

# The Epistemology of Causal Selection: Insights from Systems Biology

Beckett Sterner  
bsterner@uchicago.edu  
NSF Postdoc, Field Museum

DRAFT: DO NOT CITE WITHOUT PERMISSION  
*Accepted for publication in Causal Reasoning in Biology,  
Minnesota Studies in Philosophy of Science*

## Abstract

Among the many causes of an event, how do we distinguish the important ones? Are there ways to distinguish among causes on principled grounds that integrate both practical aims and objective knowledge? Psychologist Tania Lombrozo has suggested that causal explanations “identify factors that are ‘exportable’ in the sense that they are likely to subserve future prediction and intervention” (Lombrozo 2010, 327). Hence portable causes are more important precisely because they provide objective information to prediction and intervention as practical aims. However, I argue that this is only part of the epistemology of causal selection. Recent work on portable causes has implicitly assumed them to be portable within the same causal system at a later time. As a result, it has appeared that the objective content of causal selection includes only facts about the causal structure of that single system. In contrast, I present a case study from systems biology in which scientists are searching for causal factors that are portable across rather than within causal systems. By paying careful attention to how these biologists find portable causes, I show that the objective content of causal selection can extend beyond the immediate systems of interest. In particular, knowledge of the evolutionary history of gene networks is necessary for correctly identifying causal patterns in these networks that explain cellular behavior in a portable way.

**Keywords:** systems biology, causal selection, motif, explanation, mechanism, gene network, null model, top-down method, teleology, contingency.

## 1. Introduction

Among the many causes for an event, how do we distinguish the important ones? Finding principled grounds for drawing this distinction is the problem of causal selection. When we ask for a causal explanation, for instance, we are interested in why one thing happened rather than something else, and we expect the answer to feature the relevant, objective causal relationships between events. The notion of importance here is partly interest-relative but also depends on objective facts about the world (Woodward 2011). Psychologist Tania Lombrozo has recently suggested that causal explanations “identify factors that are ‘exportable’ in the sense that they are likely to subserve future prediction and intervention” (Lombrozo:2010gz p.327; see also Hitchcock 2012; Lombrozo and Carey 2006). Hence portable causes are more important precisely because they provide objective information to prediction and intervention as practical aims.

However, I argue that this is only part of the epistemology of causal selection. Finding portable causes may offer a principled ground for distinguishing among causes, but the goal itself does not tell us how to find portable causes. When we consider this other dimension of the problem, it becomes clear that the recent literature on causal selection harbors a crucial ambiguity.

In specific, is the aim to predict and intervene on the same system later in time or a different system of the same kind? In the first case, the portability of a cause can be understood in terms of its insensitivity to background conditions (Hitchcock 2012; Woodward 2006). That is, the cause would be an important part of an explanation because it will behave the same way in that system even if other things have changed. The second case is more complex. As an example, imagine that we find some species of bacteria that tends to move toward higher concentrations of

food over time. Biologists call this behavior chemotaxis. Suppose that we also find that some kind of component in the cell, a signal transducer, affects chemotaxis and is insensitive to background conditions. What must be true for the causal relationship between the transducer and chemotaxis in this species to export to transducers and chemotaxis in new species? It cannot be the insensitivity of the relationship between transducer X and chemotaxis in the first species, because this alone fails to even imply a causal relationship between X and chemotaxis in the new species. Something more is needed than knowledge about the first species and its characteristic causal structure.

What we need to know will also depend on the nature of the causal systems themselves. Why would it be the case that a causal relationship in one system generalizes to other instances of the kind? For one thing, it will matter whether the causal relationship is a necessary or contingent feature of the systems we generalize over. If it is contingent, then historical accidents are possible that would introduce noise into the process of selecting portable causes across a large domain. Separating signal from noise often requires additional knowledge about why the noise occurs that is only indirectly relevant to our goal of prediction and intervention. As we will see, this matters crucially for causal selection in systems with a teleological history where our goal is to identify factors responsible for satisfying imposed functional constraints.

Furthermore, scientists rely on multiple strategies for identifying exportable causal factors. For example, biologists often distinguish between bottom-up and top-down research strategies: bottom-up methods proceed inductively from individual cases to more general classes, while top-down methods presuppose a universal theory that is specialized and tested against individual cases (Boogerd et al. 2007). The two methods differ crucially in how and when they define the domain over which portability should hold. A top-down approach must define its

domain in advance in order to acquire potential empirical content, while a bottom-up approach can specify domains a posteriori based on observed patterns.

This paper presents a case study that demonstrates why these each of these underanalyzed factors matter for studying causal selection: dimension of portability, nature of the causal system, and practical methodology. I argue that these factors are not always separable and can interact to shape what objective information is needed to carry out causal selection correctly. In particular, I show that knowledge of current causal structure in a system is sometimes insufficient: we may also need to know about the causal processes that produced the system of interest.

The case study focuses on a recent research program in systems biology that aims to explain kinds of cell behaviors, such as chemotaxis, in terms of patterns of internal molecular interactions (Milo 2002; Alon 2007b; Alon 2007a). The biologists call these causal patterns “motifs,” and they hope to discover a small set of motifs that form the universal building blocks for cellular functions. In other words, these biologists are searching for an engineering vocabulary composed of patterns of molecular interactions in order to find “design principles” for the functional organization of cells. The project of identifying motifs has proceeded in a top-down manner using only knowledge of the contemporary causal structure of cells. Following critiques by other biologists, I argue that knowledge about the evolutionary history of cells is also necessary for correctly judging the portability of motifs. The contingent role that motifs play in causing cell behaviors therefore demands a broader range of background knowledge than previous discussions of causal selection have acknowledged. A bottom-up approach to motifs could potentially avoid this requirement, but at the cost of identifying portable causes at a slower rate.

## **2. Causal selection**

The philosophical problem of causal selection lies at the intersection of human practical interest and metaphysical reality. Given a commitment to the ontological reality of causation, what basis is there for distinguishing among the causes of an event? Are there principled grounds for causal selection that integrate practical relevance and objective knowledge? The last ten years have seen increasing interest in these question among philosophers and scientists (Hitchcock 2012; Rose and Danks 2012; Lombrozo 2011; Halpern 2008), in contrast to past skepticism among philosophers about the relevance of causal selection. With this interest have also come new approaches to investigating philosophical problems that draw evidence from what people actually think and do, including how causation figures in scientific practices. A core question for this work is to what extent people incorporate subjective and objective grounds in their distinctions among causes. The case study I present articulates new dimensions within this question by focusing on a critique of how some systems biologists have distinguished among causes in practice.

Many philosophers have argued that causal selection is a mistake because it involves drawing distinctions between metaphysically equivalent causes. David Lewis, for example, wrote,

“We sometimes single out one among all the causes of some event and call it ‘the’ cause, as if there were no others. Or we single out a few as the ‘causes’, calling the rest mere ‘causal factors’ or ‘causal conditions’... We may select the abnormal or extraordinary causes, or those under human control, or those we deem good or bad, or just those we want to talk about. I have nothing to say about these principles of invidious discrimination” (Lewis 1987, 162).

As Lewis frame the issue, there is nothing more to say about the relation between two events beyond whether it is causal or not. Causation should be one thing that holds in the same way

everywhere. Adding that a cause is a background condition draws an unjustifiable distinction. The concept of causation should be indifferent to how we use it and why we care about it.

Methodologically, philosophers working on causation several decades ago assumed that it would be possible to directly state and test such an account of causation. How people used causation in practice was irrelevant to this philosophical project. For better or worse, this approach to the problem did not succeed in producing an account that satisfied all standards. This lack of consensus has motivated new interest in studying how people actually use or think about causation. As a philosophical topic, then, causation has broadened to include metaphysical, epistemological, and methodological dimensions (c.f. Cartwright 2007). In this regard, how people select among causes for practical purposes may constrain or illuminate what causation could be.<sup>1</sup>

Recent work on causal selection has approached the problem from several directions. One angle investigates the issue in terms of philosophical intuitions about hypothetical test cases, such as causal preemption (Hall 2007; Halpern and Hitchcock 2010; Hitchcock 2007). This traditional approach has recently expanded to include eliciting the intuitions of philosophers and people more generally using surveys and experiments (Knobe and Fraser 2008; Hitchcock and Knobe 2009; Knobe 2009; Lombrozo and Carey 2006; Lombrozo 2006). Another angle is to examine the role that causation and related concepts play in scientific practices, such as experiment (Waters 2007; Woodward 2010).

The case study I present here from systems biology falls within this latter angle on causal selection, but it introduces new dimensions that were not addressed by Waters and Woodward.

---

<sup>1</sup> A number of philosophers have also endorsed the idea of pluralism about causation in one form or another. The argument I make here does not depend on a commitment to a singular or pluralistic view of causation at the metaphysical level.

Waters (2007) focused on causal selection in the setting of an individual experiment, where a scientist manipulates independent variables in the target system and measures the values of dependent variables over a set of initial conditions. Waters argued that we can draw an ontological distinction between the causes of the observed effects based on whether the causes were actually or only potentially responsible for real variation in the results. The distinction between actual and potential responsibility depends on whether the results would have changed if the independent variable were manipulated to take on different values. Selecting among causes on this basis depends on how the scientist designed the experiment as well as the objective causal processes at work.

Woodward (2010) focused on the epistemological value of causal stability, proportionality, and specificity for biological research. I briefly summarize the accounts he gives here: if we know that X causes Y, then stability refers to the range of conditions under which that relationship holds. Proportionality describes what we can think of as the looseness or slack between variation in X and variation in Y. That is, the level of resolution at which we define X and Y, including the properties we ascribe to their different states, should include all and only relevant details. Specificity is a combination of X having a fine-grained effect on Y and there being a one-to-one mapping between states of X and states of Y. In this way, manipulating X should ideally allow us to achieve any possible state of Y without any redundancy in the values of X.

The project of motif identification differs from the contexts considered by Waters and Woodward because it follows a top-down rather than bottom-up method that selects among causal patterns rather than individual causal relationships. Systems biologists studying motifs begin with a pre-defined space of possible explanatory causal patterns and select among these

options only those patterns that are wide ranging and have distinctive, context-independent behaviors. By contrast, Waters' distinction among actual and potential difference makers emerges bottom-up out of the particular set of initial conditions used in that experiment. If there are more generalized relationships between the experimental variables, these could be discovered only inductively through further application of experiments that generate actual versus potential difference makers.

In a related fashion, the project of identifying motifs aims to discover a general set of causal patterns in the molecular interactions within cells, not to sequentially describe the properties of these interactions one at a time. Stability, proportionality, and specificity do matter for identifying motifs but at a higher-order level: motifs need to be causal patterns that regularly combine to form stable, proportional, and specific mechanisms within and across gene networks. Finding motifs in the molecular interactions of a cell does not in itself add new empirical details to our knowledge of these interactions; instead, it identifies which kinds of interactions are most important in understanding the evolved organization of the cell.

Let me now give some background on systems biology in order to develop these points further.

### **3. The problem structure of systems biology**

The project of motif identification selects causal patterns that are portable across rather than within systems. In other words, the ultimate aim is not to identify sets of causal relationships within a single system that will be informative for predicting and controlling the same system at a later time, although this may be an incidental benefit. Instead, the ultimate aim is to find causal patterns that explain kinds of behaviors shared in common by many systems and thereby contribute to prediction and control of a whole range of phenomena. This goal fits within



the larger problem facing systems biology today: is there a general theory of living systems and how would one discover it? While motifs have been used to address this problem at the largest scale (Milo 2002), I will focus only on their use for theorizing about the organization of cells, which is also where most experimental work on motifs has been done. This section provides the theoretical background for motif identification, within which we can make sense of portability across systems as a concrete challenge for causal selection.

Systems biology is a complex field at the intersection of an old tradition of mathematical systems theory and a recent explosion of molecular level data about cells (Boogerd et al. 2007; O'Malley and Dupré 2005). One gloss on the field is that biologists have finally acquired enough data, on the order of all the molecular parts of a cell, that they can model whole cells at a molecular level in an empirically concrete way. Nonetheless, biologists disagree about how to decompose the structure of whole cells into molecular systems that can be used to predict, explain, and manipulate behavior. My discussion here will focus on systems biologists working with gene networks, a recent and prominent approach. Gene networks are new enough in biology that their ultimate usefulness and limitations as a mode of representation are still uncertain.

The relevance of systems biology for causal selection will therefore not be as an exemplar of scientific success or failure. Instead, the case will illustrate a kind of *problem*: what are the right wholes and parts biologists should use to find important causal explanations? Systems biology articulates this general problem in its own, characteristic way as a result of its aims, data, methods, and theoretical assumptions, among other things.

The best place to begin is hypostatizing the cell as a unit and level of analysis. The single cell is commonly conceived as the basic building block of life. However, cells are by no means causally isolated units: they are non-equilibrium systems that engage in a variety of signaling

interactions, stick to each other, get eaten, move around, etc. The strand of systems biology I focus on is committed to searching for explanations using methodological reductionism. The idea is fairly simple: take all the complex cellular behaviors we can observe, such as moving toward food in a directed fashion, and describe these behaviors as causal dispositions whose relevant properties are determined by the internal organization of the cells. In other words, hypostatize that there *is* a stable internal organization that can be abstracted from ongoing evolutionary and developmental processes. The hope is that this internal organization has a principled structure that will support general explanations of the functional organization of life.

Systems biologists working in the molecular tradition also stipulate that this reduction will be to the level of individual molecules as parts. The organization of these molecules is cashed out in terms of the collective structure of their pair-wise physico-chemical interactions. In other words, cellular organization becomes a network representing the physical interactions of all the molecules in the cell. There are various ways of describing this network. The dominant representation in systems biology today idealizes away the three dimensional location and spatial extension of the molecules in order to describe only their average concentration as kinds of molecules across the cell. The idealization also typically ignores constraints such as stoichiometry that matter more for metabolic models and less for gene regulation. The edges in the network then refer to rates of change in the concentration of these chemical species, averaged over all their individual interactions. We should expect this idealization to work, for example, in relative equilibrium situations where diffusion doesn't matter and the number of molecules of each kind is large.

The methodological reduction is hence a part-whole, interlevel reduction (Wimsatt 2006). See Figure 1 for a visualization. It asserts for the sake of research that all of the interesting

dispositions of the cell are realized by various causal structures within the molecular network. For instance, the holistic behavior of “moving in a directed fashion toward food” will get broken down into specific patterns of molecular interactions in some part of the network. If the aim is to find a theory of cells as organized systems, then the success of this reductive commitment hinges on whether one can find generalizable explanatory principles for similar behaviors across concrete networks.

Systems biologists are also typically committed to search for these explanatory principles using only the contemporary causal structure of the networks. This focuses research on the question of *how* the cells work today instead of asking *why* they work that way. The biologists want to understand a general type of cell disposition, e.g. chemotaxis, in terms of causal patterns among the interactions of molecular parts. They are not — at least initially! — inquiring into the processes that caused, organized, and maintained these parts as heritable properties in the first place.

Building a theory of cellular organization in systems biology is therefore a problem of finding systematic causal relationships between kinds of behavior at the cellular level and kinds of interactions at the molecular level. It is at once more general and more specific than demonstrating that any one kind of molecular interaction has the property of stability or specificity. Most or all of the molecular interactions should be subsumable under patterns of interactions that exhibit specificity or stability in different ways. Ideally, the unique features of each causal pattern should combine to form a general vocabulary for explaining common cell behaviors across all life. In this way, motifs form a leading proposal for a general systems theory of cells.

#### **4. Motifs and gene networks**

Before describing in technical terms what a motif is, let me first back up and talk about gene networks as a framework for representing causal structure in cells. Gene networks extend the traditional molecular theory of protein synthesis to include dynamic regulatory interactions between genes, RNA, proteins, and other molecules such as metabolites. Emmert-Streib and Glazko give a convenient definition: “A gene network is a graph whose nodes represent genes, gene products, or metabolites and edges correspond to molecular interactions which can be observed experimentally” (Emmert-Streib and Glazko 2010). The network is called a *gene* network because the effects of genes are assumed to be central to understanding changes in other molecules’ concentrations over time. See Figure 2 for an example of a gene network from *E. coli*.

In the network, each kind of molecule forms one node that also represents the concentration of that molecule across the cell. So glucose would have its own node, as would the protein DNA polymerase. The edges of the network represent directed causal interactions between kinds of molecules that increase or decrease the concentration of the target molecule. For example, an RNA molecule might bind to a gene’s promoter region and inhibit expression of its protein product. Alternatively, an enzyme might catalyze the phosphorylation of a signaling protein, changing the concentrations of the modified and unmodified versions. Gene networks typically abstract away differences in mechanisms and represent interactions generically as directed edges. Each edge of the network is also associated with a parameter describing the average rate of change the interaction produces in the concentration of the affected molecule.

The experimental data behind gene networks is a messy and complicated affair. Biologists can’t measure causal interactions directly, of course, so they have to infer them to produce gene networks. Microarrays are a common experimental technique that measure changes

in molecular concentrations over time when applied across multiple samples (for a critical review in a clinical context see Keating and Cambrosio 2012). There is a range of computer algorithms now available that compute correlations between the concentrations and use heuristics to infer genuine causal interactions (Markowitz and Spang 2007; Bansal et al. 2007). Indeed, some of them use the same Bayes networks underlying Woodward's interventionist theory of causation (e.g. Friedman et al. 2000). Nonetheless, these algorithms are heuristics and sometimes have high error rates. For our purposes we are interested only in the problem of whole-part decomposition that biologists would face if they did get good data. This problem is relevant to causal selection in all areas of biology, so I will set aside issues of data quality and proceed by just presuming that one can get reasonably accurate gene networks.

Given this framework of gene networks, we can then define motifs as distinctive patterns of edges connecting a predetermined number of nodes within a gene network. Figure 3, for example, shows all the possible motifs with three nodes (Milo 2002). In paradigmatic work, Uri Alon and collaborators developed dynamic models for one class of motifs called feed-forward loops, or FFLs (Mangan and Alon 2003; Mangan, Zaslaver, and Alon 2003; Alon 2007b). As shown in Figure 4, a three-node feed forward loop can come in eight different varieties, depending on whether the interactions cause an increase or decrease in concentration. Alon et al. have split these varieties into two groups, called coherent FFLs and incoherent FFLs, based on whether the direct effect of X on Z is in the same direction as its indirect effect via Y. In (Mangan and Alon 2003), Alon et al. used mathematical modeling to argue that all incoherent FFLs can function as accelerators for transcription response to input changes. For the type-1 coherent FFL, (Mangan, Zaslaver, and Alon 2003) offered *in vivo* experimental evidence that the motif functioned to filter out positive bursts from inputs to X while responding sensitively to

negative inputs. They called this pattern a “sign-sensitive delay element” and a “persistence detector” for positive signals to  $X$ . Following this example, motifs serve to organize the causal interactions of a gene net into a selection of circuit elements that contribute distinctive features to the cell’s overall functional dynamics.

A variety of other motifs have also turned up as important. An important example for our later discussion below is the bifan motif, which can be seen in Figure 1. The bifan has four nodes organized into two pairs. Each of the upstream pair has a causal effect on both nodes in the downstream pair. One can distinguish bifans into coherent and incoherent types in a parallel way as FFLs based on whether the two upstream nodes regulate each downstream node in the same way. (One can also have partial incoherence here if one downstream node is regulated coherently but the other is not.) Bifans can be generalized topologically to involve an indefinite number of nodes arranged on two tiers with dense connecting edges. Alon et al. call this generalized motif a dense overlapping regulon (DOR), where a regulon is a set of genes regulated by a single transcription factor (Shen-Orr et al. 2002; Alon 2007a). Finding generalized motifs in gene networks involves a somewhat different technical procedure since the number of nodes is not fixed in advance (Shen-Orr et al. 2002).

Obviously there are many more motifs one could find in gene networks. The space of motifs grows exponentially with the number of nodes, and simply enumerating all the nodes of size ten in some gene network is almost prohibitively expensive in computing resources. In order to select just those motifs that are generally valuable for explaining cell behavior, Alon et al. have imposed several additional criteria. FFLs and other important motifs are supposed to have fixed causal dispositions no matter the structure of the larger network. In other words, input levels to the motifs may vary, but the mapping each motif establishes between inputs and outputs

should not. Which genes or molecules realize the motif as a causal pattern and what the rate constants are for the reactions involved should be less important than the formal structure of the motif in determining relevant causal effects. For example, variations in the rate constants might modulate the behavior of the motif, accelerating its dynamics or minimizing absolute changes in concentrations, but the structure of the mathematical mapping between input and output states should remain invariant.

However, it is impossible in practical terms to select motifs one by one based on whether they exhibit distinctive functions based on context-independent causal mappings. Instead, Alon et al. apply a prior condition: the motif must be statistically enriched in a gene network compared to a background distribution. Only if this holds is it worth investigating the motif's causal structure in detail. In the next section, I describe the process of selecting among motifs as applying an iterative sequence of filters that winnow down the initial, huge space of possibilities.

## **5. The search for motifs**

The idea that motifs are the engineering vocabulary of cellular structure is a hypothesis. In order to test the claim's adequacy, one must go to experimental data about actual gene networks. By definition, any gene network will contain motifs: they are minimally just formal patterns of causal interactions after all. The question is whether the behavior of cells is best explained in terms of these patterns and no others. In order to answer this, systems biologists must choose a practical method for ascertaining whether any motifs do indeed meet their explanatory aims. As I already noted, Alon et al. have chosen a top-down method that uses a statistical procedure for selecting among motifs.

The promise of motif identification as a research program is that it would deliver a big theoretical payoff using an efficient research strategy. In order to achieve both requirements,

motifs need to meet several important conditions, which I will list here. The most important methodological constraints are: 1) it must be possible to identify a motif by a low-cost diagnostic character, in this case the formal pattern of nodes and edges in the network, and 2) it must be possible to test the explanatory importance of a motif without needing to know how it affects the surrounding gene network. As I noted, there are too many motifs for systems biologists to first characterize each motif in terms of its causal behaviors under all conditions and only then turn to see which of these motifs occur in real gene networks. The other problem is that scientists typically lack empirical data about the rate parameters in gene networks. This makes it very difficult to model or simulate the precise differences any motif makes. Experimental testing of motifs is possible but again highly time intensive (Mangan, Zaslaver, and Alon 2003).

On the theoretical end, I have already discussed a couple constraints: 3) Motifs must have distinctive invariant behaviors that characterize their effects on the network. For example, coherent FFLs serve as a “sign sensitive delay element.” 4) These invariant behaviors should depend solely on the internal causal structure of the interactions in the motif, so that the effects of some motif are independent from its context in the network.

Further constraints come from the explanatory aims of motifs as a general theoretical vocabulary for systems biology: 5) There have to be motifs fitting the above four constraints that make a causal difference to the behaviors of cells across many species. For convenience, let’s call these “large-scope motifs.” One might find that coherent FFLs always figure in causal explanations of chemotaxis in single-celled microorganisms, for example. 6) It should be possible to organize the motifs fitting the first four constraints into “design vocabularies” that have a stable composition over domains of life or kinds of cellular functions. 7) For any given gene network, the set of large-scope motifs should be sufficient to cover most of the network’s



functionally important behaviors. In other words, we can use one engineering vocabulary to analyze many different design features across gene networks.

As a research program, work on motifs can be broken down along several directions. The third and fourth conditions above can be investigated for each possible motif using experiments and mathematical modeling (Mangan et al. 2006; Mangan, Zaslaver, and Alon 2003; Mangan and Alon 2003; Kremling, Bettenbrock, and Gilles 2008). The first condition, however, limits the need for this work by allowing Alon et al. to design computer programs that can efficiently identify and count motifs within and across gene networks. It is in this filtering step that we will see how the history of the gene network becomes relevant for causal selection.

The short answer for how this first step works is statistics: given all the modules or motifs that can be enumerated in a network, are there statistically significant patterns in how they occur? The long answer requires specifying what counts as “statistically significant.” In particular, it involves designing and validating a null model against which the occurrence of some motifs will stand out as important. I want to emphasize that statistics is not serving here to discover causal relationships, since these have already been determined experimentally and are given in the structure of the gene network. (So I have granted for the sake of argument.) The role of the null model is solely to select important patterns of causal relationships from those interactions already given.

The overall search for motifs thus proceeds as follows: given an experimentally determined gene network, measure the frequency of different motifs of a particular size occurring in the network. Randomize the network many times under certain constraints and calculate a null distribution using these randomized networks for the background frequencies of each motif. Calculate the statistical likelihood of the actually observed motif frequencies given

the null distribution. Then, any motifs that stand out as significant will be candidates for further examination using mathematical modeling and experiments to study their internal causal structure and dynamic behavior. Once a set of motifs have been found to be statistically overrepresented and to possess the appropriate causal structure, one can look for general “design principles” that use motifs to explain cellular behaviors (e.g. Milo, Itzkovitz, Kashtan, Levitt, Shen-Orr, et al. 2004).

The null model is computed using random re-arrangements of the actual, experimentally-determined network. As Kashtan et al. (2004) put it: “There are therefore two main tasks in detecting network motifs: (1) generating an ensemble of proper random networks... and (2) counting the subgraphs [motifs] in the real network and in random networks.” The second task is primarily a challenge of efficiently estimating motif frequencies in networks and has little biological interest.<sup>2</sup> I will therefore focus on the procedure for generating the random networks.

The biological content of the null model is implicit in the constraints placed on the randomization process. One wants to compare the actually observed frequency of motifs with frequencies in a population of similar but different networks. What properties of the actual network are held invariant under randomization affects the distribution of frequencies in the simulated population. Due in part to limitations of knowledge and computing power, systems biologists have chosen to preserve only certain formal (topological) properties of the network, such as the number of incoming and outgoing edges at each node. The computational procedure for the randomization involves probabilistically swapping edges between pairs of nodes, so that if node A points to B and node C to D, then the edges are switched so A points to D and C to B

---

<sup>2</sup> The complexity of the counting problem increases prohibitively with the dimensions of the graph unless one can find statistical sampling heuristics that estimate the true value without needing to count every motif exhaustively.

(Shen-Orr et al. 2002). The networks are therefore similar in the degree of connectedness of their nodes but different with regard to how they are connected.

Given this, the top-down character of motif identification derives from several features. First, the search process presupposes a universally sufficient space of possible causal patterns for explaining cell behaviors. Second, it selects important patterns from this space using a series of progressively more demanding tests that apply uniformly to each motif.<sup>3</sup> Third, the structure of the target explanations are largely fixed in advance: the aim is to explain kinds of cell behaviors in terms of statistically unlikely motifs and their combined causal effects. In a bottom-up approach, each of these dimensions would be left open-ended to be determined by inductive evidence and abductive inferences as research progressed. For instance, one would not presume that small parts of gene networks with fixed internal structure would always be explanatorily sufficient, and the domain of generalization for some causal pattern would be determined a posteriori rather than in advance. To be sure, the contrast between top-down and bottom-up methods is a matter of degree rather than a strict dichotomy. Nonetheless, the way that Alon et al. select among motifs imposes and then tests a universal organization in the structure of gene networks. By contrast, a bottom-up approach would draw distinctions among causes on a local, case-by-case basis and would define domains for generalization during the course of research.

## **6. Why evolutionary history matters for the selection of motifs**

Does randomizing gene networks while preserving only the number of incoming and outgoing edges make for the right comparison of real to simulated distributions? Does this randomization select *all* and *only* those motifs that can provide a theoretical vocabulary for the

---

<sup>3</sup> Ideally, one could run the above statistical test on many gene networks simultaneously and aggregate the results. In practice, only a few high-quality gene networks are available so the testing runs in a more piece-meal fashion.

design principles of cells? Notice that the randomization does not preserve the functionality of the network. Probably all of the simulated networks would be lethal for an actual cell, since randomization does not respect the stoichiometry, thermodynamics, or kinetics of the interactions.<sup>4</sup> There is no guarantee that metabolic reactions would still flow in the correct direction, or that the signals from transduction mechanisms wouldn't get jumbled together. As a contrast class, then, these randomized networks highlight the functional organization of the actual network. The parts of the actual networks that solve functional design problems will not be preserved under randomization and should therefore turn up as significant against the null distribution.

What if the actual network became enriched in certain motifs for non-functional reasons, though? Simply re-arranging the edges wouldn't account for this inflation in the frequency of motifs in the actual network. I'll argue in this section, following critiques from biologists, that evolutionary processes affecting the past evolution of the gene network turn out to be one important source of such functionally neutral enrichment. As a result, we will see that the evolutionary contingency of the relationship between motifs and cell functions matters for the correctness of selecting just those motifs that are portable, explanatory causes across gene networks.

Alon et al.'s published method for selecting among motifs does not directly model evolutionary processes or incorporate facts about the evolutionary history of gene networks. *Prima facie*, this doesn't seem to be a problem. Motif identification proceeds by fixing its target domain of phenomena to be explained: cell behaviors. Then it searches within that domain to

---

<sup>4</sup> It also completely ignores the material parts of the cell that actually realize the network. One cannot take some kinase protein and expect it to influence the expression of any random gene, yet this is how the randomization operates at the level of the network's formal structure.

find unexpected structural patterns, i.e. statistically over-enriched motifs. These patterns are then candidates for design principles if the high frequency and location of certain motifs (with their distinctive effects) in gene networks explains how cells solve some functional problem.

However, the concept of “design principles” is grounded in the process of natural selection and the evolutionary history of gene networks. For instance, Alon et al. have modeled computationally what environmental conditions might lead to natural selection for motifs (Kalisky, Dekel, and Alon 2007; Kashtan and Alon 2005; Dekel, Mangan, and Alon 2005). In the case of the coherent feed-forward loop, they “find conditions that the environment must satisfy in order for the FFL to be selected over simpler circuits: the FFL is selected in environments where the distribution of the input pulse duration is sufficiently broad and contains both long and short pulses” (Dekel, Mangan, and Alon 2005, 81) Natural selection is thus the reason one would expect such design principles to exist, but motifs are not a necessary consequence of natural selection per se. Indeed, they can also be produced by evolutionary neutral processes such as gene duplication and genome duplication.

Perhaps the most important problem with the network randomization used by Alon et al. is that it treats nodes with the same number of incoming and outgoing edges equivalently, no matter their detailed causal role or context in the network. Superficially, this might seem to be a positive feature, since one goal of the null model is to describe the network’s structure separate from any influence of natural selection. However, the absence of selection does not imply equal probabilities of attachment between topologically similar nodes. Neutral processes of evolution can introduce significant biases to the network structure over time due to structural constraints on the variation produced by mechanisms such as mutation or gene duplication. Fairly simple models for these neutral processes have succeeded in producing the statistical signals used to

identify motifs (e.g. Solé and Valverde 2006; Lynch 2007). The randomization step in detecting motifs is supposed to focus future research on structural elements with generalizable causal effects, but neutral evolutionary processes carry important information for maximizing this projectability.

Heterogeneous background signal in gene networks can come from a variety of sources. The likelihood of simple motifs such as auto-regulation (a one-node feedback loop) depends strongly on the size of the regulatory region of DNA in front of a gene, which varies by orders of magnitude across species (Lynch 2007). Alternatively, Artzy-Randrup et al. have shown that enrichment in FFLs can be produced simply if networks grow by preferentially attaching new nodes to already highly connected nodes (Artzy-Randrup et al. 2004). This simulated, evolutionarily neutral process did not recreate the distribution of all motifs accurately (Milo, Itzkovitz, Kashtan, Levitt, and Alon 2004), but it still demonstrates that mechanisms of network evolution can confound the signal for design principles.

Empirical analyses of gene networks have also shown that motifs can be generated in aggregate groups. Figure 2 from (Dobrin et al. 2004) illustrates how the vast majority of feed-forward loop and bifan motifs in *E. coli* overlap to form “homologous motif clusters.” This clustering raises an important question about whether each occurrence of a motif can be accurately taken as an independent statistical event in the network. Ward and Thornton (2007) analyzed the effects of an ancient genome duplication event in the *Saccharomyces* clade and argued that the aggregate clusters of FFL and bifan motifs in the clade are due both to selection and to neutral duplication of genes. In particular, they find that “many of the bi-fan arrays and the motifs within them can be attributed to the [genome duplication] event that occurred recently in the evolution of *Saccharomyces*, with the overwhelming majority of these structures arising

from duplication of TFs [transcription factors]” (Ward and Thornton 2007, 1999). Moreover, the increase in bi-fan arrays appears to have facilitated the emergence of feed-forward loops connecting the bi-fans.

In other words, the enrichment of bifan and feed-forward loop motifs is not adequately explained by independent natural selection for each instance of the motifs. Instead, evolutionary mechanisms in gene networks produce motifs in a correlated, aggregate fashion, and natural selection may act to conserve or alter these chunks. One-off, historically contingent events such as genome duplications can strongly affect the future prevalence of motifs in a network. Randomize the network the wrong way and this background context will show up as a candidate design principle (or true design principles may remain hidden).

The challenge for correctly identifying important motifs, then, is to account for these background influences on the structure of gene networks. In this case, the background noise derives from historical processes acting on the systems of interest. In order to find motifs that are portable as causal explanations across systems, systems biologists must incorporate knowledge about the mechanisms driving heritable variation in gene networks. Without this knowledge, the null distribution will be biased in unpredictable ways, rendering the results of the causal selection process suspect.

## **8. Conclusion**

As a way of summarizing my argument, let me point out that the challenge I described for finding motifs in systems biology is an instance of a more general problem, which arises essentially out of a statistical conception of causal selection among systems with a teleological

history.<sup>5</sup> Imagine that one has measurements of some variable for a collection of distinct systems. Suppose that the causal structures of these systems are known, so that one can identify all the causally relevant factors in each system for the measured variable. Also assume that the structure of each system is constrained so that it must produce a common characteristic effect in the measured variable. For instance, the variable must always take on some value under certain conditions.

The problem, then, is whether there are general explanations for how the set of systems produce this effect. A top-down statistical approach would answer this question by looking for pre-determined kinds of causal factors or patterns of factors that occurred more often than expected by chance. It would then infer that these patterns are statistically enriched because they are consequences of the teleological constraint on the systems' behavior. To make this inference, however, we must also assess how often those factors or patterns could arise by mechanisms that are independent of the teleological constraint. These sources of noise could be historical, like whole genome duplications in gene networks, or they could be acting concurrently on the systems as they are measured.<sup>6</sup> In either case, accounting for these sources of noise requires additional knowledge about causal influences external to the original systems of interest.

This push beyond the causal structure of the systems themselves is a joint consequence of the underappreciated dimensions of causal selection I introduced earlier: top-down versus bottom-up methodology, an across-system versus within-system explanatory target, and a contingent versus necessary presence of the causal factors within these systems. Also important

---

<sup>5</sup> The teleology arises for motifs in systems biology because they participate in design principles that solve evolutionary problems for cells, but more broadly we can also include teleology here in the sense of systems built or manipulated according to human goals (e.g. Lombrozo and Carey 2006).

<sup>6</sup> Another source of noise would arise if the teleological constraint were only sometimes effective in guaranteeing the characteristic output effect.



in this regard is the teleological background of the causal systems of interest. Together, they motivate the use of statistical procedures as a tool for efficiently generalizing over noisy data, which imports a new issue into the process of causal selection: one must know about the noise as well as the signal.

In this way, the epistemic scope of causal selection depends on the practical choice of methods as well as the pragmatic aim. One cannot guarantee that causal selection for the sake of prediction and intervention can be done correctly using only the internal structure of the target systems. In order to say what is needed, one must fully specify both the ends and the means of the selection.

## 9. Acknowledgments

My sincerest thanks to Ken Waters and Alan Love for their ongoing interest and support while this paper went through drafts, to Scott Lidgard for his detailed comments and suggestions, to feedback from the participants of the Minnesota workshops on causal reasoning in biology in 2011 and 2012, and to Devin Gouvea, William Wimsatt, and Bill Sterner for reading and commenting on an earlier draft. This research was supported in part by an NSF graduate research fellowship and post-doctoral grant SES-1153114.

## References

- Alon, Uri. 2007a. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Vol. 10. Boca Raton, FL: CRC Press.
- Alon, Uri. 2007b. "Network Motifs: Theory and Experimental Approaches." *Nature Reviews Genetics* 8 (6) (June): 450–461. doi:10.1038/nrg2102.
- Artzy-Randrup, Yael, Sarel J Fleishman, Nir Ben-Tal, and Lewi Stone. 2004. "Comment on "Network Motifs: Simple Building Blocks of Complex Networks" and "Superfamilies of Evolved and Designed Networks"." *Science* 305 (5687) (August 20): 1107. doi:10.1126/science.1099334.
- Bansal, Mukesh, Vincenzo Belcastro, Alberto Ambesi-Impiombato, and Diego di Bernardo. 2007. "How to Infer Gene Networks From Expression Profiles.." *Molecular Systems Biology* 3: 78. doi:10.1038/msb4100120.

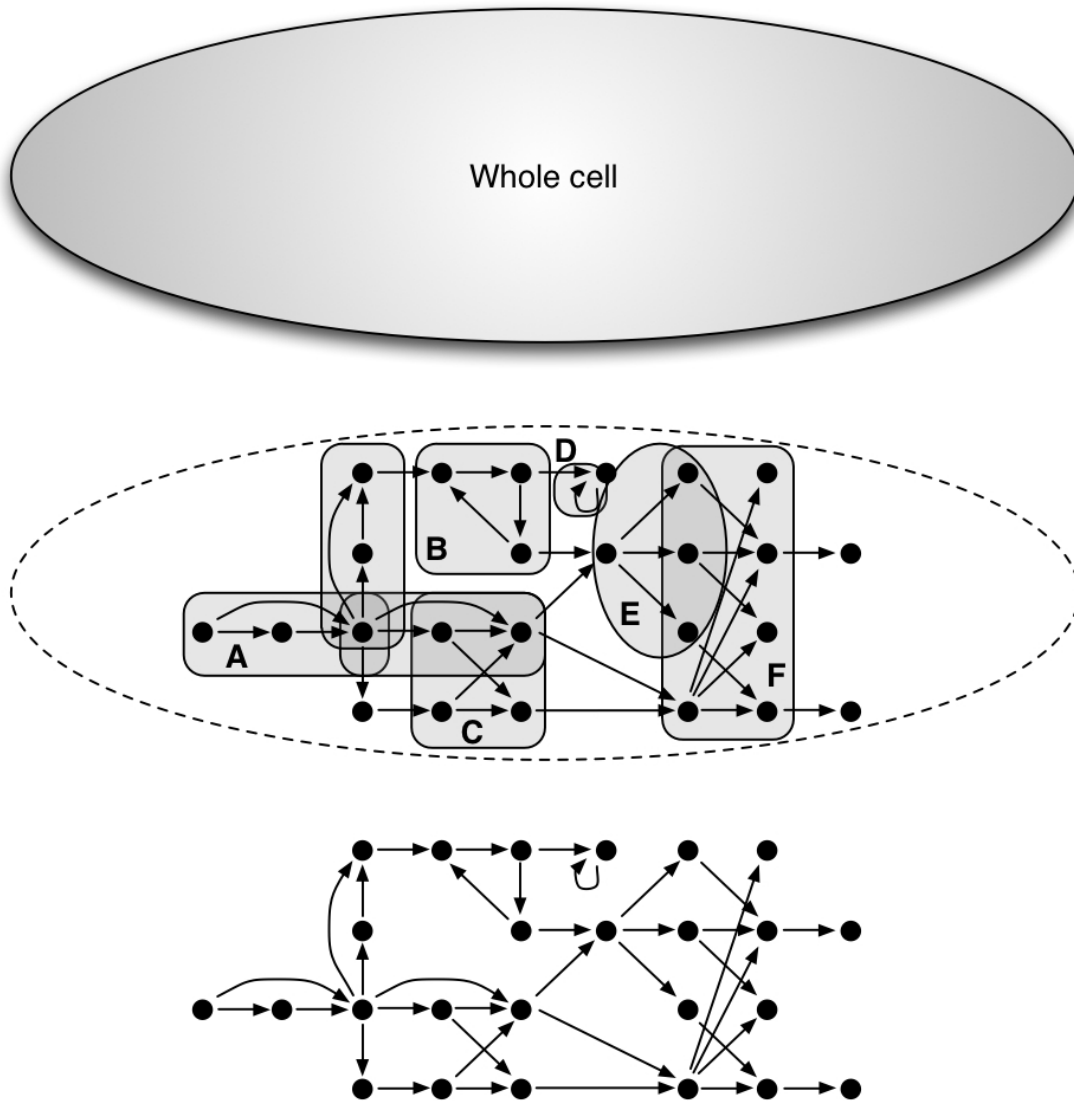
- Boogerd, Fred C, Frank J Bruggeman, Jan-Hendrik S Hofmeyr, and Hans V Westerhoff, eds. 2007. *Systems Biology: Philosophical Foundations*. New York: Elsevier.
- Cartwright, Nancy. 2007. *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press.
- Dekel, Erez, Shmoolik Mangan, and Uri Alon. 2005. "Environmental Selection of the Feed-Forward Loop Circuit in Gene-Regulation Networks." *Physical Biology* 2 (2) (June 1): 81–88. doi:10.1088/1478-3975/2/2/001.
- Dobrin, Radu, Qasim K Beg, Albert-László Barabási, and Zoltán N Oltvai. 2004. "Aggregation of Topological Motifs in the Escherichia Coli Transcriptional Regulatory Network." *BMC Bioinformatics* 5 (January 30): 10. doi:10.1186/1471-2105-5-10.
- Emmert-Streib, Frank, and Galina V Glazko. 2010. "Network Biology: a Direct Approach to Study Biological Function." *WIREs Systems Biology and Medicine* 3 (4) (December 31): 379–391. doi:10.1002/wsbm.134.
- Friedman, Nir, M Linial, I Nachman, and Dana Pe'er. 2000. "Using Bayesian Networks to Analyze Expression Data.." *Journal of Computational Biology* 7 (3-4): 601–620. doi:10.1089/106652700750050961.
- Hall, Ned. 2007. "Structural Equations and Causation." *Philosophical Studies* 132 (1) (January 13): 109–136. doi:10.1007/s11098-006-9057-9.
- Halpern, Joseph Y. 2008. "Defaults and Normality in Causal Structures." In, 198–208.
- Halpern, Joseph Y, and Christopher R Hitchcock. 2010. "Actual Causation and the Art of Modeling." In *Heuristics, Probability, and Causality: a Tribute to Judea Pearl*, edited by R Dechter, H Geffner, and Joseph Y Halpern, 383–406. London: College Publications.
- Hitchcock, Christopher R. 2007. "Prevention, Preemption, and the Principle of Sufficient Reason." *The Philosophical Review* 116 (4) (October 13): 495–532. doi:10.1215/00318108-2007-012.
- Hitchcock, Christopher R. 2012. "Portable Causal Dependence: a Tale of Consilience." *Philosophy of Science* 79 (5) (December): 942–951. doi:10.1086/667899.
- Hitchcock, Christopher R, and J Knobe. 2009. "Cause and Norm." *Journal of Philosophy* CVI (11): 587–612.
- Kalisky, Tomer, Erez Dekel, and Uri Alon. 2007. "Cost–Benefit Theory and Optimal Design of Gene Regulation Functions." *Physical Biology* 4 (4) (December 1): 229–245. doi:10.1088/1478-3975/4/4/001.
- Kashtan, Nadav, and Uri Alon. 2005. "Spontaneous Evolution of Modularity and Network Motifs." *Proceedings of the National Academic of Sciences* 102 (39) (September 27): 13773–13778. doi:10.1073/pnas.0503610102.
- Kashtan, Nadav, Shalev Itzkovitz, Ron Milo, and Uri Alon. 2004. "Efficient Sampling Algorithm for Estimating Subgraph Concentrations and Detecting Network Motifs." *Bioinformatics* 20 (11) (July 21): 1746–1758. doi:10.1093/bioinformatics/bth163.
- Keating, Peter, and Alberto Cambrosio. 2012. "Too Many Numbers: Microarrays in Clinical Cancer Research." *Studies in the History and Philosophy of Biological and Biomedical Sciences* 43 (1) (March 1): 37–51. doi:10.1016/j.shpsc.2011.10.004.
- Knobe, Joshua. 2009. "Folk Judgments of Causation." *Studies in History and Philosophy of Science* 40 (2) (June 1): 238–242. doi:10.1016/j.shpsa.2009.03.009.
- Knobe, Joshua, and Ben Fraser. 2008. "Causal Judgment and Moral Judgment: Two Experiments." In *Moral Psychology*, edited by Walter Sinnott-Armstrong. Moral psychology.

- Kremling, Andreas, Katja Bettenbrock, and E D Gilles. 2008. "A Feed-Forward Loop Guarantees Robust Behavior in Escherichia Coli Carbohydrate Uptake." *Bioinformatics* 24 (5) (February 28): 704–710. doi:10.1093/bioinformatics/btn010.
- Lewis, David. 1987. "Causation." In *Philosophical Papers*. Vol. II. New York: Oxford University Press.
- Lombrozo, Tania. 2006. "The Structure and Function of Explanations." *Trends in Cognitive Sciences* 10 (10) (October): 464–470. doi:10.1016/j.tics.2006.08.004.
- Lombrozo, Tania. 2010. "Causal-Explanatory Pluralism: How Intentions, Functions, and Mechanisms Influence Causal Ascriptions." *Cognitive Psychology* 61 (4) (December): 303–332. doi:10.1016/j.cogpsych.2010.05.002.
- Lombrozo, Tania. 2011. "The Instrumental Value of Explanations." *Philosophy Compass* 6 (8) (August 8): 539–551. doi:10.1111/j.1747-9991.2011.00413.x.
- Lombrozo, Tania, and Susan Carey. 2006. "Functional Explanation and the Function of Explanation." *Cognition* 99 (2) (March): 167–204. doi:10.1016/j.cognition.2004.12.009.
- Lynch, Michael. 2007. "The Evolution of Genetic Networks by Non-Adaptive Processes." *Nature Reviews Genetics* 8 (10) (October): 803–813. doi:10.1038/nrg2192.
- Mangan, Shmoolik, Alon Zaslaver, and Uri Alon. 2003. "The Coherent Feedforward Loop Serves as a Sign-Sensitive Delay Element in Transcription Networks." *Journal of Molecular Biology* 334 (2) (November 21): 197–204.
- Mangan, Shmoolik, and Uri Alon. 2003. "Structure and Function of the Feed-Forward Loop Network Motif." *Proceedings of the National Academic of Sciences* 100 (21) (October 14): 11980–11985. doi:10.1073/pnas.2133841100.
- Mangan, Shmoolik, Shalev Itzkovitz, Alon Zaslaver, and Uri Alon. 2006. "The Incoherent Feed-Forward Loop Accelerates the Response-Time of the Gal System of Escherichia Coli." *Journal of Molecular Biology* 356 (5) (March 10): 1073–1081. doi:10.1016/j.jmb.2005.12.003.
- Markowitz, Florian, and Rainer Spang. 2007. "Inferring Cellular Networks—a Review." *BMC Bioinformatics* 8 Suppl 6: S5. doi:10.1186/1471-2105-8-S6-S5.
- Milo, Ron. 2002. "Network Motifs: Simple Building Blocks of Complex Networks." *Science* 298 (5594) (October 25): 824–827. doi:10.1126/science.298.5594.824.
- Milo, Ron, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, and Uri Alon. 2004. "Response to Comment on "Network Motifs: Simple Building Blocks of Complex Networks" and 'Superfamilies of Evolved and Designed Networks'." *Science* 305 (5687) (August 20): 1107d–1107d. doi:10.1126/science.1100519.
- Milo, Ron, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai S Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, and Uri Alon. 2004. "Superfamilies of Evolved and Designed Networks." *Science* 303 (5663) (March 5): 1538–1542. doi:10.1126/science.1089167.
- O'Malley, Maureen A, and John Dupré. 2005. "Fundamental Issues in Systems Biology." *BioEssays* 27 (12): 1270–1276. doi:10.1002/bies.20323.
- Rose, David, and David Danks. 2012. "Causation: Empirical Trends and Future Directions." *Philosophy Compass* 7 (9) (August 22): 643–653. doi:10.1111/j.1747-9991.2012.00503.x.
- Shen-Orr, Shai S, Ron Milo, Shmoolik Mangan, and Uri Alon. 2002. "Network Motifs in the Transcriptional Regulation Network of Escherichia Coli." *Nature Genetics* 31 (1) (April 22): 64–68. doi:10.1038/ng881.
- Solé, Ricard V, and Sergi Valverde. 2006. "Are Network Motifs the Spandrels of Cellular Complexity?" *Trends in Ecology & Evolution* 21 (8) (August): 419–422.

doi:10.1016/j.tree.2006.05.013.

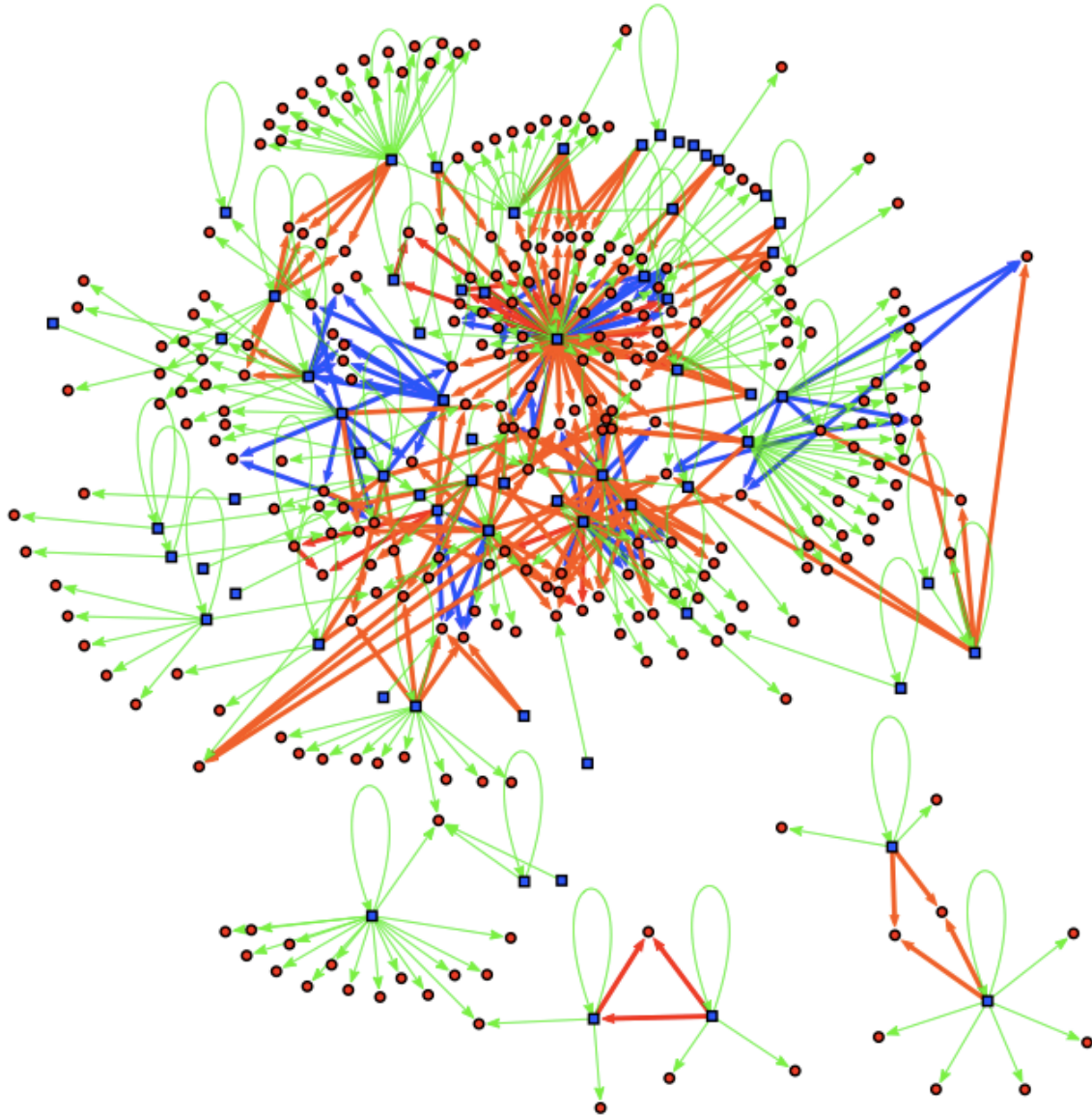
- Ward, Jonathan J, and Janet M Thornton. 2007. "Evolutionary Models for Formation of Network Motifs and Modularity in the *Saccharomyces* Transcription Factor Network." *PLoS Computational Biology* 3 (10) (October): 1993–2002. doi:10.1371/journal.pcbi.0030198.st003.
- Waters, C Kenneth. 2007. "Causes That Make a Difference." *The Journal of Philosophy* CIV (11): 551–579.
- Wimsatt, William C. 2006. "Reductionism and Its Heuristics: Making Methodological Reductionism Honest." *Synthese* 151 (3) (August 8): 445–475. doi:10.1007/s11229-006-9017-0.
- Woodward, James. 2006. "Sensitive and Insensitive Causation." *The Philosophical Review* 115 (1): 1–50. doi:10.1215/00318108-2005-001.
- Woodward, James. 2010. "Causation in Biology: Stability, Specificity, and the Choice of Levels of Explanation." *Biology & Philosophy* 25 (3) (February 6): 287–318. doi:10.1007/s10539-010-9200-z.
- Woodward, James. 2011. "Causes, Conditions, and the Pragmatics of Causal Explanation." In *Philosophy of Science Matters: the Philosophy of Peter Achinstein*, edited by Gregory J Morgan, 247–257. Oxford: Oxford University Press.

## Figures



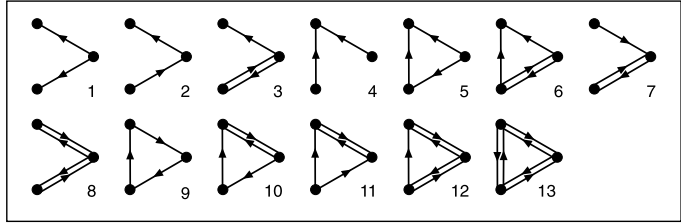
**Figure 1: Motifs as an intermediate level of structure**

Motifs are higher-order units of structure between gene networks and whole cells. A hypothetical gene network is shown on the bottom level, where nodes are kinds of molecules and edges are causal interactions affecting concentration levels of those molecules. In the middle level, these interactions are aggregated into causal patterns called motifs. (A) Three feed forward loops intersecting over the same node. (B) A three-node feedback cycle. (C) A bifan. (D) A feedback loop. (E) A single-input module. (F) A densely overlapping regulon, defined as a generalization of the bifan motif to more nodes where some edges can be missing. Note that motifs can overlap and usually do not cover the whole network structure.



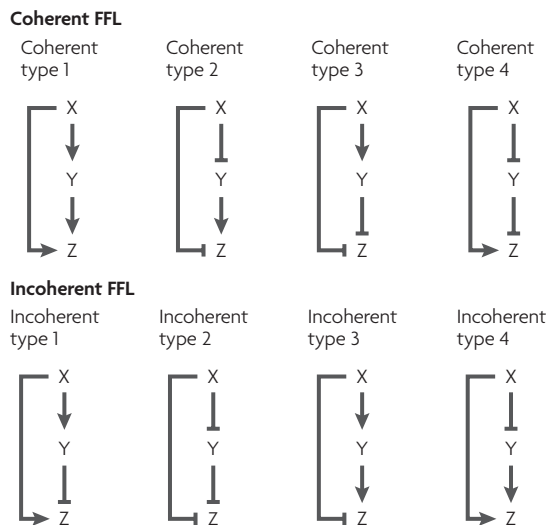
**Figure 2: The *E. coli* transcriptional regulatory network**

Illustrates a gene network specialized to only show kinds of molecules directly acting on genes to regulate expression levels. Thick lines represent edges participating in feed-forward loops and bifan motifs. Thick blue lines are edges shared between multiple motifs, while thick orange lines are edges participating in only one motif. The remaining edges in the network are shown with thin green lines. Note how the motifs form an interconnected and partly overlapping aggregate covering the main cluster of nodes in the network. From (Dobrin et al. 2004).



**Figure 3: Three-node motifs**

All thirteen possible motifs with three nodes and directed edges. From (Milo 2002).



**Figure 4: The eight possible feed-forward loops**

Note that the directed edges in this graph contain additional information compared to Figure 3. The edge with an arrow indicates that one molecule causes an increase in the other, while an edge with a line indicates that it causes a decrease. The coherent FFLs have the same sign between the direct edge from X to Z and the indirect path through Y. (Positive, negative, negative, and positive, respectively for each of the four motifs.) The incoherent FFLs have opposite signs in the two paths from X to Z. From (Alon 2007).