. In general, if you have an "encoded message", to use a fashionable molecular terminology, as a sequence of signs, "decoding" means its translation into a language and context that is *completely meaningful* to the intelligence agent or the (biological) structure using it. Now, the "meaning", in a cell, in an organism, of a segment of DNA should at least be the *function* of its chemical structure in said context – that is what the DNA *does*. Baverstock illustrates how far we are from this, that is, from associating, in general and not in a few special cases, "DNA sequence information into the functional information that informs the phenotype."

Unfortunately, but perhaps just to conform to the dominant fashion, Baverstock continues to use "informational" language. I have criticized in several articles, including, most recently, (Longo 2019), the consequences of a terminology borrowed from other sciences. With an implicit reference to the two main theories of elaboration (Turing) and transmission (Shannon) of digital data, one imports a Laplacian "structure of determination", as Turing and Schroedinger explicitly acknowledge. That is, one thinks, first and foremost, that determination implies predictability, since this is proper to both theories, and, again, Turing and Schrödinger say this with great lucidity – Shannon assumes it, since one of the purposes of his theory is exactly to reduce noise, that is to allow to predict the result of a well determined transmission (see (Longo 2019) for references). In short, the first, Turing, stresses the Laplacian nature of his 1936 Logic Machine to process information, at the origin of computer science, in 1952; Schrödinger observes, in 1944 "What is life?", that " ...In calling the structure of the chromosome fibers a code-script we mean that the all penetrating mind, once conceived by Laplace, ... could tell from their structure whether the egg would develop, under suitable conditions, into a black cock or into a speckled hen..." (p. 7). Not only that, but the information theoretic language suggests the idea that dimensionality and materiality (hardware) do not matter. In fact, both digital theories allow to encode any finite dimension in one, a fundamental property of processing and transmission of discrete data types (digital data) that allows to process and transmit as sequences of digits even three-dimensional pixelated images, in perfect fidelity: in the mathematical discrete, there are no dimensional differences, everything is completely encoded in a linear sequence, an essential property of transmission and of computers' elaboration of data.

In addition, the independence of software from hardware makes computer science possible, as a general theory of processing sequences of signs, independently of the hardware. The same can be said of the transmission of information that can be done with smoke signals, frequency modulations, drums, sequences of bits The living is instead of a "radical materiality", that is, it is made only of these molecules, those that make up DNA, RNA and proteins, these membranes, and not others, in short, of this only "flesh" that makes us alive. At most, the non-specificity of macromolecular interactions makes it possible to replace a molecule (a hormone, for example) with another similar one, which may interact, with more or less probabilities, with the same cellular receptor (an endocrine disruptor, typically) – yet the non-specificity of macromolecular interactions should not be confused with the independence of information from the material support, it simply stresses the prevailing stochasticity of these interactions, that must be given in probabilities whose value depend on the context.

This idealistic obliteration of the dimensionality and materiality of the living, at all levels of organization, is one of the most serious distortions of knowledge that we owe to dominant molecular biology, along with the myth of the "exact stereospecificity" of molecular interactions, which has expelled randomness from molecular analyses, an ideological absurdity among the most serious, see (Paldi 2020) for a recent introduction. Again, informational language has contributed to this expulsion, since digital information is not processed nor transmitted at random: at most, probabilities, as a frequentist analysis, can help give syntactic relevance to sequences of signs, in Shannon's approach.

Let us therefore avoid to treat material flows and their gradients as "information", while stressing the contribution that this approach has given to the most deleterious of modern mechanisms, that of the cell as a "Cartesian mechanism", today a digital computer, a "boolean algebra", of which the Central Dogma is at the heart, although accompanied by the feedbacks described by Crick and Monod. The latter, if expressed in a digital language as they are, are forms of *recursion*, therefore perfectly described by Turing's calculus or, even better, by Church's lambda-calculus, another wonderful theory of computability with its usual structure of Laplacian determination (Longo 2018, 2019), without bothering the differential theory of control in the continuum by Wiener. The computational language is perfectly closed on discrete data types and it is fantastic for its purposes, which need to by-pass dimensions, materiality ... historicity (see below), that is all what matters in the analysis life.

In short, Baverstock falls into the linguistic trap that has done so much to structure the approach from which he wants to move away. Fortunately, this does not prevent him from telling very well a famous experience that provides a robust counter-example the Modern Synthesis in Evolutionary Theories (MS) – this "synthesis" has become the core informational approach to Evolution since the '60s (Gouyon et al 2002). It is in fact very interesting the analysis conducted in the paper concerning the Long-term Evolution Experiment (LTEE), a laboratory experience that followed 66,000 generations of a population of *E. coli* bacteria. The choice of *E. coli* is part of the experimental practices, as this bacterium is "the standard model" of many laboratories. The standardization of the model is a good practice, but has also other motivations. Very few unicellular organisms

survive and reproduce in solitude, as *E. coli* can do: the vast majority of them in fact needs a diverse ecosystem, populated by thousands or millions or more of different (micro-)organisms. These organisms horizontally exchange fragments of DNA, RNA, proteome... increasing at the same time both the production of diversity, through hybridization and contagions, and the forms of channeling and mutual control of evolutionary dynamics. The excellent arguments by Baverstock that use the LTEE to criticize the Modern Synthesis are a fortiori valid, if followed one by one, in reference to natural contexts, far from the artificial, though very interesting, evolution in the LTEE flasks: on the one hand the production of diversity and novelty is much greater, on the other hand the natural ecosystems constrains Evolution by excluding the incompatible in an even more effective way. Both phenomena, in nature, induce the change that matters most in an evolutionary dynamics, as "history": the change of the "space of the possible", or of the "phase space", we will return to this. In fact, this change, unlikely, but not impossible, in a monospecies' ecosystem, is the focus of a much more radical criticism, which we will mention, of the MS and its "mathematical" heart, the simplistic Fisher's law - sometimes even called "theorem", while it is only an abstract principle, expressed in mathematical writing, and that the LTEE is sufficient to demonstrate of no empirical foundations.

The history of physics and physicalism in biology

I will address a methodological critique, as an a priori of any discussion concerning an analysis dealing with the *historicity* of Evolution and the *specificity* of the pertinent objects, the organisms. A critique by "facts" may only follow a perspective from which one may read the facts. Indeed, the history of physics, I will hint, proposed a major diversity of perspectives into phenomena of the inert – it is a fortiori so when addressing issues pertaining the living state of matter.

The main richness of physics in the last four centuries, which has contributed to its hegemony among the sciences, is mainly due to its many "theoretical inventions". Faced with a new phenomenon or simply a change in scale or a new look at a common phenomenon, physics has been able to invent radically original theories. It was enough to look differently at a falling stone, a body in friction on an inclined plane, and a profoundly original theory of gravitation and movement (inertial) was born, which had though nothing to do with the movements of the planets, according to Galileo. Newton will *unify* the two very different scales, falling apples and moving planets, introducing a further theoretical revolution. Faced with new thermal machines, physicists then invented a new theory with different observables, such as entropy. Again, the *unity* with particles movements will be found in Boltzmann's revolution, thanks to a third theory, based on new principles and concepts, the ergodic principle, the thermodynamic integral; both hypotheses and notions are infinite constructions, that is new asymptotic principles that allow to "understand together" particles in motion and macroscopic thermodynamic quantities like temperature. So Maxwell and Einstein will unify different fields (optics, electricity and magnetism, the first; gravitation and inertia, the second), further revolutions. Few measurements in microphysics, and there are those who dare to propose principles and theories incompatible with classical physics, based on the indetermination and non commutativity of measurement, to invent new spaces for quantum dynamics, the Hilbert spaces. The unity, in this great theoretical diversity, even a physical and mathematical incompatibility, is given by common principles of "conservation" (of energy, of momentum...), or rather by very general principles of "symmetry", as H. Weyl will explain (1929).

Many physicists, instead, turning to the living, locus of a rather original phenomenality, an entanglement of almost all physical scales, think that there is nothing else to say, that physical theories are enough to derive its fundamental properties. Suffice it instead to think that in a cell ... there is a lot of water and that the hydrodynamics of incompressible fluids is not describable in terms of either thermodynamics or quantum physics, that is, in terms of known theories of "particles". Partial results give bridges between the equations of Boltzmann and Schrödinger and those of Navier-Stokes (see Chibbaro et al book, reviewed in Longo (2016)), but we are far from a unification, which, as always, would presuppose the invention of a new theory. Yet, Baverstock tells us that the properties of the living are *derivable* from a thermodynamic principle.

The starting point is certainly valid: an organism is at *least* a thermodynamically open system, to which certainly apply principles of thermodynamics, Prigogine style - systems far from equilibrium (Nicolis, Prigogine 1977). However, even if an animal falls with the acceleration proposed by Galileo, very little is *derived* of its biological properties from Galileo's theory and just some more from Prigogine's thermodynamics (although both are very important sciences and do affect organisms). Beginning with the properties of water in a cell, observe that these properties need, in addition to hydrodynamics, also Quantum Electrodynamics, see (del Giudice 2007),

all theories intractable from thermodynamics and long from being unified. Invoking "emergence" does not help much: the properties of incompressible fluids do not emerge from those of quanta, nor from statistical physics, just as Newton's laws or the relativistic field do not emerge from the properties of falling stones nor from the quantum field, respectively. To understand them, a new theory had or must be invented, which correlates these theories, or, better, unifies them - the reduction to an underlying level from which the "laws" of the one above would emerge does not exist in physics and the best approximation of such a practice, statistical physics, requires, as we said, asymptotic principles, that is, a new theoretical invention. In other words, when we pass from the theories of particles, be it statistical physics or quanta, to the hydrodynamics of incompressible fluids, we change theory by a conceptual transition to different pertinent observables that requires a new theoretical frame. The unity between phenomenal levels or scales, the only way we know to move among them, as between apples and planets, is to be invented - and much is being done. The notion of "emergence" masks an impossible reduction, and hides the need for a new unified frame, making the different phenomena and their physical transition intelligible.

Physics, or better, according to Baverstock, only one of its theories in the very rich theoretical scattering of often incompatible theories, should instead allow to derive the properties of the living on the basis of a single, in itself very interesting, principle of optimality, the principle of "least action".

Changing phase spaces, or the historicity of the living

Before further criticizing the thermodynamic perspective in biology, I would first like to recognize in its promoters important allies in the battle against the vision of the living as a "Cartesian mechanism", driven by the genetic program, a molecular machine for processing and transmitting information, fashionable distortion of biological knowledge. The understanding of the importance, in biology, of the thermodynamics of systems far from equilibrium is an important first step, an essential component of the intelligibility of life. But we must go further.

First, an organism is not a self-organizing system. It does not emerge *spontaneously* and *necessarily* under certain boundary conditions, like a hurricane or a flame. It is not a self-organizing system of flows, but it *uses* flows of energy and matter, starting from the most ancient cell capable of metabolism and reproduction. To say that this spontaneous emergence must have happened, once (!), at the origin of life, is like moving the problem of unification of statistical physics and hydrodynamics at the moment of the Big Bang or seven seconds later and call it "emergence": it does not help much and it even prevents to give an autonomous theory that unifies or at least relates the two theories, as physics has been able to do in the many examples described above, a possible and preliminary way also to understand what happened in the seven seconds after the Big Bang. So it is a good practice, closer to the method of physics, to give oneself first an adequate theory of the pertinent phenomenal level, the organism, as some attempt in many writings, see for example (Soto et al. 2016), and then try to unify, or better, to extend the pertinent physical theories, if possible, by this new theory (Longo 2020). Darwin followed the first step of this path, by a theoretical invention, and proposed a robust theory of Evolution, while avoiding to discuss the origin of life – which may require an interface with physical theories. He founded it on a "non-conservation" principle (heredity is "descent with modification"), while Hamilton was founding physics on conservation principles and deriving from them Maupertuis (optimality) principle for classical dynamics.

In a Darwinian context, time has a novel and crucial role, both in phylogenesis and ontogenesis. Every organism has a history: instead, hurricanes and flames, on Earth, are of the same type, they are thermodynamically and mathematically identical, since four billion years, they have no history and can be described in the same "phase space". The living has changed quite a bit in this time – changing or new pertinent observables and parameters must be part of the analysis. And this is a crucial point: thermodynamics offers us a theory of irreversible time, which is necessary to understand the living, well beyond the time of classical and relativistic theories (Longo 2021), but is insufficient to grasp the additional *dimension* of historical time. This further form of time, historicity, is given by the change in phase space, a change not addressed by existing physical-mathematical theories, we will return to this, and by the role of rare events in determining phylogenetic trajectories, a role unknown to physical theories, even those dealing with "large fluctuations" (Longo 2018a). I am aware that this language is foreign both to the biologists of the "genetic program" and to the best physicalist reductionism such as Baverstock's: we are a small community trying to advance in a theory *proper to* the organism and for us nothing is understandable in biology without an analysis of its historical, that is phylogenetic and ontogenetic timing. Main stream, genocentric, biology goes elsewhere, like the fantastic main

stream astronomy, trigonometry and geometry of epicycles in the Arabic world in year 1000 (Longo, Mossio 2020; Longo 2021a).

Second, let us return to the issue of "phase space". Kant beautifully framed the Newtonian practice in mathematical-physics by proposing an epistemic pre-existence of space and time to knowledge construction, as the locus of all possible trajectories, all already in potentia in space-time, or, also, he will see space and time as a condition of possibility of theorizing in physics – that is for writing equations and describing actual trajectories. The nineteenth century will extend this pre-condition of any theoretical description in physics to the "a priori" of the "space of phases", that is the choice of observables and parameters relevant to the theory intended – including space or time. A very productive theoretical boldness will lead physics to invent new spaces of phases for new theories, that is new observables, such as entropy, new pertinent parameters or to make variable the number of parameters (statistical physics), to propose spaces of infinite dimensions (Hilbert spaces in quantum mechanics), but each theory will have, and in principle, a pre-given space of phases. The symmetries proper to the description of these spaces make them always mathematically describable in the finite and a priori, even Hilbert spaces.

This is no longer valid in a historical science, such as biology. In biology, the historicity of the processes is due precisely to the dynamics of the space of the possible, a phase space, which changes along with the co-construction of phenotypes and organisms, their niches and ecosystems: for example, the observable "the placenta" was not already there, albeit in potentia, before a retroviral infection less than 100 mln years ago (Liavalle et al 2013). All forms of ex-aptation à la Gould propose new observables, if we reason at the relevant scale and biological observables, that of phenotypes and organisms (Gould 2002). Their reduction to the level of colliding molecules stops, we said, at least in front of the theoretical irreducibility of incompressible fluids to classical or quantum molecular dynamics. But there is much more. Biological relativity, (Noble et al. 2019), does not allow to privilege one causal level over others: the heart and the vascular system, in their example, are formed at a critical transition in which the interactions between different scales and levels of organization allow the formation of the organ and its function, at once. Indeed, the formation of an organ in embryogenesis, but even more so the establishment of an evolutionarily novel phenotype, feathers and, then, wings for flight in dinosaurs, requires an extended critical transition (Longo, Montévil 2014), i.e., a mathematically dense cascade of changes in symmetries, which depends on the entire ecosystem context. For feathers and functional wings, it ranges from muscular and lung structure, to... air density: the phenomenon depends on all these constraints of which the last one is a typical stable boundary condition, of physical type, the others are co-constituted at and by the evolutionary event and depend on the very formation of the possibility to fly, which, for vertebrates did not exist before – it is a new observable, depending on new observables and parameters, all of them. The new possible actually depends on the global interaction of all these constraints, some of which are due to the constitutive dynamics itself; mathematically, in my view, it depends on a *non-local variable*, that is on a parameter depending on the global network of interactions making the novelty possible, but not necessary.

This is what we would like to add to Noble's biological relativity, the non-locality of parameters or of causal dependence: at least one of the pertinent parameters that allows/govern the new observable (the heart in embryogenesis, wings in Evolution...) depends on the entire new global structure that did not exist before. As for Evolution, this is part of its historicity, which does not allow to give, a priori, a phase space. (By the way, it is possible that Cosmology, at least when dealing with the formation of novelty at the early stages after the Big Bang, such as the fundamental constants of physics, is facing a similar challenge, historicity, a major problem for physicists.)

To give a principle of optimum, instead, such as Maupertius, we know from Hamilton that it is necessary to give, a priori, a space with a partial order, in which we can speak of "extremal" values, minima or maxima, be they local or global. This is the pre-given phase space that the historicity of the living forbids. In our perspective, the analysis of the dynamics of this changing space is part of the theory of Evolution and, therefore, of the organism (similarly, no way to derive the values of the fundamental constants of physics from some "optimality principle", when these constants are not yet fixed – infinitely many values are possible, thus the "many Universes" hypothesis in Cosmology). Mathematically, some of us are trying to treat this aspect of historical dynamics in life sciences as "heterogenesis" (Sarti et al. 2019), well beyond the (very interesting) physical morphogenesis to which far from equilibrium thermodynamics belongs.

Let's put it differently, but still summarizing apodictically a lot of work: history, evolutionary history in particular, does not follow optimal trajectories, therefore specific, but possible or generic trajectories, thus not

necessary ones, in spaces co-constituted with and by the trajectories themselves. This process produces specific objects, that is, historical, individuated organisms in the case of the living. The physical trajectories, instead, are specific, geodesics in the appropriate space of phases, a space for each relevant theory, of course, but always given a priori. Physical objects are then generic, definable by a finite and a-temporal list of properties. The electron is a solution of the Dirac equation, just as a weight is perfectly described by some Galilean properties, without history. In contrast, a mouse is only definable in phylogenetic terms, telling its evolutionary story (Montévil 2020). As an example of an application of a method from physics, but not a theory, I dare to recall here the work in (Bailly, Longo 2009), where we proposed to fix a counting of some measurable features of phenotypes independently of their actual realization (number of foldings, fractal dimensions, e.g. in lungs, vascular systems..., tissue differentiations ...). The aim was to model Gould's analysis of "increasing complexity" in Evolution as a-symmetric random diffusion of "complexity" over bio-mass (Gould 1996), a new notion that we called anti-entropy. The "method" is borrowed from physics, but anti-entropy is a new observable, proper to life, to be added to entropy/negentropy, which are well defined in thermodynamics; its analysis does not assume extremal principles (the existence of maxima or minima), it thus departs from this omnipresent tool in mathematical physics. The production of anti-entropy, typically, requires the production of entropy (it does not oppose to it, like negentropy) and it is related to historical changes in the phase space (organisms' phenotypes and their pertinent parameters), a non-sense in existing theories in physics.

Let's follow the method, if inspiring, not the already given theories of physics: let's propose a theory adequate to the phenomenality of living, then we may try to correlate it, possibly as a non-conservative extension to the relevant physical theories (Longo 2020) - not easy because these theories are not unified. Non-conservative means that the extension of physics by biology can also allow us to explain, for example, the physical-chemical properties of the huge networks of molecular interactions that take place only in a living cell. These networks, from the physical point of view, have almost zero probability to exist and are found only inside cells. In short, there are macromolecular, thus physical-chemical activities, made possible only by biological contexts. That is, there is no plausible physical explanation of the origin and maintenance of huge networks of chemical interactions without considering the cellular structure in which they take place. These networks contribute to continually regenerate the cell, its membranes and its other organs, which, in turn, make the networks possible, *enable* them. This is a form of "constrained process" that is eminently biological and quite different from the ones given within boundary conditions treated by physics, since the constraints are produced by the dynamics itself, including the interaction networks they enable. In my opinion, the best treatment to date of this "closure of constraints" may be found in (Montévil, Mossio 2015). In that theoretical framework, the DNA itself is analyzed as an internal constraint of the cell, permanently reconstructed by the molecular activities that it contributes to structure and canalize, in the cell. Obviously, a change in this constraint, the DNA, will affect cellular and organismal dynamics, but does not "drive" them. No "book of life" but an amazingly important constraint, sitting within each cell, a core component of the process of protein formation, continually modified and repaired by the macromolecular dynamics within the cell – like vital tissues and organs, at their scale, are repaired in an organism. For example, double-strand breaks can be repaired by non-homologous or microhomological-mediated end joining, which may also introduce changes in the DNA. By this, the molecular networks in the cell, by their dimensions and materiality, their variations, affect the DNA itself, while pressures and torsions on the chromatin modify the opening sites and, thus, in addition to (de-)methylation, change what matters in these processes, the *function* of the DNA.

This role of the context in the functioning of the DNA, as a constraint, is outside the scope of the Central Dogma (CD) as well of its negation: it is not a matter of "information" going back from the proteins to the DNA, but of protein networks, of three-dimensional and material, physical-chemical and biological contexts that modify the DNA or what DNA *does* in the cell. In this, I disagree with both Baverstock and Noble on their critique of the validity of the CD. The CD, in the language of information, *implicitly* says or implies that the ontogenetic information and its hereditary transmission is *completely* contained in the DNA: at most the RNA may modify this information content by a retroaction on the DNA, not the proteins nor the epigenetic context, of course. I insist: the "completeness" assumption is crucial, while hidden in the language of information, as information transmission (heredity) and elaboration (development) define biological dynamics. Thus, the CD, as stated, in that language, is either insignificant, since the language of information (Turing, Shannon etc) is insignificant in these processes, as I claim, or it is true, since, clearly, there is no information sent back from proteins inside DNA. As René Thom writes, the limit of truth is not falsehood, but insignificance – possibly by

the construction of a language-frame that forces to miss what is significant. So, the language of information and programming, as a frame for the CD, has proposed a perspective that made insignificant for biology the many *material* processes that are essential components of how the DNA *works*, often referred to as “epigenetic”, for example pressures and torsions on the chromatin, RNA – proteome’s interactions and activities etc. Conversely, the CD itself is insignificant w.r. to that meaningful, material frame (including all epigenetic phenomena). By this, the CD has had a major impact on research: in a community using the language of information, it let too many disregard the physico-chemical (and biological) role of the proteome, the cell, the organism and the environment in determining biological dynamics, with serious ramifications for applied fields of biotechnology and genetic engineering in agro-food systems (GMOs are the direct children of the CD) and medicine, e.g. cancer research (Longo 2018).

Acknowledgments These views have been constructed thanks to the collaborations with F. Bailly, P.A. Miquel, M. Montévil, M. Mossio, C. Sonnenschein, A. Soto, A. Vianelli.

Minimal references (Longo’s papers and the references therein can be found in <https://www.di.ens.fr/users/longo/download.html>)

- Bailly, F., Longo, G. (2009). Biological Organization and Anti-Entropy. *J. Biological Systems*, Vol. 17, No. 1, pp. 63-96.
- Collins, F.S. (1999). Medical and societal consequences of the human genome project. *New England Journal of Medicine* 341, 28–37.
- Del Giudice E. (2007). Old and new views on the structure of matter and the special case of living matter. *Journal of Physics: conference Series*, 67, 012006.n..
- Gould, S.J. (1996). *Full House*. Three Rivers Press, New York.
- Gould S.-J. (2002). *The Structure of Evolutionary Theory*, Harvard U. Press.
- Gouyon PH, Henry JP and Arnoud, J. (2002) *Gene Avatars, The Neo-Darwinian Theory of Evolution*, Kluwer.
- Liang, F., Holt, I., Pertea, G. et al. Gene (2000) Index analysis of the human genome estimates approximately 120,000 genes. *Nat Genet* 25, 239-240 . <https://doi.org/10.1038/76126>
- Lavialle C, Cornelis G, Dupressoir A, Esnault C, Heidmann O, Vernochet C, Heidmann T. (2013) Paleovirology of ‘syncytins’, retroviral env genes exapted for a role in placentation. *Phil Trans R Soc B* 368: 20120507. <http://dx.doi.org/10.1098/rstb.2012.0507>
- Longo, G. (2016). A review-essay on reductionism: some reasons for reading "*Reductionism, Emergence and Levels of Reality. The Importance of Being Borderline*", a book by S. Chibbaro, L. Rondoni, A. Vulpiani. In *Urbanomic*, <https://www.urbanomic.com/document/on-the-borderline/> London, May 8.
- Longo, G. (2018). Information and Causality: Mathematical Reflections on Cancer Biology. In *Organisms. Journal of Biological Sciences*, vol 2, n.1.
- Longo, G. (2018a). How Future Depends on Past Histories and Rare Events in Systems of Life, “How future depends on past histories and rare events in systems of life”, *Foundations of Science* 23 (3):443-474 (2018)
- Longo, G. (2019). Information at the Threshold of Interpretation, Science as Human Construction of Sense. In Bertolaso, M. & Sterpetti, F. (Eds.) *Will Science Remain Human?* pp. 67-100, Springer, New York.
- Longo G. (2020). Naturalizing Physics. Or, embedding physics in the historicity and materiality of the living. *Deleuziana*, n. 11, special issue on “Differential Heterogenesis: Deleuze, Mathematics And The Creation Of Forms”.
- Longo G. (2021). Confusing biological rhythms and physical clocks. Today’s ecological relevance of Bergson-Einstein debate on time. In A. Campo, S. Gozzano (eds.), *Einstein vs Bergson. An enduring quarrel of time*, De Gruyter.
- Longo G. (2021a). Programming Evolution: a Crack in Science. *A Review of the book by Nobel winner, Jennifer A. Doudna, and Samuel H. Sternberg "A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution"* 2017, in *Organisms. J. Bio Sci.*, Vol. 5, No. 1.

- Longo, G. and Montévil M. (2014). *Perspectives on Organisms: Biological Time, Symmetries and Singularities*. Dordrecht: Springer.
- Longo, G., Mossio, M., 2020. Geocentrism vs genocentrism: theories without metaphors, metaphors without theories. *Interdisciplinary Science Reviews* 45, 380–405.
- Montévil, M. (2020) Historicity at the heart of biology. *Theory in Biosciences*, Springer Verlag.
- Montévil, M. & Mossio, M. (2015). Closure of constraints in biological organisation. *Journal of Theoretical Biology*, vol. 372: 179-191.
- Nicolis G., Prigogine I. (1977). *Self-organization in non-equilibrium systems*. New York, Wiley.
- Noble R., Tasaki K., Noble P., Noble D. (2019) Biological Relativity requires circular causality but not symmetry of causation: so, where, what and when are the boundaries? *Frontiers in Physiology* 10: 827.
- Paldi A. (2020). Stochastic or Deterministic? That is the Question. *Organisms. J. Bio Sci.*, vol. 4, N.1.
- Sarti, A., Citti G., Piotrowski, D (2019) Differential heterogenesis and the emergence of semiotic function. *Semiotica*. Vol 2019, Issue 230
- Soto A., Longo G., Noble D. (eds.) (2016) *From the century of the genome to the century of the organism: New theoretical approaches*, a Special issue of *Progress in Biophysics and Mol. Biology*, Vol. 122, 1, Elsevier.
- Thom R. (1991) *Predire n'est pas expliquer*, Eshel, Paris.

Cover Letter:

the following paper is an invited commentary, revised following the referees' advice

From Information to Physics to Biology

Giuseppe Longo

Commentary to "The gene: An appraisal" by Keith Baverstock,
PBMB, [Volume 164](#), September 2021, Pages 46-62.

Best regards

Giuseppe Longo

Centre Cavallès, CNRS et Ens, Paris

<http://www.di.ens.fr/users/longo>