Can biological complexity be reverse engineered?

Sara Green[1]

Abstract

Concerns with the use of engineering approaches in biology have recently been raised. I examine two related challenges to biological research that I call the synchronic and diachronic underdetermination problem. The former refers to challenges associated with the inference of design principles underlying system capacities when the synchronic relations between lower-level processes and higher-level systems capacities are degenerate (many-to-many). The diachronic underdetermination problem regards the problem of reverse engineering a system when the non-linear relations between system capacities and lower-level mechanisms are changing over time. Braun and Marom argue that recent insights to biological complexity leave the aim of reverse engineering hopeless - in principle as well as in practice. While I support their call for systemic approaches to capture the dynamic nature of living systems, I take issue with the conflation of reverse engineering with naïve reductionism. I clarify how the notion of design principles can be more broadly conceived and argue that reverse engineering is compatible with a dynamic view of organisms. It may even help to facilitate an integrated account that bridges the gap between mechanistic and systems approaches.

1. The virtues and pitfalls of reverse engineering

Reverse engineering methodologies are currently gaining terrain in biological fields such as systems biology and neuroscience. In response to these developments, experimental biologists have raised concerns regarding the associated quest for design principles that they take to imply an assumption of a rather static and modular design of organisms. A workshop in Konstanz brought together philosophers and biologists to discuss the implications of research methodologies in the life sciences.[2] This paper focuses in particular on concerns raised regarding reverse engineering of biological networks.

The experimental biologists Erez Braun and Shimon Marom (this issue) provide fascinating insights to biological complexity by stressing how living systems are characterized by a (deep) two-way degeneracy and lack of separation of time-scales. Against this complexity, they criticize so-called reverse engineering approaches for investigating biological systems as if these were programmed and fully decomposable engineering systems, designed to conduct pre-designed functions. This paper clarifies and supports their criticism of naïve

---

reductionism, but questions the argued discrepancy between reverse engineering and systemic approaches.

The notion of reverse engineering has its origin in the attempt to copy or further develop the design underlying a functional system in hardware and software engineering without access to the design protocol (Chikofsky and Cross, 1990). Although the aim to design functional systems is primarily seen in synthetic biology and bioengineering, reverse engineering has currency in many other biological fields where it commonly refers to the process of '(detailed) examination of a functional system, in the face of limited *a-priori* knowledge of its design principles' (Braun and Marom, this issue).[3] In the abovementioned definition I have bracketed the requirement of a *detailed* examination. This requirement marks an important difference between Braun and Marom's view on reverse engineering as a reductionist strategy and the account I develop in this paper. For Braun and Marom, 'design principles' seem to refer to the results of fine tuning (and selection of) specific control parameters engineered to lock the system in a given dynamic state.[4] In contrast, the view I develop is inspired by reverse engineering in systems biology where abstraction is central. In this context, design principles typically refer to general features of functional organization that are independent of system-specific contexts or particular molecular parameters. My aim in this paper is to take seriously the challenge posed for (reductionist) reverse engineering, while nuancing the description of reverse engineering approaches through examples from systems biology.

It is worth noting that the call for engineering approaches often reflects a wish for an alternative to the reductionist study of isolated molecules and pathways. The goal is to identify organizational patterns that may otherwise be lost in the preoccupation with molecular details. If the same principles can be applied in the design of different types of engineered systems from cars to computers or airplanes, it seems likely that some principles are shared among different biological systems or even among engineered and biological systems (Braillard, 2010). The objective to identify shared formal criteria for functional design is not new. It dates back at least to the early days of control system theory and modern cybernetics, where feedback control was formalized as a basic principle for maintaining stable states and oscillations (Wiener, 1948). But reverse engineering methods have recently become increasingly widespread in data-intensive biological fields concerned with the identification of non-random connectivity patterns in biological networks. When un-aided pattern detection is not analytically feasible, mathematical models can guide the search for relevant structure-function relations (Levy and Bechtel, 2013).

The choice of the graph theoretical framework as a key representational strategy reflects how reverse engineering is often conducted at a high level of abstraction. Displaying regulatory connections as nodes and edges affords a topological analysis that abstracts from the details of what these units represent to identify functional capacities that relate to the architecture of the network. For instance, many real-world networks seem to share general characteristics such as the small-world effect and scale-free connectivity distribution (Barabási & Oltvai, 2004; Bechtel, this issue). It is highly contested how much biological information one can derive from topological analysis and whether the scale-free distribution exemplifies a

---

[3] This definition differs from how reverse engineering is understood in the literature on adaptationism. For clarification see (Green, Levy and Bechtel, 2015).

[4] For a discussion of the use of the notion of design principles and other functional terms in Braun and Marom's papers, see (Krohs, this issue).

2

biological design principle (Arita, 2004; Keller, 2005). But many biologists agree that there is a connection between the robustness of biological networks and their non-random connectivity distribution and hierarchical structure (Steinacher & Soyer, 2012). Other examples of design principles are bi-stable switches (Tyson et al. 2003) and overabundant sub-circuits in gene regulatory networks, called *network motifs* (Alon, 2007a, see below). To some researchers, such findings provide optimism that there is simplicity in the apparent complexity of biological systems (Csete and Doyle, 2002; Alon, 2007c).

The quest for design principles reflects a hope that key properties of biological systems can be understood without knowing all the lower-level causal details. This is not only a point about practical convenience but also about the relevant level of analysis. The cancer biologist Lazebnik (2002) provocatively compared biomedical research strategies to the attempt to fix a radio by atomizing the system into component parts and studying these in isolation. If the malfunction of the system is connected to the orchestrated organization of parts and processes, searching for broken molecular components is bound to fail. Lazebnik therefore proposes an engineering approach to investigate how the components are wired together as a functional whole. Lazebnik's original choice of example however also clarifies why reverse engineering is often considered a reductionist strategy. Biological systems do not function like a pre-designed radio; there is no simple and static "wiring" of a living cell. Critics are concerned that engineering approaches underestimate biological complexity when assuming that living systems are similar.

To exemplify, topological network analysis is sometimes advertised as a bias-free decomposition strategy, in contrast to hypothesis-driven functional analysis. But choices of structural criteria – often inspired by design principles in electronic networks – also imply a bias. It is typically possible to fit several structural modules to a given data-set, and researchers cannot assume any tight overlap between structural and functional modules (Krohs, 2010). While the problem of underdetermination is a general problem in science, the challenge is particularly apparent in fields where data-intensive modeling is only loosely integrated with experimental analysis (Krohs, 2012). These methodological concerns are complemented by a more fundamental worry that biological design principles may not exist at all (Marom et al., 2009). The inherent plasticity, degeneracy and evolvability of the functional organization of living systems indicate that methodologies and conceptual frameworks from engineering should not be uncritically applied in biology.[5] An important question is therefore whether there are better strategies for facing not only the challenge of 'synchronic degeneracy' (the many to many mappings between lower-level functional organization and system behavior) but also of what I shall call 'diachronic degeneracy' (the change of these relations over time).[6]

To discuss the prospects of reverse engineering methodologies, I compare the concerns raised by Marom and Braun to a similar debate in systems biology. I first draw on two examples to illustrate the problem of 'synchronic underdetermination' for reverse engineering of

---

[5] In the context of evolution biology, design thinking has been criticized for having adaptationist implications. I have discussed the relation between design approaches and adaptationism elsewhere (Green 2014; Green, Levy and Bechtel, 2015), and this paper instead concerns the challenges for reverse engineering associated with synchronic and diachronic degeneracy. I shall, however, discuss the possible role of reverse engineering in evolutionary studies in Section 5.
[6] I thank Ben Sheredos for suggesting these terms to cover the challenges addressed.

biological networks (Section 2.1). I then discuss whether the biases can be accounted for by increasing the resolution of the analysis (Section 2.2). While this seems to be a feasible solution if the system is relatively stable over time, Section 3 highlights the challenges associated with what I call the 'diachronic underdetermination problem'. This problem relates to the prospects of identifying design principles at the background of changes to cross-level relations over developmental and evolutionary time-scales. Section 4.1 and 4.2 reexamines the merits of engineering approaches against this challenge. Section 5 demonstrates how reverse engineering can be productively combined with a systemic approach. Section 6 concludes with some general remarks on the prospects of reverse engineering methods in biology.

2. Engineering approaches and synchronic underdetermination

2.1. Reverse engineering biological networks

To illustrate the problem with synchronic underdetermination, this section examines two attempts to reverse engineer biological networks. The examples are taken from systems biology and neuroscience, respectively. Systems biology is a relatively new interdisciplinary approach that applies mathematical modeling and engineering approaches to the interpretation of biological datasets on regulatory interactions (e.g. gene regulation). To guide the detection of organizational patterns, a common strategy is to represent the regulatory interactions as a network model that affords a mathematically aided scan for non-random patterns. An example is the pioneering work on motif-detection in transcriptional regulatory networks by Uri Alon's group, at the Weizmann Institute of Science, Israel. The group modelled datasets of transcriptional regulation in *E. coli* as a directed graph and used algorithms to scan for common patterns of connectivity between small numbers of nodes in the network (Alon, 2007a). A few motifs turned out to be statistically overabundant compared to randomized networks. Examples of network motifs are pictured in Figure 1, representing the connectivity patterns between transcription factors (X and Y) and target genes (Z).
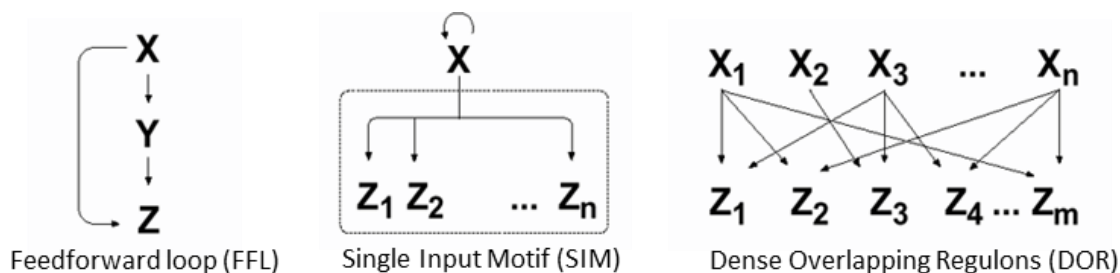
Figure 1. Examples of network motifs.

The finding that biological networks are composed of a small set of overabundant motifs, while many other logically possible motifs are absent, provides optimism that biological systems are not as complex as they could have been. For instance, the Feedforward loop pictured to the left in Figure 1 was the only statistically significant motif among 13 possible variants of three-node motifs, and only two of 199 possible four-node motifs are frequently found in biological networks (Alon, 2007a). These topological features are biologically interesting because the connectivity patterns are thought to constrain the functional capacities of the regulatory sub-circuits. For instance, mathematical modeling of the Feedforward loop suggested that this motif could act as a persistence detector that delays

protein synthesis in the presence of noisy activation signals. This prediction has been confirmed in experiments with transgenic *E. coli* where the outputs of different regulatory motifs were measured against changes in activating signal molecules (Mangan et al., 2003). The group therefore suggests that it may be possible to decompose biological networks to a set of networks motifs with characteristic functions such as signal amplification, sustained oscillations or noise filtering (Alon, 2007a).

The abovementioned approach has, however, been criticized for oversimplifying biological phenomena, and for drawing unsupported conclusions from a biased decomposition of low-resolution co-expression data. Most of the criticism has centered on the problematic inferences from the beneficial functions of network motifs to the selective origin of these (reviewed in Green, 2014). But critics have also pointed to problematic idealizations involved in the modeling of activation profiles for the motifs. Among these are assumptions that genes are either maximally on or off, and that network motifs are functionally isolated from the regulatory network they are embedded in. While the functional predictions for some motif-models were confirmed in initial experiments on transgenic bacteria, studies of more complex networks question the existence of biological counterparts to the rigid structure-function categorizations. Most biological networks have a high degree of overlap among nodes and edges involved in network motifs, and simulations suggest that the functions of motifs are sensitive to altered parameter values (Ingram et al., 2006). The possibility has been raised that functions may be realized in multiple ways from the underlying interactions between components, and the same molecular circuit may enable multiple higher-level functions (depending on parameter values of environmental stimuli).[7] Applicability of the structure-function classifications of motifs in other systems thus needs to be further investigated. In summary, the example illustrates the general problem that close integration among components in a system has the potential to undermine the ideal of non-idealized localization schemes that depict biological mechanisms (Bechtel, this issue).

A similar problem is faced in neuroscience where researchers aim to reverse engineer the underlying neural activity patterns from a behavioral response to a stimulus. Marom et al. (2009) show how a *wrong* design principle inferred for a biological 'toy system' may be accepted if state-of-the-art methods for testing representation schemes in neurophysiology are followed. The authors apply reverse engineering methodologies to identify the known (but concealed) design principle underlying neutral computation in a Braitenberg vehicle that is pre-designed to move continuously in a static environment while avoiding obstacles. The Braitenberg vehicle is in this context a Lego Mindstorms robot with two ultrasonic eyes that detect inputs from the environment and transmit electronic responses to a network of biological cortical neurons. Monitoring the spiking activity of the network of neurons, developed *in vitro* upon a substrate-embedded electrode array, allows the system to classify spatial objects and to send signals to motors attached to the wheels of the robot.

For the specific robot that is 'reverse engineered', the design principle for object classification is known in advance. A well-defined algorithm classifies spatial inputs by considering only the rank order of spikes from individual neurons. The classification of objects is based on the order of spikes from stimulus to neurons that are broadly tuned to respond to inputs from both ultrasonic eyes. Thus, the 'correct' design principle is based on the relative temporal

---

[7] Furthermore, the use of random networks as a comparison class for detecting statistically and functionally significant motifs has been challenged (for details, see Green, 2014).

order of neural recruitment. To evaluate reverse engineering methods in this context, the authors tested the validity of an alternative, but *wrong*, hypothesis. The alternative design principle is based on population response rates, i.e. the temporal profile of total spike counts is interpreted rather than the identity of single neurons.
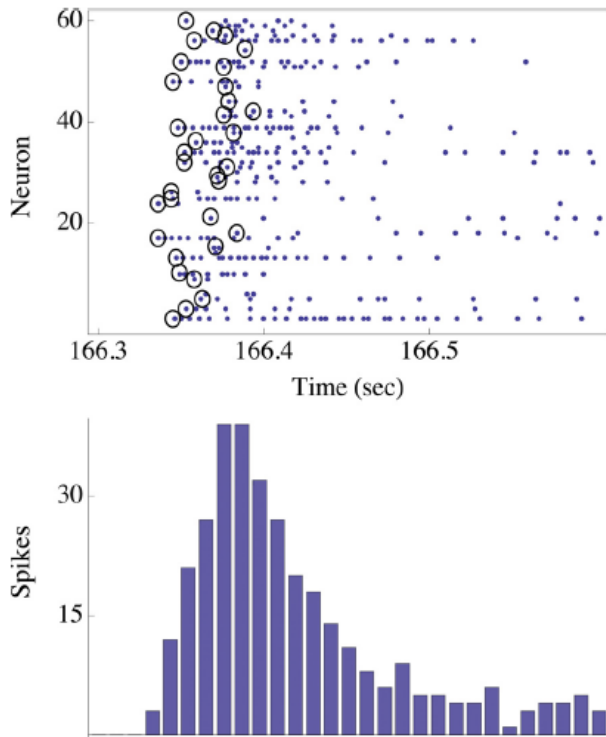
Figure 2. Spikes emitted by 60 neurons after stimulus triggered by the right eye of the robot. The figure below shows the total spike counts per time unit. See text for clarification. Figure as originally published in Marom et al. (2009). On the precarious path of reverse neuro-engineering. Front Comp Neurosci, 3:4, 1-4.

The dots on the upper diagram of Figure 2 represent the spikes emitted by individual neurons, and the first spike of each neuron is marked with a black circle. On the lower panel of the figure, the population response hypothesis is illustrated with a count histogram showing the total spike counts. In contrast, the 'correct' design principle is based only on the order of first spikes (here 24, 17, 26, 25, 48, 1 and so forth). The authors ask whether using a state-of-the art method in neurophysiology would reveal that the population response hypothesis is incorrect. To evaluate the accuracy of input classification of the eye that initiates network stimulation, they used a non-linear version of a Support Vector Machine (SVM).[8] Strikingly, the classification based on population response receives a high score for classification of the objects (0.9) even though it is wrong. This result raises the concern that this hypothesis would be considered confirmed in a more realistic situation where the correct design principle is unknown.

The authors raise the possibility that design principles inferred in reverse engineering studies may reflect methodological biases rather than inherent features of the system: 'in the possible

---

[8] SVMs are supervised, algorithm-based models used for pattern recognition and classification of new data into categories, e.g. to recognize specific objects on photographs based on examples of the categories in 'training sets'. The model can then classify new examples into categories, represented as groups of points in Euclidian space separated by a hyperplane.

absence of such principles, pushing the experiments to various limits may not necessarily lead to the selection of one universal ("true") model. In other words, different models may "win" in different extreme experimental conditions' (Marom et al., 2009, p. 3). Expectations that biological systems resemble engineered systems may be too easily confirmed, because experiments are designed from the same (reductionist) mindset. In the following section, I examine possible ways in which the heuristic value of these strategies can be defended by accounting for the biases through calibration of evidence with other methods. Section 3 discusses situations where this strategy may not suffice.

2.2. Accounting for biases

Before dismissing the heuristic value of reverse engineering, it is worth reflecting on possible ways to take advantage of and account for the biases.[9] Design heuristics can be productive despite the reliance on false assumptions if the analysis can uncover to what extent the systems are similar and different (Knuuttila and Loettgers, 2013). Sometimes ideas can fail in interesting ways, uncovering aspects that are different or more complex than assumed. In any case, some simplifying assumptions are required to constrain the overwhelming problem space of scientific analysis (Simon, 1966; Bechtel and Richardson, 1993/2010). That we need heuristic strategies to deal with complexity is also reflected in the pervasive use of metaphors and analogies that *enable,* although also possibly mislead, scientific reasoning (Hesse, 1963; Lakoff & Johnson, 1980; Nersessian, 2008).[10] Thus, the choice to make is not between neutral and biased methodologies, but between ignorance and awareness of the shortcomings of different epistemic tools. As I shall clarify below, greater awareness of the limitations of research methodologies can motivate a more careful and systematic calibration of preliminary results with independent sources of evidence (Wimsatt, 2007).

When reexamining the cases, it seems that there are several ways to account for the biases. In the case of detection of network motifs, the criticism concerned the lack of determination of functions from regulatory circuit structures for complex organisms. Responding to this challenge, Alon (2007b, 2007c) stresses his awareness of the shortcomings of the simple models in more complex biological contexts, but maintains the value of such 'first-approximations' in experimental research. The simple diagrams of network motifs may serve a fruitful role as idealized models that more realistic models can be developed from through iterative loops with experimental results (Mangan et al. 2003; Bechtel, this issue).[11] At the community level of systems biology, attempts to evaluate the effects of biases of such methods have recently emerged. An example is the DREAM project, Dialogue for Reverse-Engineering Assessment and Methods (Stolovitzky et al., 2007; Prill et al., 2010). Part of this initiative regards blind testing and evaluation of different methods to reverse engineer artificial networks with known design principles, and to discuss possible experimental designs that can

---

[9] While design-thinking arguably is insufficient for uncovering all aspects of living systems, the insights gained in successful cases provide *prima facie* reasons for not dismissing all design approaches and for evaluating the use of these in specific contexts.

[10] As pointed out in Konstanz by the philosophers and one of the biologists, it is questionable whether biological research could do without design-inspired functional metaphors. Even design-critical approaches use engineering-terms to state the flexibility of organisms, e.g. when highlighting the ability of organisms to 'rewire' their genome (Krohs, this issue)

[11] One possible way to test such functions experimentally is to draw on fluorescent staining techniques to observe regulatory responses to changes in environmental parameters.

counteract unwanted effects of biases. Thus, rather than dismissing the attempt, researchers aim to develop methods to systematically distrust the results of mathematically aided network analysis.

Similarly, in the toy example on reverse neuroengineering, the hypothesized design principles can be systematically "distrusted" in several ways. Recalling the example, the fact that the two representation schemes both perform well in classification tests made the authors question the ability of biologists to experimentally falsify wrong hypotheses. Given the problems for reverse engineering principles even in the *simple* setup described, Marom et al. (2009) see the prospects of reverse engineering as highly limited. But the results of a later study suggest that increasing the complexity of the example can provide tighter constraints on the problem space. An *in vitro* study compared the efficacy of a number of representation schemes, including the two candidates discussed in Section 2, in a series of discrimination tasks (Kermany et al., 2010). Modification of the spatial location of inputs and the time elapsed revealed that rate-based schemes have a reduced efficacy and lower stability over long time-periods compared to time-based measures in localizing the stimulus, whereas rate-based schemes are more sensitive to the history of simulation and hence to adaptation processes. Thus, the static environment and primitive task described for the toy example may be what makes it so easy to realize the task efficiently in several ways. Fine-graining the functional capacity may therefore provide tighter constraints on the problem-space, contrary to what is assumed.

Another possibility is to fine-grain experimental measures by drawing on new molecular techniques for *in vivo* testing of representation schemes. As a comment to Marom and colleagues' concerns with reverse engineering approaches, O'Connor et al. (2009, p. 923) state that '[…] inferring design principles from measurements of neural activity is treacherous. Hypotheses regarding behavior ultimately need to be tested by manipulating neural activity during behavior. The fourth step in reverse engineering therefore requires inactivation and activation of specific cell types'. Examples of methods to accommodate the fourth step include protein sensors that convert changes in membrane potentials to fluorescence signals, and genetically targeted molecular switches to manipulate the activity in defined cell types in living organisms.[12]

The aspects summarized above leave the impression that the main problem has been the *use* of design heuristics, rather than the tools themselves. But Braun and Marom raise a deeper concern that regards not only the complexity-challenge of multiple mappings between molecular structures and higher-level functions, but also the ability of organisms to rewire their functional organization over time. While the 'synchronic underdetermination problem' may be dealt with by increasing the resolution of models and experiments, this solution may not suffice for the attempt to reverse engineer diachronic relations.

3. The problem of diachronic underdetermination

To illustrate why 'looking closer' may not always suffice to account for biological complexity, consider the strategy used in the development of the first (and so far only) whole-cell model

---

[12] Cell-type-specific methods for controlling neural activity are currently being developed for studies with mice and monkeys (He et al., 2012; Gregoriou et al., 2012). The strategy of combining different kinds of evidence to overcome the problem of heuristic biases is of course a general methodological strategy that by no means is restricted to or special for engineering-approaches (Bechtel and Richardson, 1993/2010; Wimsatt, 2007).

of the human urogenital parasite *Mycoplasma genitalium*. The astonishing accomplishment to model the dynamics of the whole cell involved the integration of more than 1,900 experimental parameters in a system of coupled large-scale simulations (Karr et al., 2012). Although upscaling the complexity of models of organisms may seem like a straightforward solution to embrace biological complexity, several new challenges are faced when relying on this strategy. Despite the simplicity of *M. genitalium*, with only 525 genes, it is tremendously challenging to integrate data from different experimental contexts and models of different biological processes. No single model can accommodate all the different kinds of data. Ordinary differential equations cannot capture the spatial features of biochemical and biophysical interactions, and partial differential equations and agent-based simulations quickly lead to intractable models (cf. Wolkenhauer and Green, 2013). Karr and colleagues dealt with this challenge through reliance on a decomposition strategy akin to the one described by Bechtel and Richardson (1993/2010). The system was divided into 28 functional submodules, corresponding to different cellular functions (transcription, replication, host interactions etc.) which are modelled in isolation. These models were then recomposed through numerical integration of 16 output variables for every 1 second sub-model simulation, and their values were updated as inputs to the next simulation round. Together, the mosaic of integrated models forms the simulation of a complete cell cycle.

While the results of model of *M. genitalium* may readily be considered a success, the example also offers a lesson in modesty regarding the prospects of applying this strategy to complex systems (Gross, 2013). The example shows that also large-scale models have to rely on decomposition strategies that separate operations of functional modules on short time-scales. Researchers hope that by imposing boundaries on the system, the results may hold to a "first approximation" for models that can subsequently be modified to better encompass complex organization. It however remains an open question whether more complex systems can be successfully reverse engineered in detail like *Mycoplasma genitalium*. In particular, several researchers have stressed their skepticism regarding projects like the Blue Brain/Human Brain Project and the SyNAPSE project that aim to reverse engineer the brain in all its causal details.[13] For instance, the assumption that the neocortex is organized into modular building blocks characterized by a basic canonical pattern of connectivity has been criticized with reference to the plastic nature of brain processes (Schierwagen, 2012). Decomposition strategies (whether structural, functional or temporal) are highly productive if biological systems are hierarchically structured into functional modules, but face limitations if these are not decomposable or even near-decomposable, or if the functional organization changes over time (cf. Simon, 1962, 1966; Bechtel and Richardson, 1993/2010; Marom, 2010). It is therefore important to reflect on the ramifications of the assumption that organisms are organized into semi-static functional modules like machines.

A pervasive problem in biological research is that knowledge on the relevant level of organization and time-frames for experimental studies is often lacking. To distinguish fact from artifact, biologists must constantly evaluate the effects of conditions imposed by the observer of the system through experimental setups and modeling frameworks. Marom

---

[13] In the official video for the Blue Brain Project it is declared that in the new *in silico* neuroscience there will be "nothing we cannot measure, no aspect of the model we cannot manipulate, and there will be no question we cannot ask". Available at bluebrain.epfl.ch, accessed 15/05-2014. While this aim seems overly optimistic, a more promising middle-way to model complex systems may be to advance lower-resolution simulations of whole organs. This strategy is currently pursued in cardiac computational physiology (Carusi et al., 2012).

(2010) demonstrates this for timescales documented for neurological processes. Most experiments in neuroscience are limited to measurements on the timescale of seconds. But excitable dynamics have been observed over extended timescales, revealing a much richer dynamics than typically observed (Marom, 2010). This suggests that the time-frame of analysis constrains the observable characteristics of neurological processes, and the influence of methodological choices must be critically assessed. The issue is particularly problematic for studies of multilevel dynamic processes like brain activities and behavior that cannot be reduced to sequential information flow.[14]

Furthermore, studies of neural circuits under stable experimental settings often fail to account for the degeneracy and adaptability of the functional organization of neural processes over time. The adaptability of brain processes was recognized already in the beginning of the last century by the German neurologist and psychiatrist Kurt Goldstein who observed that the neural system of injured soldiers could readjust even to severe brain damage. Goldstein concludes on this basis that organisms are very different from machines in their capacity to reorganize organizational processes to recover critical functions whereas most machines would break down if you remove essential parts (Goldstein, 1934; Nicholson, 2013). The causal relations between a population of neurons and brain functions are not just multiply realized in the sense of many-to-one relations between microscopic and macroscopic processes. There is also one-to-many degeneracy between a given pattern of activity and the mapping of brain functions (Braun and Marom, this issue). Due to this two-way degeneracy biological networks may be able to rewire their organization in an overwhelming number of ways.

Functional plasticity in terms of many-to-many relations and reorganization of regulatory structures over time also characterize the highly non-linear relations between genotypes and phenotypes. One genotype can give rise to the realization of number of different phenotypic characteristics, but a large space of genotypes can also give rise to similar phenotypes (see also Section 5). Like neural networks, these regulatory connections may change even over short time-scales from interactions with the environment. Whereas the details of gene regulatory networks are often thought of as 'prewired' or 'programmed' by selection to deal with specific challenges, recent studies suggest that adaptation to environmental challenges sometimes happens through changes that are not caused by mutation and subsequent selection but by 'rewiring' of the genome. Phenotypic plasticity, as a facilitator of adaptive change, has been documented for genetic as well as non-genetic perturbations (Stolovicki et al. 2006; Espinosa-Soto et al., 2011; Katzir et al., 2012).

More specifically, Braun's group have used genetic engineering techniques to rewire a yeast genome by placing an essential metabolic gene, from the histidine biosynthesis pathway, under the regulation of the galactose utilization system that is repressed if glucose is available (David et al., 2010). In a glucose medium, cells are severely challenged to produce histidine and must adapt the regulatory system significantly in order to survive. The experiments showed that a high number of individual cells can adapt through yet unknown mechanisms of heritable phenotypic switching.

---

[14] For a discussion on the implications of these insights for mechanistic accounts of explanation, see Bechtel (this issue).

Studies with twin cell populations have revealed that yeast cell populations with identical genomes can exhibit similar adaptive growth phenotypes while displaying diverse expression dynamics of metabolically essential genes. Accordingly, the authors suggest that regulatory changes in gene expression should be understood as a self-organization process that is intrinsically coupled to population-environment dynamics (Stolovicki and Braun, 2011; Braun and Marom, this issue).[15] The degeneracy of living systems thus indicates that the way genomes are often conceptualized as 'programs' that generate output signals from input signals through 'cellular computing' should be revised and replaced with more adequate guiding assumptions (Braun & David, 2011, p. 190).

Braun and Marom (this issue) take the abovementioned challenges to show the limitation of engineering approaches in biology. While acknowledging that engineering approaches have often been used to push a reductionist or genetic deterministic agenda, I shall suggest what seems to be the opposite; namely that one of the strongest virtues of reverse engineering lie in its ability to make progress on the problems of understanding the non-linear relations and dynamic aspects of development and evolution.

4. Beyond naïve reverse engineering

4.1. Reexamining engineering approaches

There are many examples of how comparisons of organisms and engineered systems have led to misunderstandings, and it has accordingly been argued that the heuristic value of engineering approaches are limited and often purely metaphorical (Nicholson, 2013; Pauwels, 2013). Sometimes, the use of engineering approaches in practice however goes beyond metaphors in exploring the application of concrete principles and modeling tools in the two domains (Calcott, 2014).[16] Furthermore, these comparisons exceed the reductionist design perspective in searching for the class of models that can realize the observed system behavior, rather than by decomposing the system to detailed interactions between components. From this perspective, a search for design principles can afford insights to phenomena that Braun and Marom (this issue) consider a challenge to reverse engineering, namely the capacities of biological systems to maintain dynamic stability largely independent of specific control parameters.

To illustrate this point, I shall briefly revisit the case on network motifs. Although one fruitful way to use abstract models is to use these as templates for more realistic models, it is important to recognize that representational accuracy is not always the aim of reverse engineering. A central aim is to understand why systems with shared organizational features conduct similar operations. That is, although specific functions of concrete network motifs may depend on specific parameters, it may be possible to classify – from a coarser grained perspective – parameter spaces for which the generalized functions hold. With this aim in mind, Tyson and Novák (2010) explore the shift between two motifs functions for a Feedforward loop with an inhibitory connection, depending on the relative temporal order of activation of the two transcription factors. A similar approach is taken by Wall et al. (2005)

---

[15] Krohs (this issue) questions whether the experimental result should reflect a population-related property. I have a similar concern regarding the lack of clarity of what constitutes self-organization. However, the challenge to reductionism is well taken.

[16] Calcott (2014) points to illuminating examples and highlights parallels between a 'diachronic engineering goal' in software development and 'evolvability' in EvoDevo.

who explicitly state that they selected the reference parameter values 'in the interest of exploring model behaviors rather than in the interest of understanding the dynamics of specific natural systems' (supplementary material). Thus, the aim is to explore how a set of parameter changes affect system behavior for models with mathematically defined organizational constraints. This aim seems to be directly relevant to a goal called for by Braun and Marom, namely to understand why dynamic states of living systems are robust to changes at the molecular level.

Robustness of biological functions against perturbations to the system components and underlying connectivity is one of the major research topics in systems biology (Gross, 2013). Contrary to what one would expect from Braun and Marom's (this issue) account, a key insight provided by reverse engineering analyses is that robustness is often not attained through fine-tuning of specific mechanisms but due to features associated with the organization of the system. Architectural features of robust systems have been demonstrated for robustness to gradients in embryonic patterning (Eldar et al., 2002), concentration robustness in biochemical reaction networks (Shinar and Feinberg, 2011), for bacterial chemotaxis (reviewed in Braillard, 2010), and for the spindle assembly checkpoint mechanism in cell division (reviewed in Gross, 2013). Thus, engineering-inspired research in systems biology is directly related to the aim of understanding what robust networks have in common in terms of system-level organizational features, and to identify organizational features that may or may not work for a given property.[17] It is important to mind the level of abstraction in such analyses, since these often go beyond the specific mechanisms realized in concrete systems and instead identify possible classes of models based on observationally or operationally defined constraints. Recalling the challenge posed in Section 3 can such approaches also provide progress for understanding diachronic relations? The following section explores the merits of design thinking in studies of development and evolution.

4.2. Mechanistic and systemic approaches to development

Currently, the most popular conceptual schemes used to describe developmental processes rely on information and computation metaphors that draw on a comparison between evolving software programs and static hardware systems that run the program. As Braun and Marom (this issue) point out, the problem with such a view is that causal relations are pictured as a fixed one-way instructive information flow encoded in the genome. If the relations between different organizational levels are characterized by complex many-to-many mappings, the traditional focus on additive models of (programmatic) genetic and selective effects will provide limited insight to how genetic effects relate to the generation of phenotypic variation. Thus, genetic changes need to be contextualized by an understanding of how the regulatory (or developmental) system works as a whole.

Critics of reductionist and gene-centric approaches have for many years pointed to the existence of "invariant sets of genotypes" as a challenge to the simplistic idea of the genome as a program (e.g. Goodwin, 1982; 1994). While two-way degeneracy of the genotype-phenotype

---

[17] While this property involves ascription of functions to biological systems (Krohs, this issue), the analysis needs not be adaptationist or even focused on evolutionary questions (Green, Levy and Bechtel, 2015). Neither does the analysis necessarily entail assumptions about optimal design. It is rather ironic that discussions of modularity in biology have often focused solely on selective advantages whereas engineers have discussed the costs and trade-offs of different degrees of modularity (Krohs, 2009).

map presents a challenge to biological research, functional flexibility also provides optimism that we do not need to map biological mechanisms in all their quantitative and qualitative details to understand patterns and regularities in systems dynamics. Evolutionary systems biologists have recently found inspiration in theoretical approaches to development that draw on the conceptual framework of dynamical systems theory (Jaeger et al. 2012). I first examine the general idea of the approach, and then proceed to clarify how reverse engineering may be used to support this aim.

The basic insight that motivates the dynamical systems theory approach is that even complex systems converge to a limited set of (self-organizing) stable states as long as certain general conditions are obeyed (Goodwin, 1982; Kauffman, 1993; Oster & Alberch, 1982). To exemplify, Goodwin and colleagues used computer simulations to clarify why different algal species go through similar developmental stages where certain morphological patterns (whorls) are generated (Goodwin, 1994). The central idea is that although the possible mechanisms leading to behaviors observed in specific organisms are manifold, the inherent constraints of biochemical and biophysical interactions in the systems may greatly limit the space of possible numbers of, and transitions between, dynamical regimes. In this approach, stable (developmental) states are modelled as attractors in phase space, and changes in dynamic states are reflected by a shift between basins of attraction that separate the phase space into distinct dynamical regions. The phase space representation allows for prediction of the states that a developmental system converges to, given the initial conditions and other changes that cause the system to either stay within a basin of attraction or to shift to another dynamic state. An understanding of what enables such transitions, or buffers against these, is believed to clarify why some lower-level changes have no impact on system behavior; why other changes have a drastic effect; and why some phenotypic designs are not feasible at all (see Section 5).

The notion of design principles in the example I examine in Section 5 is inspired by structuralist envisions of general principles influencing the range of phenotypes a system can or cannot implement (Goodwin, 1994; Jaeger and Monk, 2013: Jaeger and Sharpe, 2014). The principles point to the generative constraints on phenotypic patterning. This 'thin' notion of design, for instance, is employed in the following passage by Gould: '[H]owever much we celebrate diversity and revel in the peculiarities of animals, we must also acknowledge a striking 'lawfulness' in the basic design of organisms' (Gould 1977/2006, p. 319). The lawfulness that Gould refers to includes physically and structurally constrained relations between size, shape and function of organisms (e.g. the disproportionate relations between volume, weight, surface area, strength of organic material, and between requirements for energy consumption and gas exchange for organisms of varying sizes). Design explanations from this perspective may therefore reflect insights to the synchronic constraints on the survival of organisms under certain circumstances, rather than a blueprint or programmatic principles of any specific organism (Wouters, 2007; Braillard, 2010). In the context of evolutionary systems biology, *evolutionary design principles* reflect generative aspects of constraints in determining the possibilities of diachronic relations to produce phenotypic variation (Green, Fagan and Jaeger, 2015). To exemplify the significance of such constraints for understanding evolution, there are limits to what respiratory designs (gills, lungs, or open respiratory systems) could work on land or in water for organisms of different size. Vertebrates have had the evolutionary potential to increase surface area by

compartmentalization and convolution of organs (e.g. lungs and intestines), but land-living organisms such as insects are constrained to remain small. On the other hand, the small size of insects enables these to fly using design strategies that are not feasible for bigger animals (Gould, 1977/2006).[18] Other examples of general designs in need of a developmental explanation are fixed digits of vertebrae, stripe and segment formation in insects, wing shapes, and patterns in leaf and flower structures. The point is that there is a limit to the designs that are *possible* in relation to capacities maintained in organisms, and in relation to the phylogenetic lineage and developmental systems that produce the variation. The goal of analysis is therefore akin to Hugo de Vries' call for 'laws of variation' that clarify how non-genetic as well as genetic constraints enable the generation of stable phenotypes. The thin notion of design is not only compatible with systems-theoretical views but supportive of the task to define the space of possible dynamic states at the background of system organization and environmental interactions. I therefore believe that the framework is supportive of Braun and Marom's call for insights to so-called 'universal features' of biological systems.

The existence of regularities in dynamic and phenotypic patterns provides optimism that the roads of evolution are not randomly curved, even if specific endpoints are inherently unpredictable. But the key limitation of the dynamical systems approach is - as also pointed out by Stolovicki and Braun (2011) - that it so far has remained highly theoretical and disconnected from the causal processes at work in real-world biological systems. Similarly, evolutionary systems biologists contend that general principles of functional design and development can only be found if the gap between systems-theoretical and experimental analysis can be bridged (Jaeger and Crombach, 2012; Green, Fagan and Jaeger, 2014). As I shall clarify below, reverse engineering can be used to give the otherwise highly theoretical concepts of attractors and bifurcations an empirical basis by connecting 'ensemble approaches' to detailed experimental studies.[19] Reverse engineering therefore can directly support – rather than counteract – an understanding of the dynamics of synchronic and diachronic genotype-phenotype relations.

5. Reverse engineering as a tool for integration

Johannes Jaeger and colleagues, at Centre de Regulació Genòmica in Barcelona, aim for an integrative approach to study development and evolution of dipteran insects. Inspired by dynamical systems theory, a basic assumption of their research is that the diversity of functional morphological traits, and the processes producing these, is constrained by a finite amount of dynamic behaviors that can be realized by particular biological systems (Jaeger and Crombach, 2012; Jaeger and Sharpe, 2014). For the group it is, however, important that design principles not only cite superficial similarities but refer to design themes that are experimentally supported as well as formally well-defined.

---

[18] Similarly, the laws of gravity greatly constrain the evolutionary roads for different lineages in terms of types of skeletons to support the weight of animals on land. An animal of the size of an elephant or dinosaur cannot have the body proportions of a dog.

[19] 'Ensemble approaches' here refer to mathematical modeling of networks to study properties associated with network architectures such as network motifs (Alon, 2007a), potential for evolvability (Espinosa-Soto et al., 2011) and attractors of complex (developmental) systems (Goodwin, 1982; Kauffman, 1993).

Connecting general theoretical insights to experimentally tractable biological systems has so far been challenged by discrepancies in the explanatory aims and standards of the mechanistic and systems theoretical approaches (Green, Fagan and Jaeger, 2015), and by the lack of relevant data on gene regulatory networks (GRN) to test theoretical results.[20] One of the most extensively studied developmental systems is the gap gene system that governs segment development in insects. Data on nucleus-specific information on spatial and temporal distribution of gap gene products and maternal transcription factors have recently become available. But obtaining such data requires labor intensive techniques, such as immunofluorescent staining and confocal microscopy, and detailed modeling is only feasible for a few extensively studied organisms like *Drosophila melanogaster* (Jaeger and Crombach, 2012). Jaeger's group therefore explores the use of reverse engineering methods to fit systems parameters from coarser-grained mRNA expression data to create GRN models amenable for phase space analysis.[21]

Figure 3 illustrates the procedure of the approach. The first step is extensive experimental measurements to generate quantitative mRNA expression data that enable an empirical basic for changes in expression patterns of the developing insect embryo. The next step is to draw on reverse engineering methodologies to fit computational models of the gene regulatory networks to the dataset. The system is mathematically modelled through coupled ODE-models that characterize changes in protein concentrations in each nucleus as a function of gene regulation, diffusion and decay rates. Whereas reverse engineering of GRNs traditionally have been restricted to inference of static network structure, more dynamic network models can now be obtained through global non-linear optimization algorithms (for details, see Ashyraliyev et al., 2009; Jaeger and Monk, 2010). These models can then feed empirical import to the otherwise highly abstract representations of phase space models whose dimensions correspond to the concentration of regulatory factors (state variables) as a result of the regulatory network structure and initial conditions.

The dynamical systems framework helps to clarify how regulatory interactions within the whole network constrain protein synthesis across the embryo. In phase space representations, a point corresponds to a dynamic state of the system. Because genes influence each other's transcription (through activation and inhibition) and are at the same time affected by environmental inputs (e.g. maternally expressed transcription factors), some combinations of proteins cannot be maintained over time in the network. Accordingly, depending on the initial values of the state variables, the system will converge to stable points (attractors) in the characteristic basins of attraction.

---

[20] It should be noted that gene regulatory networks in this context do not refer to a static mapping of DNA sequences but to the interconnections between nodes, i.e. genes, mRNAs and protein products. Access to genome-wide gene regulatory network data now affords modeling of the concerted changes in expression levels of the nodes in a gene expression network over time.

[21] To explore the merits of the coarser-grained analysis, they have compared results from reverse engineered networks (inferred from mRNA expression data) and GRNs made from detailed data on *Drosophila* (Jaeger and Crombach, 2012). Since they obtained a close match, the group now draws on this method to study development in other dipteran insights.
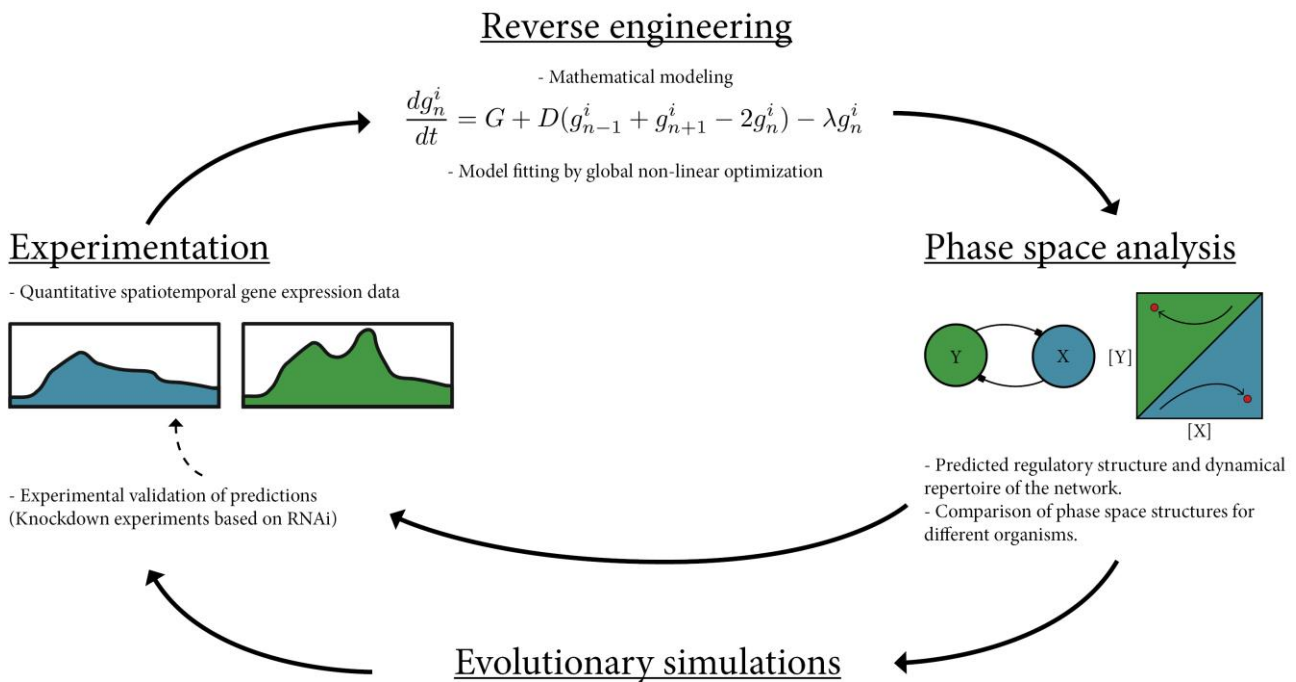
Figure 3. Illustration of the approach described in (Jaeger and Crombach, 2012), where reverse engineering is used to bridge between experimental and theoretical approaches. See text for clarification.

The representation of changes in gene expression patterns over time as trajectories in phase space can be illustrated through a simple empirical example. Staining the protein products of gene regulatory networks with fluorescent antibodies can reveal how some genes exclude the expression of other genes in the same nucleus. For instance, the gap genes *giant* and *Krüppel* never co-occur in the same nucleus of *Drosophila* embryos but block each other's expression (blocked arrows on Fig. 3; for details see Jaeger and Crombach, 2012). The relation between concentrations of such genes can be modelled as a phase space diagram, where the axes represent the concentration of gap-gene products. The scope of possible trajectories is determined by the features of the basins of attraction that divide the phase space (in this case there are two such basins, see Figure 3). Particular trajectories are influenced by initial conditions, e.g. maternally expressed proteins that differ in concentration in different nuclei. As long as the initial expression of one of the genes is higher than the other, the system will converge to high expression of the most "dominant" gene and repression of the other, corresponding to one of the two attractor states (red dots in state space).

The structure of the phase space is determined by the system components (state variables) and system parameters. In this context, system components are transcription factors that regulate the expression of genes and other transcription factors, physiological components like enzyme activities, membrane potentials and environmental features that affect regulatory activities. Systems parameters specify strengths and types of regulation (activation, repression) between transcription factors and genes, and are modelled as connectivity matrices. Although the GRN of the whole gap gene system is complex and high-dimensional, a dynamical systems analysis is still feasible with the use of advanced computer algorithms

(Manu et al., 2009: Jaeger and Sharpe, 2014).[22] Basing the phase space analysis on empirical data also enables a comparison of theoretical predictions to further experimental measures. For instance, predictions of how changes in regulatory patterns are associated with expression patterns can be tested through systematic manipulation of gene activities through RNA interference techniques that increase or eliminate gene expression (Figure 3).

(How) can the approach provide insight to evolutionary change in systems parameters whereby the developmental system is itself modified? For this purpose, the group combines the reverse engineering of phase space portraits with data-constrained evolutionary simulations. They use the phase portraits of closely related organisms or species with conserved and diverged traits as starting- and end points of the simulations to identify changes in the developmental system. Thus, it is an explicit goal of the integrative approach to study how developmental potentials affect evolutionary processes through diachronic changes in the scope of realizable relations between GRNs and system capacities. The conceptual framework of dynamical systems theory is well suited for this purpose because it illustrates why many changes will not significantly affect the phenotypic development observed (the system will converge to similar attractor states from a range of initial conditions and parameter values). In the same systems some changes, e.g. development of auto-activation of genes, can lead to dramatic changes as a system approaches a bifurcation, where basins of attraction are annihilated or new ones are created (Jaeger et al., 2012). Accordingly, the conceptual approach of attractors and bifurcations may help clarify why the genotype-phenotype map is highly non-linear and why some evolutionary events are not gradual but punctuated. Depending on the phase space portrait of specific lineages, species with similar expression profiles in one environmental setting may react differently to new environmental or genetic changes because their systems may differ in the 'hidden variables' that contribute to the evolvability of the system (Jaeger and Crombach, 2012). The analysis can therefore reveal which phenotypic transitions are possible for specific lineages if parameters of the system change.

The example shows that reverse engineering may play an integrative role in bridging between experimental practices and theoretical approaches that study development through phase space geometry. The approach does not draw on a view of living systems as pre-designed or programmed by their genomes. On the contrary, the aim of the analysis is to uncover the causal basis for the two-way degeneracy of the evolving genotype-phenotype map. Combining the different methods enables the researchers to alternate between the different levels of analysis, from general and abstract conceptions of design principles to specific causal mechanisms. Neither of these can stand alone. Insights to the higher-level dynamics of phenotypic patterning are relevant for understanding how the developmental repertoire of lineages enables the production of stable and discrete phenotypes (Jaeger and Sharpe, 2014). But gaining mechanistic details about specific developmental systems is crucial for

---

[22] Some analyses also involve evaluation of dynamic patterns in three-dimensional configuration spaces where the response of developmental transitions to variation can be compared and classified based on similarities or differences of the qualitative properties of arrangement of attractors and bifurcations. This approach has recently been used to investigate whether mechanisms involved in digit patterning in vertebrates and stripe formation are similar or form separate clusters of processes in phase space (reviewed in Jaeger and Sharpe, 2014).

constructing and validating models of developmental dynamics in real-world systems.[23] The example illustrates a productive role of reverse engineering that goes beyond reductionism, and demonstrates a regulative role of design principles to clarify general constraints on phenotypic variation. This approach is therefore radically different from (naïve) reverse engineering that is rightfully criticized by Braun and Marom.

6. Conclusion

The fact that living systems are functionally flexible and adaptive has consequences for the prospects of naïve reverse engineering in biology. As we have seen in Section 2, pattern detection in regulatory networks involves an inherent risk of inferring design principles that upon closer inspection only reflect a snap-shot of the functional organization of the system or, even worse, just reflect inherent biases of decomposition strategies. I have discussed the prospects of reverse engineering in systems biology in comparison to challenges highlighted by Braun and Marom. While acknowledging the problematic aspects of naïve reverse engineering, I take issue with the conflation of reverse engineering and naïve reductionism.

First, I have argued that the challenge of synchronic underdetermination may be dealt with if the simplicity of the hypothesized design principles is more systematically distrusted through calibration with finer-grained experimental techniques. This solution may however not suffice for the challenge of diachronic underdetermination. I have discussed how heuristics relying on assumptions of static design and programmatic relations between genotypes and phenotypes provide limited insights to phenomena such as robustness, development and evolvability. But rather than dismissing reverse engineering, I have exemplified how reverse engineering can be productively combined with a systemic approach. The advantage of reverse engineering methods in the example discussed in Section 5 is that it allows researchers to mathematically fit system parameters that are too costly or impossible to obtain experimentally but nevertheless are necessary to give theoretical analysis an empirical basis. Thereby, reverse engineering of more dynamic gene regulatory networks provides a framework for integrating theoretical and experimental approaches.

In this context, the notion of design principles is a regulative abstraction that illuminates how different mechanisms relate to general types due to common physical, functional and developmental constraints on the dynamics of the system. This view differs considerably from the view of design principles as static end-points of evolution or the surprisingly persistent idea of the genome as a program. Yet, it is important to stress that methodological power is not necessarily correlated with 'true' assumptions. Completely unbiased methodologies are not viable options because efficient heuristics must reduce the complexity of a problem to make this tractable for analysis. But just as all handicraft tools have restricted scopes of applicability, so can different types of heuristics be productive or counterproductive for addressing specific questions. Machine-inspired heuristics may serve as efficient tools for (near)decomposable systems, or as negative analogies, but face limitations if the aim is to understand the nonlinear relations in evolving genotype-phenotype maps. A systemic approach accounting for the concerted changes in expression levels of the network over time can better clarify aspects of the degeneracy, flexibility and evolvability of living systems. On

---

[23] For further discussion on what insights the mechanistic and DS theoretical view afford, and the prospects and challenges for integration, see (Green, Fagan and Jaeger, 2015).

the other hand, the dynamical systems approaches relies on a reduction of developmental processes to trajectories in phase space and is not geared for uncovering the detailed biochemical composition of specific molecular regulators. A combinatorial strategy therefore seems to be the most feasible approach.

Integrating different methodological strategies is necessary because no single framework can uncover all relevant aspects of living systems. Modeling all measurable details of organisms does not necessarily provide the information needed for understanding how the system is functionally organized. Whereas reverse engineering of network topology risks oversimplifying the system of study, the pre-occupation with details runs the opposite risk of overcomplicating the task by overlooking common dynamic patterns (Gross, 2013). In the words of Marom: 'One good reason to bother with formulation of an abstract model is the hope that it leads up to a mathematical construct that dramatically reduces the dimensionality of the problem at hand. This, in turn, allows for analyses and uncovering of principles that are otherwise masked by the details' (Marom, 2010, p. 26). Thus, the task seems to be to find a way to productively combine approaches that lie between a distorting reduction of complexity and unfruitful reproduction of complexity. Discussion among biologists and philosophers is one way by which we may increase the awareness of our guiding assumptions. But progress requires that the evaluation of research methodologies goes beyond an analysis of the truth value of guiding assumptions, to study the productivity and pitfalls of epistemic tools used in practice.

Acknowledgements

References

Alon, U. (2007a). *An introduction to systems biology: Design principles of biological circuits*. Boca Raton: Chapman and Hall.

Alon, U. (2007b). Network motifs: theory and experimental approaches. *Nature Reviews Genetics, 8*, 450-481.

Alon, U. (2007c). Simplicity in biology. *La Nature, 446*(7135), 497.

Arita, M. (2004). The metabolic world of Escherichia coli is not small. *Proceedings of the National Academy of Sciences of the United States of America, 101*(6), 1543-1547.

Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J., & Blom, J. (2009). Systems biology: parameter estimation for biochemical models. *FEBS Journal, 276*(4), 886-902.

Barabási, A., & Oltvai, Z. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics,* 5(2), 101-113.

Bechtel, W. (2015). Can mechanistic explanation be reconciled with scale-free constitution and dynamics? *Studies in History and Philosophy of Biological and Biomedical Science* (this issue)

Bechtel, W., & Richardson, R. C. (1993/2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton, New Jersey: Princeton University Press.

Braillard, P. (2010). Systems biology and the mechanistic framework. *History and Philosophy of the Life Sciences,* 32, 43-62.

Braun, E., & David, L. (2011). The role of cellular plasticity in the evolution of regulatory novelty. In S. B. Gissis, & E. Jablonka (Eds.), *Transformations of Lamarckism* (pp. 181-192). Cambridge, MA: MIT Press.

Braun, E. & Marom, S. (2015). Universality, Complexity and the Praxis of Biology: Two Case Studies. *Studies in History and Philosophy of Biological and Biomedical Sciences* (this issue).

Calcott, B. (2014). Engineering and evolvability. *Biology & Philosophy, 29*(3), 293-313.

Carusi, A., Burrage, K., & Rodríguez, B. (2012). Bridging experiments, models and simulations: an integrative approach to validation in computational cardiac electrophysiology. *American Journal of Physiology - Heart and Circulatory Physiology, 303*(2), H144-H155.

Chikofsky, E., & Cross, J. (1990). Reverse engineering and design recovery: A taxonomy. *Software, IEEE, 7*(1), 13-17.

Csete, M., & Doyle, J. (2002). Reverse engineering of biological complexity. *Science, 295*(5560), 1664-1669.

David, L., Stolovicki, E., Haziz, E., & Braun, E. (2010). Inherited adaptation of genomic-rewired cells in response to a challenging environment. *HFSP Journal, 4*(3), 131-141.

Eldar, A., Dorfman, R., Weiss, D., Ashe, H., Shilo, B., & Barkai, N. (2002). Robustness of the BMP morphogen gradient in Drosophila embryonic patterning. *Nature,* 419(6904), 304-308.

Espinosa-Soto, C., Martin, O., & Wagner, A. (2011). Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits. *BMC Evolutionary Biology, 11*(1), 5.

Goldstein, K. (1934). *Der Aufbau des Organismus. Einführung in die Biologie unter besonderer Berücksichtigung der Erfahrunen am kranken Menschen*. Den Haag: Nijhoff.

Goodwin, B. (1982). Development and evolution. *Journal of Theoretical Biology, 97*, 43-55.

Goodwin, B. (1994). *How the leopard changed its spots. The evolution of complexity*. London: Phoenix.

Gould, S. J. (1977/2006). Size and Shape. In P. McGarr, & S. Rose (Eds.), *The richness of life. The essential Stephen Jay Gould* (pp. 317-323). London: Vintage Books.

Green, S. (2014). A philosophical evaluation of adaptationism as a heuristic strategy. *Acta Biotheoretica*, *64*(4), 479-498.

Green, S., Fagan, M., & Jaeger, J. (2015). Explanatory integration challenges in evolutionary systems biology, *Biological Theory*, 10(1), *18-35.*

Green, S., Levy, A., & Bechtel, W. (2015). Design sans adaptation. *European Journal of Philosophy of Science*, *5*(1), 15-29.

Gregoriou, G., Gotts, S., & Desimone, R. (2012). Cell-type-specific synchronization of neural activity in FEF with V4 during attention. *Neuron, 73*(3), 581-594.

Gross, F. (2013). *The sum of the parts: Heuristic strategies in systems biology.* PhD thesis in Foundations of the Life Sciences and Their Ethical Consequences (FOLSATEC), European School for Molecular Medicine (SEMM) and the University of Milan, Italy.

He, M., Liu, Y., Wang, X., Zhang, M., Hannon, G., & Huang, Z. (2012). Cell-type-based analysis of microRNA profiles in the mouse brain. *Neuron, 73*(1), 35-48.

Hesse, M. (1963). *Models and analogies in science*. New York: Sheed and Ward.

Ingram, P., Stumpf, M., & Stark, J. (2006). Network motifs: structure does not determine function. *BMC Genomics, 7*(1), 108.

Jaeger, J., & Sharpe, J. (2014). On the concept of mechanism in development. In A. Minelli, & T. Pradeu (Eds.), *Towards a theory of development*. Oxford, Oxford University Press, pp. 56-78.

Jaeger, J., & Crombach, A. (2012). Life's attractors. Understanding developmental systems through reverse engineering and *in silico* evolution. In O. Soyer (Ed.), *Evolutionary Systems Biology* (pp. 93-119). London: Springer.

Jaeger, J., & Monk, N. (2010). Reverse engineering of gene regulatory networks. In Lawrence et al. (Ed.), *Learning and inference in computational systems biology* (pp. 9-34). Cambridge, MA: MIT Press.

Jaeger, J., & Monk, N. (2013). Keeping the gene it its place. In D. Lambert, C. Chetland & C. Millar (Eds.), *The intuitive way of knowing. A tribute to Brian Goodwin* (pp. 153-189). Glasgow: Floris Books.

Jaeger, J., Irons, D., & Monk, N. (2012). The inheritance of process: a dynamical systems approach. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 318*(8), 591-612.

Final version published in *Studies in History and Philosophy of Biological and Biomedical Sciences:*
http://www.sciencedirect.com/science/article/pii/S1369848615000618

Karr, J., Sanghvi, J., Macklin, D., Gutschow, M., Jacobs, J., Bolival, B., . . . Covert, M. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell, 150*(2), 389-401.

Katzir, Y., Stolovicki, E., Stern, S., & Braun, E. (2012). Cellular plasticity enables adaptation to unforeseen cell-cycle rewiring challenges. *PloS One, 7*(9), e45184.

Kauffman, S. (1993). *Origins or order in evolution: Self-organisation and selection*. New York: Oxford University Press.

Keller, E. F. (2005). Revisiting "scale-free" networks. *BioEssays, 27*(10), 1060-1068.

Kermany, E., Gal, A., Lyakhov, V., Meir, R., Marom, S., & Eytan, D. (2010). Tradeoffs and constraints on neural representation in networks of cortical neurons. *The Journal of Neuroscience, 30*(28), 9588-9596.

Knuuttila, T., & Loettgers, A. (2013). Basic science through engineering? Synthetic modeling and the idea of biology-inspired engineering. *Studies in History and Philosophy of Biological and Biomedical Sciences, 44*(2), 158-169.

Krohs, U. (2009). The cost of modularity. In U. Krohs, & P. Kroes (Eds.), *Functions in Biological and Artificial Worlds: Comparative Philosophical Perspectives* (pp. 259-276). Cambridge, MA: The MIT Press.

Krohs, U. (2010). Epistemic consequences of two different strategies for decomposing biological networks. *EPSA Philosophical Issues in the Sciences* (pp. 153-162) Springer.

Krohs, U. (2012). Convenience experimentation. *Studies in History and Philosophy of Biological and Biomedical Sciences, 43(1), 52-57.*

Krohs, U. (2015). Can functionality in evolving networks be explained reductively? *Studies in History and Philosophy of Biological and Biomedical Sciences* (this issue).

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lazebnik, Y. (2002). Can a biologist fix a radio?—Or, what I learned while studying apoptosis. *Cancer Cell, 2*(3), 179-182.

Levy, A., & Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of Science, 80*, 241-261.

Mangan, S., Zaslaver, A., & Alon, U. (2003). The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *Journal of Molecular Biolog*y, 334(2), 197-204.

Manu, S., Surkova, S., Spirov, A. V., Gursky, V. V., Janssens, H., Kim, A., . . . Reinitz, J. (2009). Canalization of gene expression and domain shifts in the Drosophila blastoderm by dynamical attractors. *PLoS Computational Biology, 5*(3), e1000303.

Marom, S., Meir, R., Braun, E., Gal, A., Kermany, E., & Eytan, D. (2009). On the precarious path of reverse neuro-engineering. *Frontiers in Computational Neuroscience, 3*(5), 1-4.

Marom, S. (2010). Neural timescales or lack thereof. *Progress in Neurobiology, 90*(1), 16-28.

Nersessian, N. J. (2008). Representation and reasoning: Analogy, imagery, thought experiment. In N. J. Nersessian (Ed.), *Creating scientific concepts* (pp. 131-181), Cambridge, MA: The MIT Press.

Nicholson, D. J. (2013). Organisms ≠ Machines. *Studies in History and Philosophy of Biological and Biomedical Sciences, 44*(4), 669-678.

O'Connor, D., Huber, D., & Svoboda, K. (2009). Reverse engineering the mouse brain. *Nature, 461*(7266), 923-929.

Oster, G., & Alberch, P. (1982). Evolution and bifurcation of developmental programs. *Evolution, 36*(3), 444-459.

Pauwels, E. (2013). Communication: Mind the metaphor. *Nature, 500*(7464), 523-524.

Prill, R. J., Marbach, D., Saez-Rodriguez, J., Sorger, P. K., Alexopoulos, L. G., Xue, X., . . . Stolovitzky, G. (2010). Towards a rigorous assessment of systems biology models: the DREAM3 challenges. *PloS One, 5*(2), e9202.

Schierwagen, A. (2012). On reverse engineering in the cognitive and brain sciences. *Natural Computing, 11*(1), 141-150.

Shinar, G., & Feinberg, M. (2011). Design principles for robust biochemical reaction networks: What works, what cannot work, and what might almost work. *Mathematical Biosciences*, 231(1), 39-48.

Simon, H. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*(6), 467-482.

Simon, H. (1966). Scientific discovery and the psychology of problem solving. In R. Colodny (Ed.), *Rational Choice and the Structure of the Environment* (pp. 286-303). Pittsburgh: University of Pittsburgh Press.

Steinacher, A., & Soyer, O. (2012). Evolutionary principles underlying structure and response dynamics of cellular networks. In O. Soyer (Ed.), *Evolutionary Systems Biology* (pp. 225-247). London: Springer.

Stolovicki, E., & Braun, E. (2011). Collective dynamics of gene expression in cell populations. *PloS One, 6*(6), e20530.

Stolovicki, E., Dror, T., Brenner, N., & Braun, E. (2006). Synthetic gene recruitment reveals adaptive reprogramming of gene regulation in yeast. *Genetics, 173*(1), 75-85.

Stolovitzky, G., Monroe, D., & Califano, A. (2007). Dialogue on Reverse-Engineering Assessment and Methods, The DREAM of High-Throughput Pathway Inference. *Annals of the New York Academy of Sciences, 1115*, 1-22.

Tyson, J., Chen, K., & Novák, B. (2003). Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Current Opinion in Cell Biology 15 (2):221-231.

Tyson, J., & Novák, B. (2010). Functional motifs in biochemical reaction networks. *Annual Review of Physical Chemistry*, 61, 219-240.

Wall, M., Dunlop, M., & Hlavacek, W. (2005). Multiple functions of a feed-forward-loop gene circuit. *Journal of Molecular Biology*, 349, 501-514.

Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Paris: MIT Press.

Wimsatt, W. C. (2007). *Re-engineering philosophy for limited beings: piecewise approximations to reality*. Cambridge, MA: Harvard University Press.

Wolkenhauer, O., & Green, S. (2013). The search for organizing principles as a cure against reductionism in systems medicine. *FEBS Journal Minireview, 280*(23), 5938-5948.

Wouters, A. (2007). Design explanation: determining the constraints on what can be alive. *Erkenntnis*, 67, 65-80.