

Two Dogmas of Biology

Leonore Fleming*

The problem with reductionism in biology is not the reduction, but the implicit attitude of determinism that usually accompanies it. Methodological reductionism is supported by deterministic beliefs, but making such a connection is problematic when it is based on an idea of determinism as fixed predictability. Conflating determinism with predictability gives rise to inaccurate models that overlook the dynamic complexity of our world, as well as ignore our epistemic limitations when we try to model it. Furthermore, the assumption of a strictly deterministic framework is unnecessarily hindering to biology. By removing the dogma of determinism, biological methods, including reductive methods, can be expanded to include stochastic models and probabilistic interpretations. Thus, the dogma of reductionism can be saved once its ties with determinism are severed. In this paper, I analyze two problems that have faced molecular biology for the last 50 years—protein folding and cancer. Both cases demonstrate the long influence of reductionism and determinism on molecular biology, as well as how abandoning determinism has opened the door to more probabilistic and unconstrained reductive methods in biology.

Keywords

reductionism • determinism • protein folding • disordered proteins • chromosomal instability • cancer

1 Introduction

Echoing W. V. O. Quine (1951), modern biology has been conditioned in large part by two dogmas: (1) One is *reductionism*—broadly understood as a belief that biological entities and processes correspond to some simpler, lower level—where “the reductive method *par excellence* is the ‘dissection of biological systems into their constituent parts’ (Van Regenmortel 2004, 1016)” (Kaiser 2015, 74); (2) The other is *determinism*—broadly understood as the necessity of fixed future states—where causal determinism is the idea that “given a specified *way things are at a time t*, the way things go *thereafter* is fixed as a matter of *natural law*” (Hofer 2016). Both dogmas play a large role in motivating and guiding biological research, although their presence is often only implied rather than stated outright. For example, in an editorial for

*Department of Philosophy, Utica College, 1600 Burrstone Rd, Utica, NY 13502, USA, lflemin@utica.edu

Received 7 September 2015; Revised 26 February 2017; Accepted 1 March 2017
doi:10.3998/ptb.6959004.0009.002



the journal *Infection and Immunity*, Fang and Casadevall claim, “Few scientists will voluntarily characterize their work as reductionistic. Yet, reductionism is at the philosophical heart of the molecular biology revolution” (2011, 1401). They go on to call molecular biology “a triumph of reductionism” claiming that reductionism “is implicit in much of molecular and cellular biology” (Fang and Casadevall 2011, 1402). And the same can be said about determinism, although such discussion is much more rare.

Continuing to parallel Quine, we can say that reductionism is “intimately connected” with determinism (cf. 1951, 38). In an article discussing systems biology, Mazzocchi says, “In modern science reductionism is usually combined with a linear and simplistic idea of causality, where the focus is restricted to the role of a single or few essential causes or factors (occurring at the fundamental physical level), *and with determinism*” (2012, 414, emphasis added). He continues by explaining that determinism “is the idea whereby any phenomenon in nature is completely determined by pre-existing causes to which it is bound by a relationship of necessity, and that only one possible future exists” (2012, 414). The main point of this paper is to argue that this “intimate connection” between reductionism and determinism is restrictive and problematic for biological research. Nevertheless, and in contrast to Quine’s original argument, my conclusion is not that both dogmas are completely ill-founded; rather, I claim we can sacrifice one to save the other. Reductionism is often criticized for being incomplete or misleading, which perhaps explains Fang and Casadevall’s statement that few scientists would openly call their research reductionistic, and while I do typically tend to agree with such criticisms, my conclusion in this paper is that the main problem with reductionism in biology is not the reductive method *per se*. Instead, the problem is the implicit attitude of determinism that usually accompanies it. Thus, I argue that the second dogma is what is really hindering to biological research and needs to be abandoned. Doing so will remove a critical constraint on biology and in turn reform the first dogma—reductionism—such that it can incorporate more probabilistic methods, and produce more adequate biological explanations.

My argument in this paper is threefold: (a) reductionism and determinism motivate much research in biology, especially molecular biology; (b) there is an intimate connection between reductionism and determinism (i.e., there is a natural association between the reductive method and a deterministic framework); and (c) this connection is problematic because it is based on a misunderstanding of determinism and deterministic systems. To support these points, the bulk of this paper is spent analyzing two problems that have faced molecular biology for the last 50 years—protein folding and cancer. Both cases demonstrate the combined influence of the two dogmas on molecular biology, the subsequent problems, and how biological practice has or has not changed in response. Before delving into such biological detail, however, I will provide some additional explanation regarding points (b) and (c) of my argument in the next two subsections, respectively.

1.1 *Reductionism and Determinism*

Methodological reductionism is intimately connected with strict determinism. The “idea that complex systems or phenomena can be understood by the analysis of their simpler components” (Fang and Casadevall 2011, 1401), is supported by the idea that lower level entities behave in simple, linear, and completely predictable ways. This reductive method of decomposition “assumes that there are but a small number of such functions that together result in the behavior we are studying, and that they are minimally interactive. We start with the assumption that interaction can be handled additively or perhaps linearly” (Bechtel and Richardson 2010, 23). Further, a key belief of methodological reductionism is that it “counts as reductionistic not only

to decompose a biological object or system into parts, but also to study the parts in isolation” (Kaiser 2015, 225). The parts of a system are understood as “intrinsically determined” and as such can be studied on their own, isolated from their original context (Bechtel and Richardson 2010, 26; Kaiser 2015, 225). The idea that any interactions between parts are predictable, simple, and sequential relies on the assumption of a strictly deterministic framework (Kaiser 2015, 233). As summed up by Soto and Sonnenschein, “those that assume a reductionist stance hope that eventually a neat, linear causal chain will be identified” (2006, 372). In other words, reductionism and determinism seem to go hand in hand fairly easily.

1.2 *Determinism and Predictability*

A common misconception is that determinism is the same thing as fixed predictability (Hoefler 2016; Werndl 2016). The two are so easily conflated that even Laplace and Popper have made the mistake (Hoefler 2016). While it is easy to confuse the notions of determinism and predictability, the important distinction is that the former is an ontological claim, while the latter is an epistemological claim. Determinism is the general idea that the state of a system (whether real or a model) at one time fixes the past and future states of that system (Butterfield 2005). Predictability, on the other hand, is the ability to compute or forecast the future based on the accurate measurement of initial states and conditions. Since human measurement will never be perfectly accurate, it should come as no surprise that deterministic systems can be unpredictable in a variety of ways (Werndl 2016), which reveals the problem with defining determinism as complete predictability.

As recent advances in mathematics have shown (the most famous being chaos theory), deterministic systems can display random and irregular behavior (Werndl 2009, 203). Probably the most well known example of a chaotic system is the weather. Imagine a model based on the amount of daily precipitation in an area. Whether or not this precipitation system actually follows deterministic laws, the fact that there is inaccuracy in initial conditions, even the tiniest amount, means that this system is unpredictable, or as Werndl explains it more specifically, the system is “asymptotically unpredictable” because the slight inaccuracies in initial conditions “eventually over time spreads out more than a specific diameter representing the prediction accuracy of interest” (2016, 671). Accurate predictability is not possible for this system, even if it is ultimately deterministic in nature. The ontology of the system ultimately doesn’t matter, because what does matter is the type of predictability or unpredictability in our system, and if we are aware of our epistemic limitations, then we should accommodate that in our modeling and our science.

Indeterministic systems involve the likelihood of many possible future states, and thus, what can be predicted from such systems are probability distributions or sets of possible outcomes. Indeterministic processes can be modeled stochastically as processes that evolve over time according to probabilistic laws. And while it might go against our intuition, Werndl demonstrates that “deterministic and indeterministic models can be observationally equivalent” (2016, 677). By comparing a deterministic daily precipitation model with an indeterministic coin-flipping model, Werndl shows that the two models give the same predictions, relative to an observation function, which means that deterministic models and stochastic models are not, in fact, mutually exclusive (2009; 2013; 2015; 2016). The moral of the story is that the assumption of determinism in biological research can be doubly problematic. First of all, it’s likely built on a mistaken definition of fixed predictability, rather than the recognition that determinism and predictability are different concepts, and that deterministic processes can be unpredictable in many different ways. Second, this typical mistaken definition incorrectly categorizes deterministic processes

and stochastic processes as mutually exclusive, which is unnecessarily hindering to biological research and methodology. However, if we abandon the dogma of determinism, our biological methods (such as reductionism) can be more open-minded to including stochastic models and probabilistic methodologies, in spite of any underlying biases or “hope that eventually a neat, linear causal chain will be identified.”

2 Protein Folding and Disordered Proteins

Over the last 50 years, the prevailing view in biochemistry has been “sequence determines structure determines function,” or The Central Dogma of Genomics, as it has been called (Petsko 2000). The basic idea is that a one-dimensional polypeptide chain (string of amino acids) codes for a unique, folded, three-dimensional functional structure (a protein). Since Christian B. Anfinsen’s groundbreaking research on protein folding in the 1960s (which led to him sharing the Nobel Prize in Chemistry in 1972), one huge question remains unanswered: How can a sequence of amino acids code for one specific 3D structure? For over 50 years biochemistry has been trying to explain how proteins that are denatured in the lab can refold into their native structure at a speed too fast to be the result of tedious searching or trial and error, given the astronomically large number of possible conformations. The breakthrough was expected to come from computer programs and algorithms built with information about physical forces and other molecular factors thought to influence proteins (e.g., hydrogen bonds, van der Waals interactions). In fact, “The early days saw hopes of finding simple sequence patterns—say of hydrophobic, polar, charged, and aromatic amino acids—that would predict protein structures and stabilities. That has not materialized” (Dill and MacCallum 2012, 1043). Such factors are clearly relevant to protein structure, however, it appears that the mechanism of protein folding is much more complicated than first thought. The now seemingly impossible task of predicting protein structure from amino acid sequence is known as the “protein-folding problem.” And although the “early days” were optimistic in assuming that computer algorithms developed within a deterministic and reductionist framework would solve the problem, the context for addressing protein folding has changed.

There are a number of worldwide competitions designed with the goal of finding the best computer algorithms to predict protein structure from amino acid sequence. The most famous is “Critical Assessment of protein Structure Prediction,” or CASP, which began in 1994 and is held every two years. Its 12th competition was scheduled to start in May 2016, and since the competition is always blind, the unpublished structures and results will be revealed in December 2016, when the 12th competition is set to complete. As Dill and MacCallum summarize it, “A grand challenge has been to develop a computer algorithm that can predict a protein’s 3D native structure from its amino acid sequence” (2012, 1043). But there is a common theme of disappointment and astonishment at how far we are from attaining this goal, especially considering the amount of time, money and importance placed on the pursuit. Additionally, with each new CASP competition there are more known protein structures to draw upon as public databanks grow rapidly. For example, as of August 18th 2016, the Protein Data Bank (Berman et al. 2000) had 113,095 proteins listed in their current holdings with 98,510 structures containing a structure factor file. Yet, according to Dill and MacCallum, since CASP6, “overall progress has slowed” (2012, 1044). Despite our increasing knowledge of protein structures, when it comes to predicting them based on amino acid sequences, “only modest gains were made over the last decade for certain classes of prediction targets” (Khoury et al. 2014, 1815).

The most recent competition, CASP11, was perhaps more encouraging than typical, due in large part to new methods of evolutionary coupling (Monastyrskyy et al. 2015). One of the

most successful groups, CONSIP2, led by David Jones of UCL, used evolutionary coupling and neural network architectures. Evolutionary coupling is based on the idea that “pairs of amino acids that interact with each other tend to ‘co-evolve’; that is, if one amino acid changes, then the second amino acid also changes in order to accommodate it” (Ovchinnikov et al. 2015, 2). Identifying these pairs and their changes by creating a contact map adds helpful constraints for predicting how proteins can possibly fold. Neural network architectures are “trained” using historical data and examples, and can automatically “learn” certain dependencies or constraints based only on those examples (Jones et al. 2014). What is obvious is that CONSIP2, and all the top contenders at CASP11, are stochastic, based on evolving probabilities (Kryshtafovych et al. 2015). The “grand challenge” might be presented deterministically, and past practice may be deterministic, but it is now clear that the best solutions are modeled stochastically, and still far away from making successful, fixed predictions. As the best model of CASP11, CONSIP2 achieved only 27% accuracy on de novo target proteins, yet such a result “indicates a breakthrough” (Monastyrskyy et al. 2015, 131).

Nonetheless, one might argue that the deterministic “sequence-structure-function” approach to protein folding still holds in-principle, if not yet in-practice. Such a claim runs into problems though if “wild-type” proteins from cell cultures are considered (Gershenson and Gierasch 2011). Unlike the protein folding that occurs in a highly diluted lab context, naturally-occurring proteins are affected by their cellular environment, including shape changes like those resulting from post-translational modifications (PTMs), which are not considered in the CASP structural predictions (Kryshtafovych et al. 2014). As discussed in Section 1.1, reductionism, when pursued under the bias of a deterministic framework, emphasizes decomposition and isolation of parts, assuming that those parts are “intrinsically determined.” What protein folding research has shown us, however, is that our best predictive models currently include dynamic information such as how amino acid pairs interact. Additionally, isolating techniques are beginning to be supplemented with “wild-type” protein experiments to further assist in understanding the age-old problem of protein folding. Even if these trends could be ignored, the deterministic framework of “sequence-structure-function” is threatened by a larger problem—intrinsically disordered proteins and intrinsically disordered protein regions.

2.1 *Intrinsically Disordered Proteins*

According to the conventional wisdom, a normal functioning protein has a predictable, folded structure. Proteins that lack regular form are considered malfunctioning, and cause neurodegeneration and diseases like Parkinson’s and Alzheimer’s, as well as contribute to various cancers, of which it has been found that 79% of proteins that are associated with cancer contain malfunctioning (i.e., disordered) regions 30 residues or longer (Liu and Huang 2014). Therefore, the common understanding is that proteins that do not fit the “sequence-structure-function” mold are defective. Yet, starting about ten years ago, evidence began to accumulate of proteins and protein regions that, despite lacking a predictable structure, are extremely functional.

Initially called “unstructured,” the literature used at least 35 other terms to describe these dynamic yet functional proteins before settling on “disordered” (Uversky 2013). Intrinsically Disordered Proteins (IDPs) and Intrinsically Disordered Protein Regions (IDPRs or IDRs) do not display a regular form, however, that does not mean they lack structure; quite the contrary, they have the potential to occupy a range of unfolded, partially folded, and fully folded, functional structures. And because IDPs can shift between these structures, they are often called “non-foldable” based on their lack of any well-defined “normal” structure. Although protein disorder sounds harmful, IDPs and IDPRs are present in all healthy organisms, and it is cur-

rently estimated that around 15%–45% of all eukaryotic proteins, and around 10%–35% of all prokaryotic proteins, contain disordered regions at least 30 residues in length (Dill and MacCallum 2012; Tompa 2012; Xue et al. 2012). Looking only at mammals, it is estimated that around 75% of mammalian signal proteins contain disordered regions at least 30 residues in length, and about half of the total proteins in mammals are estimated to have long disordered regions, with about 25% of those proteins thought to be fully disordered (Xue and Uversky 2016). Additionally, there is a growing number of examples illustrating how protein disorder is not only useful but also necessary for normal cellular functioning (Shoemaker et al. 2000; Pansca and Tompa 2012; Uversky 2013; Liu and Huang 2014). Clearly, IDPs/IDPRs are more the rule than the exception.

Examples of IDPs/IDPRs thus far have shown an astounding number of differently disordered states at all structural levels (primary, secondary, tertiary, quaternary), and well as an ability to switch between them: “Currently available data suggest that intrinsic disorder possesses multiple flavors, can have multiple faces, and can affect different levels of protein structural organization, where whole proteins or various protein regions can be disordered to a different degree” (Xue and Uversky 2016, 136). Vladimir Uversky claims that “one can even assume that every disordered protein (or at least every family of disordered proteins) is disordered in its own way” (2013, 701). This poses a large problem for the researcher following the deterministic (fixed predictability) approach, since one sequence seems to lead to an unpredictable range of structures and functions. If the idea of determinism is abandoned, however, then the variety of structures that are exhibited by an IDP/IDPR can be modeled and predicted stochastically. Furthermore, reductive methods that focus on amino acid sequence are then useful for predictive purposes, as long as they include the understanding that ‘protein structure’ is dynamic, and protein folding (or non-folding) is a probabilistic process. In other words, by removing the dogma of determinism. The 50-year deterministically motivated protein-folding problem ceases to exist because, as Dill and MacCallum say, “we believe it is no longer useful to frame the question this way. Protein folding is now a whole field of research—a large, growing, and diverse enterprise ... bigger than any individual research question” (2012, 1045).

3 Cancer Evolution and Chromosomal Instability

Reflecting back on the last 40 years of cancer research, Robert Weinberg, a leading cancer researcher and biology professor at MIT, has “witnessed wild fluctuations from times where endless inexplicable phenomenology reigned supreme to periods of reductionist triumphalism and, in recent years, to a move back to confronting the endless complexity of this disease” (2014, 267). The reductionist (and determinist) illusion was strongest for him in 1982, when it appeared that certain cancers could be attributed to a single genetic mutation. The prediction was that a small number of mutations could explain all cancerous growths; an expectation that nicely confirms Soto and Sonnenschein’s claim that “those that assume a reductionist stance hope that eventually a neat, linear causal chain will be identified” (2006, 372). However, the optimism of this theory was fleeting, and “From the point of view of the reductionist hoping that a small number of molecular events might explain cancer, things went downhill from there for the next 30 years” (Weinberg 2014, 269). Weinberg continues, explaining that “two simple notions—that small numbers of genes would suffice to transform cancer cells, and that all mammalian cells would follow the same set of genetic rules during the course of neoplastic transformation—were undermined” (2014, 269). With more and more failed predictions, the variation and complexity of cancer at the molecular level became increasingly apparent to the point that “each tumor seemed to represent a unique experiment of nature, acquiring a unique set of mutant genes and

in an unpredictable chronological order” (Weinberg 2014, 269). In other words, each mutated tumor appeared mutated in its own way.

Since the year 2000, when the era of big data was already gaining speed, reductive analyses were primary, consumed by the notion that genome sequencing would uncover hidden answers. The trend continues today, with an emphasis on sequencing and big data, yet, as Weinberg laments, we still lack the general ability to predict cancer progression (cf. 2014, 270). As an example, The Cancer Genome Atlas (TCGA) used large-scale genome sequencing and genome analysis techniques to “systematically explore the entire spectrum of genomic changes involved in more than 20 types of human cancer” (National Institutes of Health 2016). Started in 2006, TCGA officially completed in 2014 having discovered nearly 10 million cancer-related mutations after genetically profiling around 10,000 tumors. When the almost 400 million dollar project completed, Heidi Ledford reported in *Nature News* that “Although a few ‘drivers’ stood out as likely contributors to the development of cancer, most of the mutations formed a bewildering hodgepodge of genetic oddities, with little commonality between tumours” (2015, 128). Despite the very many efforts to discover those genetic mutations that “drive” oncogenesis, it appears that such a deterministic and reductive search is not only inconclusive, but also wrong-headed.

Data now show that populations of cancer cells taken from clinical settings are incredibly heterogeneous at multiple levels, much more than was expected, which directly challenges the reductionist-determinist method of decomposition and isolation discussed in Section 1.1 (Hepner 1984; Torres et al. 2007; Park et al. 2010; Navin et al. 2011; Baca et al. 2013). More than ten years ago, at the International Cancer Genomics Consortium Meeting, a point of consensus among the members was the rather implicitly deterministic view that the “Heterogeneity of a selected tumour type should be as small as possible to allow for the identification of relevant mutations. Therefore, prioritization of the initial set of tumour types must consider the genetic heterogeneity of a given tumour type” (2007, 5). In other words, mutations should be investigated based on the assumption that they are intrinsically determined to cause cancer, and thus can be investigated outside their original context. And although such minimization of variation can make initial lab work simpler, the problem, of course, is when such research is applied to “real world” cell populations in clinical situations, where it is found that it consistently fails (Stepanenko and Kavsan 2012; Watson 2013). A reductive, deterministically motivated investigation assumes that “eventually a neat, linear causal chain will be identified,” which is why cancer research is still searching for specific driver mutations that are inextricably linked with cancer development. Such a discovery would provide a way to accurately predict (and remove) with certainty the presence of cancer (assuming the system is deterministic and the measurements completely accurate). As such methods have repeatedly proven disappointing, an alternative is viewing cancer as an evolutionary process, with selection acting on heterogeneous populations at multiple levels. This more closely resembles real world cell populations as observed in clinical settings, confirming Weinberg’s claim that the recent trend in cancer research is “a move back to confronting the endless complexity of this disease” (2014, 267).

3.1 *Chromosomal Instability*

Much cancer research today focuses on the evolutionary dynamics of cancer growth at different levels of biological hierarchy: “Some investigators focus on the particular features of the cells themselves; some focus on features of whole populations of such cells; some focus on hierarchical or spatial relations between cells; some focus on interactions between cancer cells and their microenvironment” (Lean and Plutynski 2016, 45). Christopher Lean and Anya Plutynski

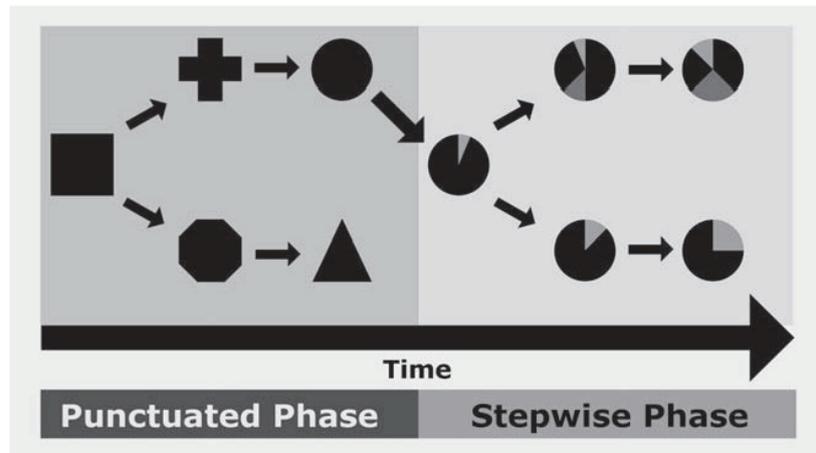


Figure 1: The two phases of cancer evolution. The punctuated phase represents a group of cells with high Chromosomal Instability (CIN), which results in rapid, genome-level changes that are unique to each individual cell. There is competition among these unstable cells, and the survivor (the circle) expands clonally during the stepwise phase. Genome-level changes from the first phase are stabilized, and minor, lower-level genetic changes evolve gradually (pie slices). Since research on cancer evolution is almost entirely focused on the second phase, it is good to remember that those mutations recorded from the second phase, are the result of the very random and erratic process from phase one. Figure from Horne et al. (2015), reproduced with permission.

argue that all of these factors likely affect cancer development, and “The key is to see this as a multilevel selection process, with selection operating at distinct levels simultaneously” (2016, 53). This includes individual selection, such as among cells, as well as group selection, like at the level of cell populations or lineages, or at the level of cell colonies. It is this last example, at the level of cell populations or cell colonies, i.e., the genome level, that Henry Heng and colleagues found what they claim is a unifying feature of all cancers—Chromosomal Instability (CIN) (2013). CIN is measured by the cell-to-cell variability of karyotypes (the number and structure of chromosomes) within a specific cell population. Karyotype variation can be due to the loss of entire chromosomes or fractions of chromosomes, or due to the increase or decrease of nonclonal chromosomal aberrations (NCCAs), which are seemingly random variations in karyotypes that alter the entire karyotype landscape. Such karyotype alterations, in turn, cause the individual genes within those chromosomes to change unpredictably.

Single-cell or single-nucleus genome sequencing has shown that from the perspective of the genome, cancer appears to evolve in two very distinct phases—punctuated and stepwise (Heng et al. 2006; Heng et al. 2013; Horne and Heng 2014; Wang et al. 2014; Horne et al. 2015). As illustrated in Figure 1, the first stage is “punctuated” because of rapid and unstable genome-level changes (NCCAs) that are unique to each individual cell. During this first phase, individual cells are competing with each other. The survivor(s) propagate during the second, stepwise stage of cancer evolution, when minor lower-level genetic changes evolve gradually. The erratic genome-level changes from the first stage are now stable and have been duplicated uniformly via clonal reproduction. In other words, the first stage is characterized by dynamic nonclonal chromosomal aberrations (NCCAs), while the second stage consists of established clonal chromosome aberrations (CCAs).

It is during the second phase of cancer evolution that the majority of cancer research has

focused. However—and *this is the important point*—the mutations in the second phase that are used to distinguish and classify a cancer cell population, are unpredictable, and the result of the random, genome-level changes from the chaotic first evolutionary phase. Thus, the mutations studied in the second phase that are investigated as potential “drivers” of cancer development, are the result of a highly stochastic and noisy process of individual selection among a lot of precarious, precancerous cells. There is nothing deterministic, linear or predictable about the way these mutations come about. In fact, it has been shown that cell populations in the punctuated phase are so unstable and unpredictable that they cannot be cloned, even in a short amount of time. According to Batoul Abdallah and colleagues, “Each cell sampled exhibited a unique karyotype within a short time-span, making it clear that these karyotypes are not clonable” (2013, 3643). And whereas the population in the stepwise phase exhibited uniform proliferation among cells, the unstable cell populations from phase one were quite variable in cell cycle times and overall population growth (Abdallah et al. 2013). For example, the doubling times in unstable cell populations were unique, leading to very unequal growth distributions as there were often a few extreme outliers. In other words, cancer research needs to shed any remaining assumptions of determinism or strict predictability, because the process is very random and irregular.

Clearly, shifting the focus of cancer research from the second evolutionary phase to the first evolutionary phase is not a complete or good solution, but neither is continuing to pursue a deterministic, reductive search for specific cancer-causing “driver” mutations in the second phase. If cancer is understood evolutionarily, then it must also be understood probabilistically. The composition of a tumor cell population is the result of a dynamic, stochastic process; hence the reason Weinberg would claim that “each tumor seemed to represent a unique experiment of nature” (2014, 269). When determinism, and the related goal of finding specific cancer-causing mutations, is abandoned, cancer research is less constrained and more open-minded. Reductionism, without the burden of determinism, can still search for cancer-causing mutations in cell populations; however, the relevant context for such a search is broadened to include real world cell populations, guided by an updated understanding that cancer development is best understood as a stochastic, multi-level, evolutionary process among heterogeneous populations. According to Lean and Plutynski, “What started as a minority view has now become an active area of research, drawing upon mathematical modeling, computer simulation, experimental evolution in model systems, and even the generation of phylogenies, or branching evolution in cancer progression ...” (2016, 40). Though perhaps not yet the dominant framework for investigating cancer, abandoning the dogma of determinism gives rise to a less restricted, interdisciplinary, enlightened view of cancer as a multi-level evolutionary process.

4 Biology without Determinism

Deborah Gordon, whose current research projects include studying ant behavior in space (cf. Countryman et al. 2015), said “the limits to what we know about ants are set as much by how we frame the problem as by the number of person-hours spent getting the answer” (2010, 13). As a graduate student in the early 1980s, Gordon worked under the fairly deterministic and reductive paradigm set by E.O. Wilson that ants are genetically programmed to fill certain roles and carry out specific tasks optimal for the functioning of the entire colony (cf. Gordon 2010, 5-6). However, Gordon’s research, then and now, challenges the algorithmic perception of a functioning ant colony, replacing it with the view that underneath the fairly predictable and efficient colony system is a noisy bunch of erratic ants that “bump into each other as they travel through an unpredictable world ...” (2010, 146). According to Gordon, if we observe ants in great detail, we notice that “There is stochasticity, randomness and noise, both in how ants react

and in what they are reacting to” (Gordon 2010, 142–143). Thus, predicting the behavior of an individual ant with any real accuracy is rather impossible since “an ant’s reactions are not deterministic” (Gordon 2010, 143).

Much stochasticity and unpredictability has also been documented at the molecular level (as a few examples, in cellular decisions, Balázsi et al. 2011; cell death, Stevens et al. 2013; and gene expression and RNA splicing, Marinov et al. 2014). However, a sense of determinism still lingers near to the genetic code, as if the amino acid sequence can only be understood as a hard-and-fast instruction manual of predictable future states. Yet, as Weinberg points out, molecular research has shown that our reductive attempts to gather and hoard large amounts of data “overwhelm our abilities of interpretation” (2014, 270) to the point that “The coupling between observational data and biological insight is frayed if not broken” (2014, 271); sending us down the wrong path, time and time again. The problem is not with methodological reductionism, or with amassing huge quantities of data at the molecular level, but rather, as I argue, doing so within a strictly deterministic framework, assuming fixed predictability and misunderstanding the difference between determinism and predictability. As reductive methods easily cohere with expectations of simple linearity, this mistaken view of determinism adversely affects the pursuit of reductionism as well. Luckily, with the removal of determinism and the false promise of complete predictability, a critical constraint is lifted from biological and reductionistic methodologies. In the case of the protein folding problem, the removal of determinism welcomed in dynamic predictive models of protein structure; the recognition that naturally occurring proteins behave much differently than proteins isolated in the lab; and the acknowledgement that disordered proteins are fairly common, and that different “flavors” of disorder can be modeled stochastically. Similarly in the case of cancer research, the removal of determinism opened the door to investigating real world, heterogeneous cell populations; and the stochastic modeling of cancer populations as irregular, dynamic, multi-level, evolutionary processes ushered in a new wave of understanding cancer, which is much more complex than the search for specific genetic mutations. Falling is the framework of understanding proteins in terms of “sequence-structure-function.” Weakening is the framework of understanding cancer as the result of “specific cancer-causing driver mutations.”

I have argued that there are two dogmas in biology that influence biological research and methodology—reductionism and determinism. Two examples from molecular biology—protein folding and cancer—have helped me demonstrate that the dogma of determinism is problematic and hindering to biological research, and thus should be abandoned. Furthermore, because “those that assume a reductionist stance hope that eventually a neat, linear causal chain will be identified,” reductionistic methods are particularly vulnerable to deterministic ways of thinking (Soto and Sonnenschein 2006, 372). However, if the assumption of determinism is removed, reductionistic methodology is less restricted and better able to incorporate irregular, dynamic behaviors. While I am advocating for the abandonment of determinism, especially as it influences reductive methods in biology, it is good to remember Bechtel and Richardson’s claim about the strategies of decomposition and localization: “We may be led to erroneous explanations, but it may be the only way to begin the task of explaining and understanding complex systems. The failure of decomposition is often more enlightening than is its success ...” (Bechtel and Richardson 2010, 23–24).

Acknowledgments

This paper was inspired by a biochemistry lecture on protein folding given by Daniel Barr, who I thank for many conversations about protein structure and for discussing with me the indeterminism of this world.

I thank Robert Brandon for his many helpful comments on this paper as it evolved via many drafts, and Dan McShea for comments on early drafts of this paper, as well as a very supportive journal editor and two reviewers, who were extremely patient and helped push me to clarify a number of my points. Lastly, I thank Richard Fleming for his help in thinking through the particulars.

Literature cited

- Abdallah, B. Y., S. D. Horne, J. B. Stevens, G. Liu, A. Y. Ying, B. Vanderhyden, S. A. Krawetz, R. Gorelick and H. H. Q. Heng. 2013. "Single Cell Heterogeneity: Why Unstable Genomes Are Incompatible With Average Profiles." *Cell Cycle* 12: 3640–3649.
- Baca, S. C., D. Prandi, M. S. Lawrence, J. M. Mosquera, A. Romanel, Y. Drier, K. Park, N. Kitabayashi, T. Y. MacDonald, M. Ghandi, E. Van Allen, G. V. Kryukov, A. Sboner, J.-P. Theurillat, T. D. Soong, E. Nickerson, D. Auclair, A. Tewari, H. Beltran, R. C. Onofrio, G. Boysen, C. Guiducci, C. E. Barbieri, K. Cibulskis, A. Sivachenko, S. L. Carter, G. Saksena, D. Voet, A. H. Ramos, W. Winckler, M. Cipicchio, K. Ardlie, P. W. Kantoff, M. F. Berger, S. B. Gabriel, T. R. Golub, M. Meyerson, E. S. Lander, O. Elemento, G. Getz, F. Demichelis, M. A. Rubin and L. A. Garraway. 2013. "Punctuated Evolution of Prostate Cancer Genomes." *Cell* 153: 666–677.
- Balázsi, G., A. van Oudenaarden, and James J. Collins. 2011. "Cellular Decision Making and Biological Noise: From Microbes to Mammals." *Cell* 144: 910–925.
- Bechtel, W., and R. C. Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, MA: MIT Press.
- Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. 2000. "The Protein Data Bank." *Nucleic Acids Research* 28: 235–242. <www.rcsb.org>.
- Butterfield, J. 2005. "Determinism and Indeterminism." *Routledge Encyclopaedia of Philosophy Online*.
- Countryman, S. M., M. C. Stumpe, S. P. Crow, F. R. Adler, M. J. Greene, M. Vonshak and D. M. Gordon. 2015. "Collective Search by Ants in Microgravity." *Frontiers in Ecology and Evolution* 3 (25).
- Dill, K. A., and J. L. MacCallum. 2012. "The Protein-Folding Problem, 50 Years On." *Science* 338: 1042–1046.
- Fang, F. C. and A. Casadevall. 2011. "Reductionistic and Holistic Science." *Infection and Immunity* 79: 1401–1404.
- Gershenson, A., and L. M. Gierasch. 2011. "Protein Folding in the Cell: Challenges and Progress." *Current Opinion in Structural Biology* 21: 32–41.
- Gordon, D. 2010. *Ant Encounters: Interaction Networks and Colony Behavior*. Princeton, NJ: Princeton University Press.
- Heng, H. H., S. W. Bremer, J. B. Stevens, S. D. Horne, G. Liu, B. Y. Abdallah, K. J. Ye and C. J. Ye. 2013. "Chromosomal Instability (CIN): What It Is and Why It Is Crucial to Cancer Evolution." *Cancer and Metastasis Reviews* 32: 325–340.
- Heng, H. H. Q., J. B. Stevens, G. Liu, S. W. Bremer, K. J. Ye, P.-V. Reddy, G. S. Wu, Y. A. Wang, M. A. Tainsky and C. J. Ye. 2006. "Stochastic Cancer Progression Driven by Non-Clonal Chromosome Aberrations." *Journal of Cellular Physiology* 208: 461–472.
- Heppner, H. G. 1984. "Tumor Heterogeneity." *Cancer Research* 44: 2259–2265.
- Hofer, C. 2016. "Causal Determinism." In *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), edited by E. N. Zalta. <https://plato.stanford.edu/archives/spr2016/entries/>

determinism-causal/.

- Horne, S. D., and H. H. Q. Heng. 2014. "Genome Chaos, Chromothripsis, and Cancer Evolution." *Journal of Cancer Studies and Therapy* 1: 1–6.
- Horne, S. D., S. A. Pollick, and H. H. Q. Heng. 2015. "Evolutionary Mechanism Unifies the Hallmarks of Cancer." *Int. J. Cancer* 136: 2012–2021.
- Kaiser, M. I. 2015. *Reductive Explanation in the Biological Sciences*. New York, NY: Springer Berlin Heidelberg.
- International Cancer Genomics Consortium (ICGC). 2007. *International Cancer Genomics Consortium Meeting Report*. Toronto, Canada: ICGC.
- Jones, D. T., T. Singh, T. Kosciolk and S. Tetchner. 2014. "MetaPSICOV: Combining Coevolution Methods for Accurate Prediction of Contacts and Long Range Hydrogen Bonding in Proteins." *Bioinformatics* 31 (7): 999–1006.
- Khoury, G. A., A. Liwo, F. Khatib, H. Zhou, G. Chopra, J. Bacardit, L. O. Bortot, R. A. Faccioli, X. Deng, Y. He, P. Krupa, J. Li, M. A. Mozolewska, A. K. Sieradzan, J. Smadbeck, T. Wirecki, S. Cooper, J. Flatten, K. Xu, D. Baker, J. Cheng, A. C. B. Delbem, C. A. Floudas, C. Keasar, M. Levitt, Z. Popović, H. A. Scheraga, J. Skolnick, S. N. Crivelli and F. Players. 2014. "WeFold: A Cooperation for Protein Structure Prediction." *Proteins* 82 (9): 1850–1868.
- Kryshtafovych, A., A. Barbato, B. Monastyrskyy, K. Fidelis, T. Schwede and A. Tramontano. 2015. "Methods of Model Accuracy Estimation Can Help Selecting the Best Models From Decoy Sets: Assessment of Model Accuracy Estimations in CASP11." *Proteins* 84 (Suppl 1): 349–369.
- Kryshtafovych, A., J. Moult, P. Bales, J. F. Bazan, M. Biasini, A. Burgin, C. Chen, F. V. Cochran, T. K. Craig, R. Das, D. Fass, C. Garcia-Doval, O. Herzberg, D. Lorimer, H. Luecke, X. Ma, D. C. Nelson, M. J. van Raaij, F. Rohwer, A. Segall, V. Seguritan, K. Zeth and T. Schwede. 2014. "Challenging the State of the Art in Protein Structure Prediction: Highlights of Experimental Target Structures for the 10th Critical Assessment of Techniques for Protein Structure Prediction Experiment CASP10." *Proteins* 82: 26–42.
- Lean, C. and A. Plutynski. 2016. "The Evolution of Failure: Explaining Cancer as an Evolutionary Process." *Biology & Philosophy* 31: 39–57.
- Ledford, H. 2015. "End of Cancer Atlas Prompts Rethink: Geneticists Debate Whether Focus Should Shift From Sequencing Genomes to Analysing Function." *Nature* 517: 128–129.
- Liu, Z. and Y. Huang. 2014. "Advantages of proteins being disordered." *Protein Science* 23: 539–550.
- Marinov, G. K., B. A. Williams, K. McCue, G. P. Schroth, J. Gertz, R. M. Myers and B. J. Wold. 2014. "From Single-Cell to Cell-Pool Transcriptomes: Stochasticity in Gene Expression and RNA Splicing." *Genome Research* 24: 496–510.
- Mazzocchi, F. 2012. "Complexity and the Reductionism–Holism Debate in Systems Biology." *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 4: 413–427.
- Monastyrskyy, B., D. D'Andrea, K. Fidelis, A. Tramontano and A. Kryshtafovych. 2015. "New Encouraging Developments in Contact Prediction: Assessment of the CASP11 Results." *Proteins* 84 (Suppl. 1): 131–144.
- National Institutes of Health. 2016. "The Cancer Genome Atlas: Mission and Goal." <https://cancergenome.nih.gov/abouttcga/overview/missiongoal>.
- Navin, N., J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W. R. McCombie, J. Hicks and M. Wigler. 2011. "Tumour Evolution Inferred by Single-Cell Sequencing." *Nature* 472: 90–94.

- Ovchinnikov, S., L. Kinch, H. Park, Y. Liao, J. Pei, D. E. Kim, H. Kamisetty, N. V. Grishin and D. Baker. 2015. "Large-Scale Determination of Previously Unsolved Protein Structures Using Evolutionary Information." *eLife* 4: e09248.
- Pansca, R., and P. Tompa. 2012. "Structural Disorder in Eukaryotes." *PLoS ONE* 7: e34687.
- Park, S. Y., M. Gönen, H. J. Kim, F. Michor and K. Polyak. 2010. "Cellular and Genetic Diversity in the Progression of in Situ Human Breast Carcinomas to an Invasive Phenotype." *The Journal of Clinical Investigation* 120: 636–644.
- Petsko, G. A. 2000. "Dog Eat Dogma." *Genome Biology* 1: 1002.1001–1002.1002.
- Quine, W. V. O. 1951. "Two Dogmas of Empiricism." *Philosophical Review* 60: 20–43.
- Shoemaker, B. A., J. J. Portman and P. G. Wolynes. 2000. "Speeding Molecular Recognition by Using the Folding Funnel: The Fly-Casting Mechanism." *Proceedings of the National Academy of Science USA* 97: 8868–8873.
- Soto, A. M., and C. Sonnenschein. 2006. "Emergentism by Default: A View From the Bench." *Synthese* 151: 361–376.
- Stepanenko, A. A., and V. M. Kavsan. 2012. "Immortalization and Malignant Transformation of Eukaryotic Cells." *Cytology and Genetics* 46: 96–129.
- Stevens, J. B., B. Y. Abdallah, G. Liu, S. D. Horne, S. W. Bremer, K. J. Ye, J. Y. Huang, M. Kurkinen, C. J. Ye and H. H. Q. Heng. 2013. "Heterogeneity of Cell Death." *Cytogenetic and Genome Research* 139: 164–173.
- Tompa, P. 2012. "Intrinsically Disordered Proteins: A 10-Year Recap." *Trends in Biochemical Sciences* 37: 509–516.
- Torres, L., F. Ribeiro, N. Pandis, J. Andersen, S. Heim and M. Teixeira. 2007. "Intratumor Genomic Heterogeneity in Breast Cancer With Clonal Divergence Between Primary Carcinomas and Lymph Node Metastases." *Breast Cancer Research and Treatment* 102: 143–155.
- Uversky, V. N. 2013. "A Decade and a Half of Protein Intrinsic Disorder: Biology Still Waits for Physics." *The Protein Society* 22: 693–724.
- Van Regenmortel, M. H. 2004. "Reductionism and Complexity in Molecular Biology. Scientists Now Have the Tools to Unravel Biological and Overcome the Limitations of Reductionism." *EMBO Rep* 5: 1016–1020.
- Wang, Y., J. Waters, M. L. Leung, A. Unruh, W. Roh, X. Shi, K. Chen, P. Scheet, S. Vattathil, H. Liang, A. Multani, H. Zhang, R. Zhao, F. Michor, F. Meric-Bernstam and N. E. Navin. 2014. "Clonal Evolution in Breast Cancer Revealed by Single Nucleus Genome Sequencing." *Nature* 512: 155–160.
- Watson, J. 2013. "Oxidants, Antioxidants and the Current Incurability of Metastatic Cancers." *Open Biology* 3.
- Weinberg, R. A. 2014. "Coming Full Circle—From Endless Complexity to Simplicity and Back Again." *Cell* 157: 267–271.
- Werndl, C. 2009. "What are the New Implications of Chaos for Unpredictability?" *The British Journal for the Philosophy of Science* 60: 195–220.
- . 2013. "On Choosing Between Deterministic and Indeterministic Models: Underdetermination and Indirect Evidence." *Synthese* 190: 2243–2265.
- . 2015. "Determinism and Indeterminism." In *The Oxford Handbook of Philosophy of Science*, edited by P. Humphreys, 210–32. Oxford, Oxford University Press.

-
- . 2016. “Determinism.” In *The Routledge Companion to Free Will*, edited by K. Timpe and M. Griffith, 669–79. New York: Routledge-Taylor & Francis.
- Xue, B., K. A. Dunker and V. N. Uversky. 2012. “Orderly Order in Protein Intrinsic Disorder Distribution: Disorder in 3500 Proteomes From Viruses and the Three Domains of Life.” *Journal of Biomolecular Structure and Dynamics* 30: 137–149.
- Xue, B., and V. N. Uversky. 2016. “Unfoldomes and Unfoldomics: Introducing Intrinsically Disordered Proteins.” In *Molecular Science of Fluctuations Toward Biological Functions*, edited by M. Terazima, M. Kataoka, R. Ueoka and Y. Okamoto, 125–50. Tokyo: Springer Japan.
-

© 2017 Author(s)

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license, which permits anyone to download, copy, distribute, or display the full text without asking for permission, provided that the creator(s) are given full credit, no derivative works are created, and the work is not used for commercial purposes.

ISSN 2475-3025