Programming Evolution: a Crack in Science

Giuseppe Longo

Sapienza Università di Roma

Commentaries

Vol. 5, No. 1 (2021)
ISSN: 2532-5876

DOI: 10.13133/2532-5876/17538

Programming Evolution: A Crack in Science

Giuseppe Longo [a]*

[a] *CNRS – Ecole Normale Supérieure, Paris, France*

***Corresponding author:** Giuseppe Longo, giuseppe.longo@ens.fr

Abstract

Nobel Prize winner, Jennifer Doudna, and Samuel Sternberg survey recent advances in a pioneering area of molecular biology. In an accessible and elegant style, the authors present the successes and challenges of a new DNA-modifying technique: CRISPR. They transmit their emotions of discovery, passion for research, and intellectual audacity. While greatly admiring the technical skills of the authors, who are among the best researchers in the field, this review critically stresses the limits of their experimental practices, namely: a vague or incomplete theoretical frame; often unreachable genetic targets; off-target effects; prior failures to deliver by other forms of genetic manipulation, and, finally, the intrinsic unpredictability of many phenotypic consequences of such a powerful technique. Due to these concerns, the authors' approach to organisms and Evolution is questioned with the purpose to generate an open debate.

Keywords: CRISPR, gene-editing, genocentrism, off-target effects, Evolution, theory of organisms

Citation: Longo G 2021, "Programming Evolution: A Crack in Science", *Organisms: Journal of Biological Sciences*, vol. 5, no. 1, pp. 5-16. DOI: 10.13133/2532-5876/17538.

Commentary on: Jennifer A. Doudna & Samuel H. Sternberg 2017, *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution*. Boston, Ma: Houghton Mifflin Harcourt Publishers.

1. The Global Judgment: Vulgarization and Ethics

The book under review is a highly effective account of an extraordinary personal adventure in the invention and use of the latest genetic manipulation techniques. Despite having two authors, it was written in first person. This adds a personal touch to a highly readable style. In fact, one catches a glimpse into the passion of a selfless and very capable researcher immersed in a difficult world of biochemical techniques. One grasps moments of not only joyful success but also perplexing disappointment. In short, this book expresses a beautiful mind that is deeply dedicated to laboratory work. The author/narrator takes the reader, even an inexperienced one, by the hand on a difficult journey to "discover,"—or rather, invent the technical potential of biological mechanisms that are specific to the interaction between viruses and bacteria. This is then extended to the manipulation of DNA in eukaryotic cells. To this purpose, the book contains interesting information on viruses and bacteria, making it accessible to anyone. I will not further comment on the many fascinating details illustrated, for example, on how bacteria defend themselves from viruses, and how the chemical structures implied in this process can be reconstructed and used in the laboratories through insights and work.

The book also features the successes plus long lists of possible future applications of the manipulation made possible by the new DNA-editing techniques: "scientists

can now manipulate and rationally modify the genetic code that defines every species on the planet, including our own". Before discussing the proposed techno-scientific framework, let us move directly to the final part of the book. This addresses the ethical issues that relate to the potential of genetic manipulation in humans—especially the "improvement" of the species. Here, despite her enthusiasm for the techniques in which she contributed, the author stops short of the ethical challenge posed by such manipulations. With great humanity and intimate concern, the book presents the possible risks and abuses of such activities and proposes strict ethical limits to manipulations in humans. To this end, it leads us through the drama of the possible violence done to our species, and we can sense the peculiar sensitivity of a woman, the main author, faced with the manipulation of the embryonic genome of a future child.

2. Theoretical Problems

Having sincerely appreciated the book's merits in terms of writing, passion, and ethics, we now turn to a critique of its scientific content. Here, too, the authors' great intellectual honesty must be valued. Without hesitation, they take the Central Dogma of Molecular Biology (Crick 1958, p. 11) as a pillar of their theoretical framework. Today, this is often not the case. Even those who still and *de facto* base their work on it, especially in the laboratory practice of molecular biology, mostly refrain from mentioning it. If asked, they often present the Central Dogma as "a figure of speech" or a "simplification" of reality. Thus, we welcome a courageous and precise choice that does not leave us in vague, ill-defined theoretical frameworks. Of course, a problem arises: How is this dogma interpreted? Although not explicitly stated, there is no doubt that the book's interpretation refers to the harder version proposed by Watson in the 1960's. Such a version considers the DNA to contain the *complete* coding of genetic information, therefore, hereditary transmission. One cannot reproach the authors for a little vagueness in this respect since the notion of "(in-)completeness," which is clear and precise in mathematics, is unusual in the natural sciences. An exception was the 1935 seminal article by Einstein, Podolsky, and Rosen (known as the EPR paradox) which dealt with the "incompleteness" of quantum mechanics, providing a very rigorous and constructive critique of its foundations. Everything suggests that Doudna and Sternberg consistently consider DNA as *complete* in its prescriptive ontogenetic potential. Accordingly, the writing in the genes contains the complete set of instructions, it prescribes ontogenesis, and is at the core of phylogenesis.

However, a certain vagueness soon appears: the notions of "(genetic) information" and "program" are as ubiquitous as they are undefined. Since we are dealing with information encoded on discrete data bases (the chemical structure of DNA), we are led to believe that we are dealing with Shannon information (transmission) and/or Turing-Kolmogorov information (processing). As it is typical of biology, this lacks any precise reference to other notions of information. Let us not go into the diversity of the two notions here. For good reasons, these deal in a dual way with the relationship between the notions of entropy and complexity, therefore, of information that is usually seen as negentropy (Longo 2019). In fact, the lack of correlation between the "complexity" of an organism, however defined, and its DNA, does not seem to concern advocates of the genocentric approach. Although the authors consider DNA as a complete encoding of the organism, they recall that, for instance, the genome is hundreds of times larger in plants than in humans. Note that in 1999, the Director of the Human Genome Project, Francis Collins wrote that he expected to find 80,000 genes in man considering, not without pride, that the much less complex *C. elegans* (a microscopic worm of 1,000 cells) had 16,000 genes. Two years later, he recognized that there seemed to be 25,000 genes in man, or, as he later claimed along other authors, 21,000. The notion of a genetic program is even more vague. No attempt is made to identify the compiler, the interpreter, or the operating system. When an attempt was made by a few biologists using the most adequate language for string manipulation and term rewriting or "*term-editing*" (Church's lambda-calculus, which has been my specialty for long (Barendregt 1984; Kreisel 1982)), the use of recursion was still abusive (see Longo 2018; 2019 for a critique and sources). In sum, main stream molecular biology tends to fuzzily refer to precise notions such as information and program, while these notions are mathematically committed to a strong and specific form of "determination" (what and how determines what). This implicitly filters into views, experiments, and the interpretation of measurements.

This is rather inadequate for a text so rich in rigorous descriptions of viruses and bacteria, which aims at a global presentation, and calls for a clear definition of terms so liberally used in the discipline (including the foundational notion of gene).

In fact, what is a gene? In her book, *The Century of the Gene*, Evelyn Fox-Keller notes that our understanding of gene changed five times in the 20th century. In fact, the notion of gene is not defined in Doudna and Sternberg's book. However, the reader is lead to think that they consider it to be a segment of DNA to be associated not only with a protein but also a phenotype. This is at odds with their acknowledgement that some phenotypes are the result of a network of genetic expression, as it is the case for long-identified phenomena such as "alternative splicing" (Leff *et al.* 1986; see also Brett *et al.* 2001; Nilsen & Graveley 2010). These alternative initiations of transcription and translation (de Klerk & 't Hoen 2015) call for a revision of the "dogmatic" view of the correspondence of one mRNA to one protein in eukaryotes (Mouilleron *et al.* 2016). Such a further complexity goes beyond the concept of networks in the genotype-phenotype relationship (Brunet *et al.* 2018; 2020). A particularly telling example involves "overlapping genes." This phenomenon was discovered in the 1970's through the first-ever sequencing of a DNA genome (Barrell *et al.* 1976) and has been neglected since. Even now, some researchers (Schlub & Holmes 2020) consider it a typical feature of viruses, while many are starting to recognize it as a very relevant feature among the general category of "alternative proteins" in cellular organisms (Mouilleron *et al.* 2016; Brunet *et al.* 2018; Pavesi *et al.* 2018; Meydan *et al.* 2019). Overall, it is clear that these phenomena falsify the idea that genes are segments of DNA with a precise beginning and end, like software designed instructions. Indeed, the ENCODE project already highlighted "the complex patterns of dispersed regulation and pervasive transcription" and proposed to define a gene as "a union of genomic sequences encoding a coherent set of potentially overlapping functional products." Yet, the researchers involved are aware that their "definition sidesteps the complexities of regulation and transcription by removing the former altogether from the definition" (Gerstein *et al.* 2007).

In summary, the exact meaning of not only "information" and "program" but also "gene" is unclear. Oftentimes, the vagueness of these notions leave room for the attribution of extraordinary power to "genes." Everything is in the genetic information and elaborated by the genetic program. Both the program and the information are completely written in the genes. Of course, the authors point out that "in an individual, all the somatic cells have the same DNA." However, the contribution of the context in the control of gene expression is never referred to—perhaps because mentioning it would question the driving role of DNA in phenotype determination. Therefore, it is assumed that a very detailed program controls genetic expression in the DNA itself, from the zygote to the adult. This also means assuming that being human is written mostly in the 5,000 genes in excess of those of *C. elegans*, which causally contribute to each cell to take on very different forms and functions, from heart cells to, neurons and liver cells. The editing of this program would allow the organism to be completely steered in the ecosystem by the rational will of man, which is ethically acceptable and even necessary, according to the authors, at least in plants and animals.

A further theoretical gap in the book is the implicit use of another property that is essential to the proposed genocentric determinism: the exact stereo-specificity of macromolecular interactions and, therefore, of all the cascades from DNA to the proteins' functions to the phenotypes. Monod, in his 1970 book, Chance and Necessity, recognizes with great intellectual coherence that this property is "necessary for the transmission of information." Even more strongly, Monod claims that "the cell is a Cartesian mechanism," a clockwise chain of gears and pulleys. Macromolecular stereospecificity in a cell, as exact as the "Boolean algebras ... in our computers," says he, makes us understand how the processing and transmission of the genetic information contained in DNA may work. The first problem that arises from such a tenet is that physical chemistry has been treating interactions between macromolecules in a statistical way for long. Molecular interactions in a cell are no exception, as noted for genetic expression as early as 1983 by Kupiec (1983; 2010). Since then, the stochasticity of all steps of gene expression, from transcription to translation plus alternative splicing, has been extensively confirmed (see Elowitz 2002; Paldi 2003; Raj & Oudernardeen 2008; Waks *et al.* 2011 and more recently Boersma *et al.* 2019).

Generally speaking, macromolecular interactions are stochastic, they must be given in probabilities, and these

probabilities depend on the context. There are many references that justify this strong theoretical principle, but they are overlooked by the dominant genocentrism. In fact, the picture changes completely if one considers that, in this spirit, almost every "gene" is transcribed in *almost every* cell. Chelly *et al.* (1989) highlighted this long ago and this has been extensively confirmed since then: it is a matter of different probabilities (see also the references above on stochasticity). Moreover, twisting and pressing on chromatin changes the sites of DNA access, altering its expression (Cortini *et al.* 2016). This is certainly a crucial issue in embryogenesis, even though it hardly applies to computers. Similarly, many highlight "nongenetic cellular diversity" and "the role of regulatory network structure and molecular noise" (Balazsi *et al.* 2011). As stressed in (Braun 2015): "The genome does not determine the ordered cell state. Rather, it participates in this process by providing a set of constraints on the spectrum of regulatory modes, which are analogous to boundary conditions in physical dynamical systems." Clearly, this is a radical perspective shift from the genocentric approach: in this frame, the "boundary conditions" and their modifications, though still relevant for the dynamics, require a different kind of analysis. Typically, no single component of the dynamics has "completeness." Moreover, in physics, a difference in the boundary conditions may induce a difference in the dynamics or in its result. However, boundary conditions are analyzed differently from the "causes" of the dynamics itself. That is, these are clearly (mathematically) distinct from boundary conditions and are usually and beautifully framed in terms of conservation laws or symmetries, so that the notion of cause may be even avoided (a stone falls for *symmetry reasons* according to the theory of relativity).

In physics, though, the boundary conditions are supposed to be pre-set with respect to the intended process. In biology, instead, these "boundary conditions" are *co-constructed constraints*. They also depend on the constrained process that produces them: even the DNA, this fundamental, physico-chemical trace of history, undergoes a constant reconstruction. It is a massive *constraint* to the dynamics and the construction of macromolecules. It dynamically changes and differentially applies during ontogenesis, as well as, dramatically, in embryogenesis. More generally, the molecular, cellular, and organismal processes continually reconstruct membranes, microtubules, and other cellular components, as well as all the functional parts of the organism. These constitute constraints that contribute to the biological dynamics at all levels of organization. If so, they also affect the many macromolecular network that, though highly improbable from the point of view of physics, exist and work, but only in living cells, with a history. The original notion of a "closure of constraints" by Montévil & Mossio (2015) elegantly introduces the approach hinted here (see also (Deacon 2015)). Of course, modifying any of these constraints, especially one as important as DNA, leads to a change. However, this is because the change in the constraints turn out to re-channel the macromolecular processes, which, per se, are at least non-linear or, more generally, stochastic.

Of course, this analysis departs from Doudna and Sterner's determinism based on the genetic program, the Central Dogma, and the (unfortunately implicit) idea that macromolecular stereospecific interactions are exact. These theoretical assumptions are not simplifications for the sake of vulgarization. Rather, they are at the core of the book's perspective. These shaky foundation undermine the entire conceptual edifice of strict genocentrism, which has been presented to the reader as the only way of thinking. The different theoretical approach that we follow here, as proposed by many and discussed by Soto el al. (2016), offers another perspective when analyzing the evidence and the promises made in the book as for the role that CRISPR can play in "reprogramming" the living.

3. Theories versus Empirical Evidence

In science, as observed by Boltzmann, there is nothing more practical than a good theory. Can empirical evidence falsify the genocentric approach of the book? I think so, but this is not so obvious. Longo & Mossio (2020) present a close analogy between the genocentric view and the geocentric, Ptolemaic, perspective on the planetary system. In particular, the extraordinary progress in the knowledge of the skies due to the great Islamic astronomy and mathematics from the 8th to 14th centuries is acknowledged. The astronomers of Arabic language described all visible celestial bodies and their movements, especially the planetary system, from a geocentric perspective. No empirical evidence could falsify their account of the planets' movements since, mathematically, any finite number of points in

an ellipsis around the Sun can be interpolated by enough epicycles centered on the Earth. A change of perspective, actually, a metaphysical one, was required in order to consider the planets from the Sun's point of view. Only a dramatic change in theoretical principles could then falsify the geocentric perspective, such as the invention of the first fundamental conservation principle of physics, i.e., inertia, by Galileo. Then the "retrograde movements" of the planets, so closely described in Arabic, became totally impossible in the absence of masses in all the centers of epicycles, particularly after Newton's work. Note that inertia is a limit principle. It never applies in practice since visible movements are always constrained by gravitations and frictions. However, it allows us to understand all physical movements at once and analyze what constrains them: gravitations and frictions–since Galileo. In a sense, inertial movement is a "default" state of inert matter. Below, we will refer to a proper "default state" of living organisms, following Soto *et al*. (2016).

The relevance of the change of perspective and the invention of a "conservation principle" became clear when the new theoretical frame allowed for unifying falling apples and planetary movements (Newton, Hamilton), thus avoiding *ad hoc* descriptions and epicycles on top of epicycles. This recalls the *ad hoc* alphabetic writing in the zygote's DNA program that supposedly allows each cell to differentiate into a neuron or a leucocyte because genes control gene expressions, one on top of the other. In a context dominated by Monod, Jacob, and Lwof, the discovery of the epigenetic control of gene expression by Barbara McClintock has not been cited for 20 or more years (Fox-Keller 2003). Of course, some epicycles do exist, for example, the stationary satellites around the Moon or the satellites of planets with respect to the Sun.

In reference to the book under review, the authors further explain that the genes' alphabetic writing, with its complete control of ontogenesis, is "as editable as a simple piece of a text." Therefore, the fate of the embryo may be programmed at our rational will, at least for many traits. We can "imagine that the human genome is a large piece of software." As a reader of the book, I do understand the enthusiasm of a talented bio-chemist that suddenly sees, in the laboratory, the explosion of her combinatorial power over sequences of DNA bases. Yet, as a theoretician, I radically disagree with the loss of the sense of organismal life in a historical context that such a position transmits to the reader.

Is there empirical evidence confirming at least some actual achievements of the genocentric-programming perspective? Yes, and the authors provide long lists of results and much longer ones of future, potential applications. What is the problem then, at least with the results? Indeed, there are several.

First, like with the Islamic astronomers, some applications can work and may turn out to be very useful. We owe Ibn Yunus (Egypt, ca. 1000 A.D.) and many other great Islamic scientists for major advances in spherical trigonometry and the celestial observations that led to the Alfonsine Tables (Catholic Spain, 1483), which were successfully and widely used for navigation. However, generalizing their point of view and promises would be a major mistake, let alone their predictions entangled with astrology. Today, they are comparable to the ones in (Plomin 2019), where human behavior is also claimed to be written in a newborn's DNA.

Moreover, one should consider that observations and experiments in (molecular) biology suffer from the most severe irreproducibility crisis (Begley & Ionnidis 2014). As a matter of fact, biology is the theoretical place of diversity, variability, and historical specificity of organisms, which result from a phylo- and onto-genetic history. This means that one cannot (easily) generalize individual cases (or not in the same way as in physics, (Montévil, 2019)). As a discipline, molecular biology endures a high pressure to "publish or perish," which is disastrous for critical and time-intensive scientific insight and integrity (Longo 2014) and produces results with the shortest time validity (della Briotta *et al*. 2015).

Secondly, "measurement in biology is methodized by theory," as closely analyzed in (Montévil 2019). The fuzzy theoretical background of information and programming contributes to make results and data too often unreliable or uncertain when it comes to interpretation. Before Newton's theory, astronomers had experienced major problems with data on planets' Keplerian orbits whose irregularities were due to planetary gravitational interactions. Until Einstein's theory, measurements of the perihelion of Mercury were unintelligible. Data do not speak by themselves, even less very big sets of data, as they necessarily contain lots of spurious correlations (Calude & Longo 2017). As for our object of study, the "re-writing" of DNA may not only achieve its goal and modify the intended

phenotype but also scramble other parts of the DNA and thus affect the organism. There may be more than the expected changes induced by CRISPR in the DNA. Different genetic changes may be due to the diversity of nucleotide modifications in the target sequence, as well as a varying spectrum of sites that have been changed. Since unwanted effects could arise from both the target and off-target sites, the detection and measurement of unintentional or off-target changes may be much more difficult than that of changes at target sites. In fact, it turns out that this is the case (Chaudhari *et al.* 2020; Höijer *et al.* 2020, Modrzejewski *et al.* 2020) because the number and location of nucleotide changes are unknown, particularly if they occur with lower but non-zero probabilities in non-specific sites. Moreover, the changes may not depend on the nucleic acid sequence modified. Rather, they may depend on the scale of the induced modification (e.g. the level of the organism or the ecosystem), as well as on its (temporary or permanent) timing and duration (Adikusuma *et al.* 2018). Information theories of macromolecular exact editing of alphabetic codes do not allow to see these phenomena nor to interpret them.

Critical observations increase with time, including remarks on low efficiency of mutation repair, high rates of mosaicism, and the possibility of unintended editing outcomes that may have pathologic consequences (National Academies of Sciences, 2020; Alanis-Lobato *et al.* 2021 and references 10-14 therein). Recently, Leibowitz *et al.* (2021) have shown that "CRISPR–Cas9 editing generates structural defects of the nucleus, micronuclei and chromosome bridges, which initiate a mutational process called chromothripsis. Chromothripsis is an extensive chromosome rearrangement restricted to one or a few chromosomes that can cause human congenital disease and cancer. These results demonstrate that chromothripsis is a previously unappreciated on-target consequence of CRISPR–Cas9-generated DSBs."

We are far from the authors' claim of "the remarkable ability to rewrite the code of life with surgical precision and astonishing simplicity." Indeed, the techniques invented by the authors and their collaborators modify the DNA, can guide the production of a specific functional molecule, and induce, among others, a "gain-of-function" at the cellular level. However, their off-target or unappreciated on-target effects, and their entangled, non-compositional consequences over the different levels of an organism's organization—which are embedded in an ecosystem—are far from under control. We can heavily affect Evolution, not control it. In fact, we may succeed in modifying a few constraints to complex processes, but we never achieve the full control of them. We can act on nature, but cautiously. At least from now on, we should only do so based on robust practices and good theories—not vague, metaphorical conceptual frames for life.

In short, the CRISPR technology does modify the DNA, but where, and with what consequences over time? The belief that we can precisely cut macromolecular interactions is a delusion belonging to the myth of the cell as a "Cartesian mechanism" with computers and software replacing Descartes' clocks. Therefore, the key issue involves shifting from a genocentric perspective to a vision centered on the organism in its relation to the ecosystem, where the DNA represents a fundamental internal and historical constraint, in the sense of (Montévil & Mossio 2015). I believe and hope that the remarkable technical invention of CRISPR may be used in a sound way for knowledge and therapies, at least for rare monogenetic diseases. Most pathologies, however, even where DNA plays a key role, are due to the deformation of a wide network of gene expressions and molecular activities that interact within an organismal and ecosystemic context.

4. Previous Cases

The exuberant expectations of CRISPR has major precedents in the prevailing genocentric view. Revisiting a few of them may help in understanding the limits of today's promises. Based on my indirect personal experience, I will refer to cancer gene therapies. These have been expected for about a century and promised for at least 50 years as the age of the Somatic Mutation Theory of cancer (SMT). Such a frame refers to cancer as an entirely genetic problem and explicitly counts on CRISPR to solve it.

Since 1971, generously funded projects have heralded the final victory against cancer thanks to genetic therapies that can "reprogram" the "deprogrammed DNA." The former U.S. President Richard Nixon's "war on cancer" aimed to provide these therapies by 1976, the bicentenary of the American Revolution. By the year 2000, the major technological achievement of "decoding" the human genome was

seen as a further tool to solve the cancer puzzle and, once again, allow genetic therapies. Hanahan & Weinberg (2000), with over 20,000 quotations in a few years, and many other authors, promised genetic therapies for "eliminating suffering and death due to cancer by 2015," as the then Director of the National Cancer Institute, Andrew von Eschenbach (2003) put it. Indeed, within a few years, DNA analyses should have led to diagnosis and prognosis.

Many of us, unfortunately, have had a direct or indirect experience of this life threatening disease. Therefore, we know that in 2021 only the histologist at the light microscope can recognize if a cancer is primary, metastatic, benign, or malignant. Moreover, no plausible gene-based cancer therapy exists (see Baker 2014; Huang 2014; Maeda & Katami 2018). Eventually, Weinberg (2014), in a severe self-critique of his previous approach (see the 2000 paper above with Hanahan), acknowledges that "Genome sequencing also came of age and documented myriad mutations afflicting individual cancer cell genomes." Moreover, "63 to 69% of all somatic mutations [are] not detectable across every tumor region… Gene-expression signatures of good and poor prognosis were detected in different regions of the same tumor" (Gerlinger *et al.* 2012). "Sequencing has revealed that healthy cells in all tissues bear heavy mutational burdens and that mutations are not exceptional, but normal" (Mustjoki, Young 2021). Versteg (2015) also mentions tumors without mutations, while Gatenby (2017) observes that "cancer cells can display a seemingly paradoxical state in which their mutational burden is similar to and perhaps even lower than that of adjacent normal cells." On this basis, Gatenby hypothesizes that the tissue and the organismal environment drive the process, following (Sonnenschein & Soto 1999). Moreover, as Weinberg (2014) dares to admit, "most human carcinogens are actually not mutagenic." Forty years of contradictory analyses on asbestos (Huang *et al.* 2011), plus the aforementioned evidence, opened to the idea that, when the frequent and heavy mutational burden in cancer occurs, it is mostly a *consequence* rather than a cause of the disruption in cell control of reproduction (see also (Mally, Chipman 2002)). This phenomenon brings a specific diversity and leads to looking at cancer as a systemic problem (Bizzarri 2014; Baker 2021). Finally, in view of the mutational confusion in cancer, (Weinberg 2014) refers to it as "infinite complexity", thus some now bet on Big Data for machines to mend the human failure in understanding cancer's etiology. Unfortunately, mathematics shows that this is nonsense (Calude & Longo 2017; Montévil & Longo 2018). Despite the failure to deliver, too many—mostly avoiding any explicit reference to the central dogma or even denying its role in private conversations—continue to research or fund research *only* on cancer causing mutations, oncogenes, proto-oncogenes, or onco-suppressors (Kato *et al.* 2016; Rohan *et al.* 2018).

With a more robust organismal perspective, the Tissue Organization Field Theory (TOFT) (Sonnenschein & Soto 1999) allows us to understand why mutated cells from a cancer tissue may functionally normalize when transferred in a healthy tissue. For example, cells from a mammary neoplasm relocated in a healthy mammary gland stroma, functionally normalize (Maffini *et al.* 2005; Soto & Sonnenschein 2011). TOFT focuses on the failure of the triangular relation tissue/organism/environment in cancer formation. It also highlights the role of endocrine disruptors and other ecosystemic causes that affect the tissular and organismal control of somatic cell reproduction. Instead, the totalizing focus on DNA when studying and curing cancer keeps diverting attention and research from environmental causes, which are rarely mentioned by the tenants of the SMT. In this sense, the environment is not mentioned once in this book, despite about one hundred references to "cancer." As a matter of fact, the search for a genetic "magic bullet" has financially dominated for 50 years. This has largely excluded other research paths and minimized environmental analyses.

5. Remarks on the Method

Some may observe that I mentioned the frequent unreliability or irreproducibility of experimental results in the perspective I critique, while I attributed more validity to empirical evidence that aligns with my point of view. This depends on explicit theoretical analyses. Namely, I have stressed in several writings, often in collaboration with biologists, the inconsistency or incompleteness of genocentric determinism. These theoretical gaps result from vague or inconsistent notions of the gene, the information, and the program (see Longo 2019 for a synthesis on the misuse of "information" and "program"), as well as their implicit causality or determinism. Notwithstanding,

experiments are designed on the base of these vague or implicit notions and their strong consequences. These include the idea (deemed "necessary" by Monod 1970) of exact macromolecular interactions at the core of huge macromolecular networks. These networks would be designed like electronic circuits and would elaborate "Boolean algebras" (and this is even not meant to be "on average"). Such ideas are spread throughout molecular biology university textbooks and shape minds forever. This has led me to raise more a priori doubts on both the experiments and the measurements carried out in the information-genocentric framework. In fact, as stressed by Einstein in physics, theory decides the observables and the pertinent parameters. It proposes measurement tools and methods, as well as interpretations of data, as mentioned above in relation to the planets' orbits: vague or inconsistent theories undermine measurements, methods and interpretations.

Now, biology suffers even more from these biases because the historical and contextual specificity of organisms requires both diachronic and synchronic knowledge and measurement, as mentioned above—see also Longo 2017; Montévil 2019, and Montévil & Mossio 2020. Accordingly, a more explicit, well-defined, and robust theoretical frame justifies a greater reliance on empirical results. For example, despite branching into at least two different approaches (Gould 2002), Evolutionary Theories now make fantastic use of DNA fingerprints in paleontology. Oftentimes, this is done in mitochondrial DNA, which allows for reconstructing phylogenetic paths in theoretically well-construed perspectives. In the case of cancer, after 50 years of failed SMT-based promises of genetic therapies, TOFT has been explicitly based on Darwin's first principle (heredity as "descent with modification"), interpreted as a "default state" ("reproduction with variation") for all sufficiently fed organisms, and applied to somatic cells under massive differential constraints (constraining reproduction as well as motility in varying ways, according to the context). This seems more convincing than SMT principles, independently of the empirical failures of the latter. In fact, SMT implicitly refers to the Central Dogma and its set of biologically fuzzy notions of information and program. TOFT refers to Darwin and, today, to an increasingly robust theory of a "closure of constraints" in biology. Its theoretical frame no longer depends on Shannon's nor Turing, Church, and Gödel's information or programming theories (see Longo 2018 for a critique of the "Gödelitis" affecting some biologists). TOFT provides a relevant understanding of endocrine disruptors as carcinogens (Sweeney *et al.* 2015; Paulose *et al.* 2015) and prevention tools, thus opening to new therapeutic paradigms (Baker 2014; Bizzarri *et al.* 2014; Proietti *et al.* 2019), such as tumor reversion.

Second, I consider "negative results" particularly interesting in science since they have always opened the way to new paths of knowledge building (Longo 2018). At the theoretical level, randomness, in particular, is subtly related to undecidability, if understood as unpredictability in the intended theory (Calude & Longo 2016). If well defined, it thus provides a precise limit to knowledge. Now, the construction of undecidability is the "negative" result, which is the origin and pillar of the theory of computability or "elaboration of information" (Gödel, Church & Turing in the 1930's), so often cited in mainstream molecular biology. Of note is that in biology, randomness is not "noise" (Bravi & Longo 2015; Calude & Longo 2016). Rather, it is an essential component of the production of variability and diversity, and therefore, of the adaptivity and stability of organisms and ecosystems (a typical "information-theoretic" bug in biology is that it cannot distinguish randomness from noise–except by the notion of "incompressible sequence", a nonsense in biology). In other words, if one can "do something" or understand more through an insight into the limitations of knowledge, such as unpredictability (randomness), then I view this as a major theoretical advancement. Provable limits and constraints require precise definitions and structure theories and objects of knowledge. I Insist, the world-changing notions of programs and computation were defined in the 1930's to demonstrate incomputability. This involved clarifying the limits of knowledge and praxis instead of claiming the theoretical completeness of the analysis of this or that component concurring to a process. Such a method is thus fundamental in reinforcing the knowledge frame and in opening to new theories and applications. For these reasons, acknowledging the stochasticity of genetic expression and macromolecular interactions, channeled by biological constraints, is a convincing methodological pathway. Given the huge enthalpic oscillations of (not crystallized) macromolecules in a cell at a viable temperature, it is also empirically convincing. Yet it is also theoretically more robust than the vague

theories that envision the programmable genetic information to fully determine biological processes up to scattered noise.

This perspective shift suggests fundamental dualities. For example, the physical, highly improbable molecular networks in a cell do not completely determine bottom-up cellular activities and components. Rather, they are enabled by the very cellular constraints that they produce (Montévil & Mossio 2015). Indeed, there is no spontaneous generation from molecules to life, except for the totally unknown "singularity" at the origin of life. Existing and even artificial life is the result of a history, where each phylogenetic trajectory is triggered by rare events (Longo 2017). Accordingly, we better focus on how to understand and act on constraints, including the most fundamental one: DNA. This way, we can *canalize* processes by modifying constraints of various nature. In reference to the previous discussion, a typical example is "tumor reversion" (Bizzarri *et al.* 2014; Proietti *et al.* 2019; Kuchling *et al.* 2020; Sonnenschein & Soto 2020). Such a totally different approach contrasts decades of claims and failures about "rewriting tumor's scrambled genetic program." Furthermore, I think that this approach may shed a light also on our relationship with the ecosystem: we mostly acted and act on it by modifying constraints to its processes–with the effectiveness and the limits in understanding and prediction that are proper to this kind of actions.

As for theorizing, Weyl (1949) points out that the main methodological teaching of the theory of relativity, beginning with Galileo's relativity, is about moving from the "subjective-absolute" (so similar to the geocentric and genocentric approaches) to the "relative-objective" perspective. The construction of scientific objectivity requires analyzing the invariants, i.e. what is stable with respect to transformations of reference systems. In biology this should mean stability with respect to a "relativization" of levels of organization and scales, for integrating them. While considering DNA an amazingly important internal constraint to cellular dynamics, we must be able to move from the point of view of DNA to the organismal and ecosystemic perspectives and vice versa. Then, we must understand their integration and respective roles in the structure of biological determination (Noble *et al.* 2019).

As stated at the beginning of this note, I greatly appreciated the book for making some theoretical principles explicit. I also criticized it for leaving others implicit. Despite my admiration for the authors' experimental talent and insights, I wanted to express my disagreement with the framework of biological thinking they propose. Should Ibn Yunus (Egypt, ca. 1000 A.D.) be awarded the Nobel Prize for his contribution to astronomy? Definitely yes, despite the shortcomings of his theoretical vision. However, I think that now we should further investigate the practical and theoretical relevance of the analogue of Galileo's asymptotic principle of inertia in organismal biology, the default state of "reproduction with variation," an application of Darwin's first principle of Evolution, "descent with modification," which Darwin considered pervasive in all species (and that he discussed at length in four out of the first six chapters of *On the Origin of Species*). Note that somatic cells' "reproduction with variation" in a (healthy) tissue is a limit state, like inertial movement in physics. This is because reproduction in somatic cells is *always* (yet differently) constrained. By posing this Darwinian principle for all cells, including somatic cells, one follows in the footsteps of 150 years of microbiology and can better understand what constrains them within an organism, as well as the failure of these constraints in controlling cell reproduction, as it seems to mostly happen in the case of cancer (Soto & Sonnenschein 2011). This principle should combine with the unifying vision of organisms as a "closure of constraints," applied to all levels of organization. Both require scientists to specify the constraints of the largely brownian or chaotic molecular dynamics, as well as cells' reproduction and motility, i.e., their functional activities in an organism (Montévil & Mossio 2015; Soto *et al.* 2016; Bizzarri *et al.* 2020).

I believe that organismal biology will achieve further relevant results. The knowledge and techniques generated by the authors' and many others' work on CRISPR has contributed and may further contribute to this. A very interesting example has already been provided by fundamental studies, where "the CRISPR-based studies have surprisingly revealed that… effects on gene expression that are not mediated by the RNA transcript itself … occur in many loci that produce lncRNAs as well as in many loci that encode mRNAs" (Engreitz *et al.* 2016; Engreitz *et al.* 2019, p. 237). Following also the work in Cortini *et al.* (2016), Ramdas & Shivashankar (2015), and others, this confirms that the physico-chemical and context-dependent actions, including the structure of (long non-coding) lncRNAs,

may have a key regulating role, well beyond the genocentric informational approach. Understanding by both robust theories, instead of vague "metaphors", and by their experimental counterpart, while framing also the remarkable results obtained by the authors, should be an essential component of science, well before acting on nature.

Acknowledgments

I am very grateful to Alberto Vianelli and Andras Paldi for many helpful references and suggestions.

References

Adikusuma F, Piltz S, Corbett M A *et al.* 2018, "Large deletions induced by Cas9 cleavage", *Nature* vol. 560, E8-E9. https://doi.org/10.1038/s41586-018-0380-z.

Alanis-Lobato G *et al.* 2021, "Frequent loss-of-heterozygosity in CRISPR-Cas9–edited early human embryos", *Pnas*, vol. 118, no. 22, art. no. e2004832117 https://www.pnas.org/content/early/2021/04/09/2004832117.

Balazsi G, Oudenardee A, Collins J 2011, "Cellular decision making and biological noise: From microbes to mammals", *Cell* vol. 144.

Barendregt H 1984, *The Lambda-Calculus: Its Syntax, its Semantics*. Amsterdam: North-Holland.

Begley C & Ioannidis J 2014, "Reproducibility in science: Improving the standard for basic and preclinical research", *Circulation Research*, vol. 116, no. 1, pp. 116-126. doi:10.1161/CIRCRESAHA.114.303819.

Baker S G 2014, "Recognizing paradigm instability in theories of carcinogenesis", *British Journal of Medicine and Medical Research*, vol. 4, no. 5, pp. 1149-1163.

Baker S G 2021, "The case for a cancer paradox initiative", *Carcinogenesis*, https://doi.org/10.1093/carcin/bgab052.

Barrell B G, Air G M, & Hutchison C A III 1976, "Overlapping genes in bacteriophage fX174", *Nature* vol. 264, pp. 34-41.

Bizzarri M 2014, "System biology for understanding cancer biology", *Curr Synthetic Sys Biol*, vol. 2, no. 1.

Bizzarri M & Cucina A 2014, "Tumor and the micro-environment: a chance to reframe the paradigm of carcinogenesis", *Biomed Res Intl.* 934038.

Bizzarri M, Giuliani A, Minini M, Monti N, & Cucina A 2020, "Constraints shape cell function and morphology by canalizing the developmental path along the Waddington's Landscape", *Bioessays*. vol. 42, no. 4, art no. e1900108.

Boersma S, Khuperkar D, Verhagen B M P, Sonneveld S, Grimm J B, Lavis L D, Tanenbaum M E 2019, "Multi-color single-molecule imaging uncovers extensive heterogeneity in mRNA decoding", *Cell* vol. 178, pp. 458-472.

Braun E 2015, "The unforeseen challenge: From genotype-to-phenotype in cell populations", *Rep. Prog. Phys.* vol. 78, 036602.

Bravi B & Longo G 2015, "The unconventionality of nature: Biology, from noise to functional randomness". In: *Unconventional Computation and Natural Computation*, pp 3-34, C S Calude & M J Dinneen (eds.), Springer.

Brett D, Pospisil H, Valcárcel J, Reich L, & Bork P 2001, "Alternative splicing and genome complexity", *Nature Genetics*, vol. 30, no. 1, pp. 29-30.

della Briotta Parolo P, Kumar Pan R, Ghosh R, Huberman B A, Kaski K, & Fortunato S 2015, "Attention decay in science", *Journal of Informetrics* vol. 9, pp. 734-745.

Brunet M, Levesque S A, Hunting D J, Cohen A A , & Roucou X 2018, "Recognition of the polycistronic nature of human genes is critical to understanding the genotype-phenotype relationship", *Genome Res* vol. 28, pp. 609-624.

Brunet M, Leblanc S, & Roucou X 2020, "Reconsidering proteomic diversity with functional investigation of small ORFs and alternative ORFs", *Exp Cell Res* vol. 393, no. 1, 112057.

Calude C & Longo G 2016, "Classical, quantum and biological randomness as relative unpredictability", *Natural Computing*, vol. 15, no. 2, pp. 263–278.

Calude C & Longo G 2017, "The deluge of spurious correlations in big data", *Foundations of Science* vol. 22, pp. 595–612.

Casali M & Merlin F 2020, "Rethinking the role of chance in the explanation of cell differentiation". In: *Phenotypic switching*, pp. 23-51, H Levine, M Jolly, P Kulkarni, & V Nanjundiah (eds), Elsevier.

Chaudari H G, Penterman J, Whitton H J, Spencer S J, Flanagan N, Lei Zhang M C, Huang E, Khedkar A S, Toomey J M, Shearer C A, Needham A W, Ho T W, Kulman J D, Cradick T J, & Kernitsky A 2020, "Evaluation of homology-independent CRISPR-Cas9 off-target assessment methods", *The CRISPR Journal*, vol. 3, pp. 440-453.

Chelly J, Concordet J P, Kaplan J C, & Kahn A 1989, "Illegitimate transcription: transcription of any gene in any cell type", *PNAS* vol. 86 no. 8, pp. 2617-2621.

Collins F S 1999, "Medical and societal consequences of the human genome project", *New England Journal of Medicine* vol. 341, pp. 28–37.

Cortini R, Barbi M, Caré B, Lavelle C, Lesne A, Mozziconacci J, & Victor J M 2016, "The physics of epigenetics", *Reviews of Modern Physics* vol. 88, art. no. 025002.

Deacon T 2015, "Steps to a science of biosemiotics", *Green Letters: Studies in Ecocriticism* vol. 19, no. 3, pp. 293–311.

Elowitz M B, Levine A J, Siggia E D, & Swain P S 2002, "Stochastic gene expression in a single cell", *Science* vol. 297, pp. 1183-1186.

Engreitz J, Abudayyeh O, Gootenberg J, & Zhang F 2019, "CRISPR Tools for Systematic Studies of RNA Regulation".

In: *RNA Worlds. New Tools for Deep Exploration*, pp. 1-23, T R Cech & J Steitz, CSHL Press.

Engreitz J M, Haines J M, Perez E M, Munson G, Chen J, Kane M, McDonel P E, Guttman M, & Lander E S 2016), "Local regulation of gene expression by lncRNA promoters, transcription and splicing", *Nature* vol 539.

von Eschenbach A 2003, "NCI sets goal of eliminating suffering and death due to cancer by 2015", *J Natl Med Assoc.* vol. 95, pp. 637-639.

Fox-Keller E 2003, *A Feeling for the Organism Life and Work of Barbara McClintock*, New York: Freeman.

Gatenby R A 2017, "Is the genetic paradigm of cancer complete?", *Radiology*, vol. 284, pp. 1–3.

Gerlinger M *et al.* 2012, "Intratumor heterogeneity and branched evolution revealed by multiregion sequencing", *Engl J Med* vol. 366.

Gerstein M *et al.* 2007, "What is a gene, post-ENCODE? History and updated definition", *Genome Research*, vol. 17, pp. 669-681.

Gould S-J 2002, *The Structure of Evolutionary Theory*, Harvard: Harvard University Press.

Hanahan D & Weinberg R A 2000, "The hallmarks of cancer", *Cell* vol. 100, 57–70.

Höijer I, Johansson J, Gudmundsson S, Chin C-S, Bunikis I, Häggqvist S, Emmanouilidou A, Wilbe M, den Hoed M, Bondeson M-L, Feuk L, Gyllensten U, & Ameur A 2020, "Amplification-free long-read sequencing reveals unforeseen CRISPR-Cas9 off-target activity", *Genome Biol* vol. 21, p. 290.

Honegger K & de Bivort B 2018, "Stochasticity, individuality and behavior", *Curr Biol* vol. 28, no. 1, pp. R1-R16.

Huang S 2014, "The war on cancer: lessons from the war on terror", *Front Oncol.* vol. 4, p. 293.

Huang S, Jaurand M C, Kamp D, Whysner J, & Heil T 2011, "Role of mutagenicity in asbestos fiber-induced carcinogenicity and other diseases", *J Toxicol Environ Health B Crit Rev*. vol. 14, no. 1-4, pp. 179–245.

Kato S, Lippman S M, Flaherty K T, & Kurzrock R 2016, "The conundrum of genetic 'Drivers' in benign conditions", *J Natl Cancer Inst* vol. 108, no. 8, art. no. djw036. PMID: 27059373; PMCID: PMC5017937.

de Klerk E, 't Hoen P A 2015, "Alternative mRNA transcription, processing, and translation: insights from RNA sequencing", *Trends in Genetics* vol. 31, pp. 128-139.

Kuchling F, Friston K, Georgiev G, & Levin M 2020, "Morphogenesis as Bayesian inference: A variational approach to pattern formation and control in complex biological systems", *Phys Life Rev* vol. 33, pp. 88-108. doi: 10.1016/j.plrev.2019.06.001. Epub 2019Jun 12. PMID: 31320316.

Kupiec J-J 1983, "A probabilistic theory for cell differentiation, embryonic mortality and DNA C-value paradox", *Specul Sci Technol* vol. 6, pp. 471-478 (reprinted in Organisms, n. 2, 2020).

Kupiec J-J 2010, "On the lack of specificity of proteins and its consequences for a theory of biological organization", *Prog Biophys Mol Biol* vol. 102, pp. 45-52.

Kreisel G 1982, *Four letters to G. L.* http://www.di.ens.fr/users/longo/files/FourLettersKreisel.pdf.

Leff S E, Rosenfeld M G, & Evans R M 1986, "Complex transcriptional units: Diversity in gene expression by alternative RNA processing", *Ann Rev Biochem* vol. 55, pp. 1091-117.

Leibowitz M L *et al.* 2021, "Chromothripsis as an on-target consequence of CRISPR–Cas9 genome editing" *Nature Genetics* vol. 53, pp. 895–905.

Longo G 2014, "Science, Problem solving and bibliometrics". In: *Bibliometrics: Use and Abuse in the Review of Research Performance*, W Blockmans *et al.* (eds), Portland Press.

Longo G 2017, "How future depends on past histories and rare events in systems of life", Foundations of Science, pp. 1-32.

Longo G 2018, "Interfaces of Incompleteness". In: *Systemics of Incompleteness and Quasi-systems*, G Minati, M Abram, & E Pessa (eds.). New York: Springer.

Longo G 2018t, "Letter to Alan Turing", *Theory, Culture and Society, Special Issue on Transversal Posthumanities*. vol. 36, no. 1, pp. 73-94.

Longo G 2019, Information at the threshold of interpretation, science as human construction of sense". In: *Will Science Remain Human?*, pp. 67-100, M Bertolaso & F Sterpetti (eds.). New York: Springer.

Longo G & Montévil M 2014, P*erspectives 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*, vol. 45, no. 3, pp. 380-405.

Maffini M V, Calabro J M, Soto A M, & Sonnenschein C 2005), "Stromal regulation of neoplastic development: Age-dependent normalization of neoplastic mammary cells by mammary stroma", *Am J Pathol* vol. 167, pp. 1405-1410.

Maeda H & Khatami M 2018, "Analyses of repeated failures in cancer therapy for solid tumors: Poor tumor-selective drug delivery, low therapeutic efficacy and unsustainable costs", *Clin Transl Med* vol. 7, no. 1, art no. 11.

Meydan S, Marks J, Klepacki D, Sharma V, Baranov P V, Firth A E, Margus T, Kefi A, Vazquez-Laslop N, & Mankin A S 2019, "Retapamulin-assisted ribosome profiling reveals the alternative bacterial proteome", *Molecular Cell* vol. 74, no. 3, pp. 481-493.e6

Mally A, Chipman J K 2002, "Non-genotoxic carcinogens: Early effects on gap junctions, cell proliferation and apoptosis in the rat", *Toxicology*, vol. 180, no. 3, pp. 233-48. doi: 10.1016/s0300-483x(02)00393-1. PMID: 12393293.

Montévil M 2019, "Measurement in biology is methodized by theory", *Biology & Philosophy*, vol. 34, no. 3, art. no. 35.

Montévil M & Longo G 2018, "Big Data and biological knowledge". In: *Predictability and the Unpredictable. Life, Evolution and Behaviour*, D Ceccarelli & G Frezza (eds). Rome: CNR Edizioni.

Montévil M & Mossio M 2015, "Closure of constraints in biological organisation", *Journal of Theoretical Biology*, vol. 372, pp. 179-191.

Montévil M & Mossio M 2020, "The identity of organisms in scientific practice: Integrating historical and relational conceptions", *Front Physiol* vol. 11, art. no. 611.

Modrzejewski D, Hartung F, Lehnert H, Sprink T, Kohl C, Keilwagen J, & Wilhelm R 2020, "Which factors affect the occurrence of off-target effects caused by the use of CRISPR/Cas: A systematic review in plants", *Front Plant Sci* , vol. 11, art. no. 574959.

Mouilleron H, Delcourt V, & Roucou X 2016, "Death of a dogma: eukaryotic mRNAs can code for more than one protein", *Nucl Ac Res* vol. 44, pp. 14-23.

Mustjoki S & Young N 2021, "Somatic mutations in 'benign' disease", *N Engl J Med* vol. 384, pp. 2039-2052.

National Academy of Sciences 2020, *Heritable Human Genome Editing*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25665.

Nilsen T W & Graveley B R 2010, "Expansion of the eukaryotic proteome by alternative splicing", *Nature* vol. 463, pp. 457-463.

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* vol. 10, art. no. 827.

Paulose T, Speroni L, Sonnenschein C, & Soto A M 2015, "Estrogens in the wrong place at the wrong time: fetal BPA exposure and mammary cancer", *Reprod Toxicol* vol. 54, pp. 58-65.

Pavesi A, Vianelli A, Chirico N, Bao Y, Blinkova O, Belshaw R, Firth A, & Karlin D 2018, "Overlapping genes and the proteins they encode differ significantly in their sequence composition from non-overlapping genes", *PloS one* vol. 13, art. no. e0202513.

Paldi A 2003, "Stochastic gene expression during cell differentiation: order from disorder?" *Cell Mol Life Sci*, vol. 60, pp. 1775-1779.

Plomin R 2019, *Blueprint: How DNA Makes Us Who We Are*. Boston: MIT Press.

Proietti S, Cucina A, Pensotti A, Biava P M, Minini M, Monti N, Catizone A, Ricci G, Leonetti E, Harrath A H, Alwasel S H, & Bizzarri M 2019, "Active fraction from embryo fish extracts induces reversion of the malignant invasive phenotype in breast cancer through down-regulation of TCTP and modulation of E-cadherin/β-catenin pathway", *Int J Mol Sci* vol. 20, no. 9, art. no. 2151. PMID: 31052313; PMCID: PMC6539734.

Raj A & van Oudernardeen A 2008, "Stochastic gene expression and its consequences" *Cell* vol. 135, pp. 216-226.

Rohan T E, Miller C A, Li T, Wang Y, Loudig O, Ginsberg M, Glass A, & Mardis E 2018, "Somatic mutations in benign breast disease tissue and risk of subsequent invasive breast cancer", *Br J Cancer* vol. 118, no. 12, pp. 1662-1664. Epub 2018 Jun 6. PMID: 29872146; PMCID: PMC6008400.

Ramdas N M & Shivashankar G V 2015, "Cytoskeletal control of nuclear morphology and chromatin organization", *J Mol Biol* vol. 427, no. 3, pp. 695-706. Epub 2014 Oct 2. PMID: 25281900.

Schlub T & Holmes E C 2020, "Properties and abundance of overlapping genes in viruses", *Virus evolution* vol. 6 no. 1, art. no. veaa009.

Sonnenschein C & Soto A M 1999, *The Society of Cells: Cancer and Control of Cell Proliferation*, Oxford: Bios Scientific Pub Limited.

Sonnenschein C & Soto A M 2020, "Over a century of cancer research: Inconvenient truths and promising leads", *PLoS Biol* vol. 18, no. 4, art no. e3000670.

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, no. 1, Elsevier.

Soto A M & Sonnenschein C 2011, "The tissue organization field theory of cancer: A testable replacement for the somatic mutation theory", *BioEssays* vol. 33, pp. 332–340.

Sweeney M F, Hasan N, Soto A M, & Sonnenschein C 2015, "Environmental endocrine disruptors: Effects on the human male reproductive system", *Rev Endocr Metab Disord* vol. 16, pp. 341-357.

Versteg R 2014, "Cancer: Tumours outside the mutation box", *Nature* vol. 506, pp. 438–439.

Waks Z, Klein A M, & Silver P A 2011, "Cell-to-cell variability of alternative RNA splicing", *Molecular Systems Biology* vol. 7:506.

Weinberg R 2014, "Coming Full Circle - from endless complexity to simplicity and back again", *Cell* vol. 157, no. 1, DOI:10.1016/j.cell.2014.03.004.

Weyl H 1949, *Philosophy of Mathematics and of Natural Sciences*, Princeton, NJ: Princeton University Press.