Network motifs: theory and experimental approaches

Uri Alon

Abstract | Transcription regulation networks control the expression of genes. The transcription networks of well-studied microorganisms appear to be made up of a small set of recurring regulation patterns, called network motifs. The same network motifs have recently been found in diverse organisms from bacteria to humans, suggesting that they serve as basic building blocks of transcription networks. Here I review network motifs and their functions, with an emphasis on experimental studies. Network motifs in other biological networks are also mentioned, including signalling and neuronal networks.

Transcription regulation networks describe the interactions between transcription factor proteins and the genes that they regulate[1–5]. Transcription factors respond to biological signals and accordingly change the transcription rate of genes, allowing cells to make the proteins they need at the appropriate times and amounts.

Recent work indicates that transcription networks contain a small set of recurring regulation patterns, called network motifs[1,6,7]. Network motifs can be thought of as recurring circuits of interactions from which the networks are built. Network motifs were first systematically defined in *Escherichia coli*, in which they were detected as patterns that occurred in the transcription network much more often than would be expected in random networks. The same motifs have since been found in organisms from bacteria[8,9] and yeast[7,10] to plants and animals[11–16]. This Review focuses on experimental studies of network motifs; a comprehensive treatment with quantitative models can be found in REF. 1.

The Review discusses two types of transcription network: sensory networks that respond to signals such as stresses and nutrients, and developmental networks that guide differentiation events. I will first consider sensory networks, the motifs of which are common to both types of network. I will then turn to motifs that are specific to developmental networks. This Review focuses on transcription networks because they are the most studied. Network motifs are also found in other biological networks, such as those that involve protein modifications or interactions between neuronal cells. I will briefly describe the motifs that are found in these biological networks.

Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel.
e-mail: urialon@weizmann.ac.il
doi:10.1038/nrg2102

The main idea that is presented in this Review is that each network motif can carry out specific information-processing functions. These functions have been analysed using mathematical models and tested with dynamic experiments in living cells. Still, there is much to be done: it is important to further experimentally test the functions that each network motif can perform. Such experiments could illuminate the dynamics of the many systems in which each motif appears. Furthermore, it is important to test whether motifs can help us to understand the densely connected networks of higher organisms.

Simple regulation

Let's begin by understanding the dynamics of a basic transcription interaction, a single arrow in the network, which is referred to here as 'simple regulation' (FIG. 1a). Simple regulation can serve as a reference for understanding the dynamic functions of network motifs. Simple regulation occurs when transcription factor Y regulates gene X with no additional interactions (FIG. 1a). Y is usually activated by a signal, S_y. The signal can be an inducer molecule that directly binds Y, or a modification of Y by a signal-transduction cascade, and so on. When transcription begins, the concentration of the gene product X rises and converges to a steady-state level (FIG. 1d). This level is equal to the ratio of the production and degradation rates, where degradation includes both active degradation and the effect of dilution by cell growth. When production stops, the concentration of the gene product decays exponentially. In both cases, the response time, which is defined as the time it takes to reach halfway between the initial and final levels, is

equal to the half-life of the gene product[17]. The faster the degradation rate, the shorter the response time. For proteins that are not actively degraded, as is the case for most proteins in growing bacterial cells, the response time is equal to one cell-generation time. This is a result of the dilution effect from cell growth.

Negative autoregulation. Negative autoregulation (NAR) occurs when a transcription factor represses the transcription of its own gene[17–19] (FIG. 1b). This network motif occurs in about half of the repressors in *E. coli*[5,17], and in many eukaryotic repressors[10]. NAR has been shown to display two important functions.

First, NAR speeds up the response time of gene circuits. This occurs when NAR uses a strong promoter to obtain a rapid initial rise in the concentration of protein X. When X concentration reaches the repression threshold for its own promoter, the production rate of new X decreases. Thus, the concentration of X locks into a steady-state level that is close to its repression threshold. This steady-state level can be adjusted over evolutionary time by mutations that cause variation in the repression threshold of X to its own promoter. By contrast, a simply regulated gene that is designed to reach the same steady-state level must use a weaker promoter. As a result, an NAR system reaches 50% of its steady state faster than a simply regulated gene (FIG. 1d). The dynamics of NAR show a rapid initial rise followed by a sudden locking into the steady state, possibly accompanied by an overshoot or damped oscillations. Response acceleration (or speed-up) by NAR has been demonstrated experimentally[17] using a fluorescently tagged repressor, TetR, that was designed to repress its own promoter (FIG. 1e). Speed-up in a natural context was demonstrated in the SOS DNA-repair system of *E. coli*, in which the master regulator, LexA, represses its own promoter[20]. These and many of the other experiments discussed in this Review were made possible by fluorescent-reporter assays, which allow the transcription dynamics of living cells to be measured with high resolution and accuracy[21].

In addition to speeding responses, NAR can reduce cell–cell variation in protein levels. These variations are due to an inherent source of noise: the production rates of proteins fluctuate by tens of percents (reviewed in REF. 22) (FIG. 1f). This noise results in cell–cell variation in protein level. NAR can, in many cases, reduce these variations: high concentrations of X reduce its own rate of production, whereas low concentrations cause an increased production rate. The result is a narrower distribution of protein levels than would be expected in simply regulated genes (FIG. 1f), as demonstrated experimentally by Besckei and Serrano[19,20,23]. However, if the NAR feedback contains a long delay, noise can also be amplified.

Positive autoregulation. Positive autoregulation (PAR) occurs when a transcription factor enhances its own rate of production (FIG. 1c). The effects are opposite to those of NAR: response times are slowed and variation is usually enhanced.

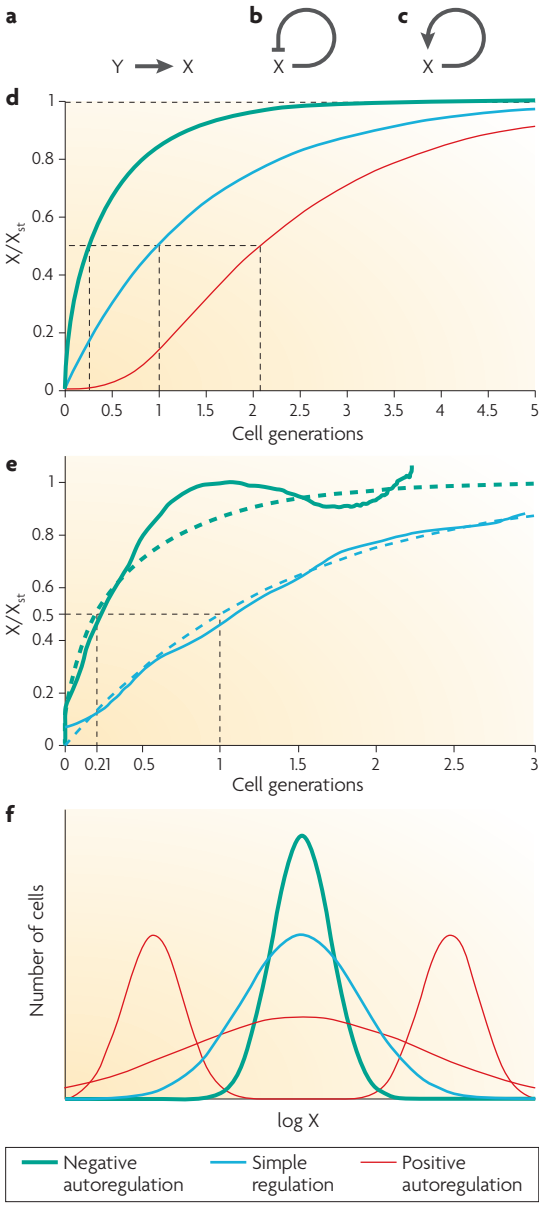


Figure 1 | **Simple regulation and autoregulation.** **a** | In simple regulation, transcription factor Y is activated by a signal S_y. When active, it binds the promoter of gene X to enhance or inhibit its transcription rate. **b** | In negative autoregulation (NAR), X is a transcription factor that represses its own promoter. **c** | In positive autoregulation (PAR), X activates its own promoter. **d** | NAR speeds the response time (the time needed to reach halfway to the steady-state concentration) relative to a simple-regulation system that reaches the same steady-state expression. PAR slows the response time. **e** | An experimental study of NAR, using a synthetic gene circuit in which the repressor TetR fused to GFP represses its own promoter. High-resolution fluorescence measurements in living *Escherichia coli* cells show that this NAR motif has a response time about fivefold faster than a simple-regulation design. **f** | A schematic cell–cell distribution of protein levels. NAR tends to make this distribution narrower in comparison with simple regulation, whereas PAR tends to make it wider, and in extreme cases bimodal with two populations of cells. X/X_{st}, X concentration relative to steady state X_{st}.

PAR slows the response time because at early stages, when levels of X are low, production is slow. Production picks up only when X concentration approaches the activation threshold for its own promoter. Thus, the desired steady state is reached in an S-shaped curve (FIG. 1d). The response time is longer than in a corresponding simple-regulation system, as shown theoretically[24] and experimentally by Maeda and Sano[25].

PAR tends to increase cell–cell variability. If PAR is weak (that is, X moderately enhances its own production rate), the cell–cell distribution of X concentration is expected to be broader than in the case of a simply regulated gene (FIG. 1f). Strong PAR can lead to bimodal distributions, whereby the concentration of X is low in some cells but high in others. In cells in which the concentration is high, X activates its own production and keeps it high indefinitely. Strong PAR can therefore lead to a differentiation-like partitioning of cells into two populations[25–27] (FIG. 1f). In some cases, PAR can be useful as a memory to maintain gene expression, as mentioned below (see the section on developmental networks). In other cases, a bimodal distribution is thought to help cell populations to maintain a mixed phenotype so that they can better respond to a stochastic environment (reviewed in REF. 28).

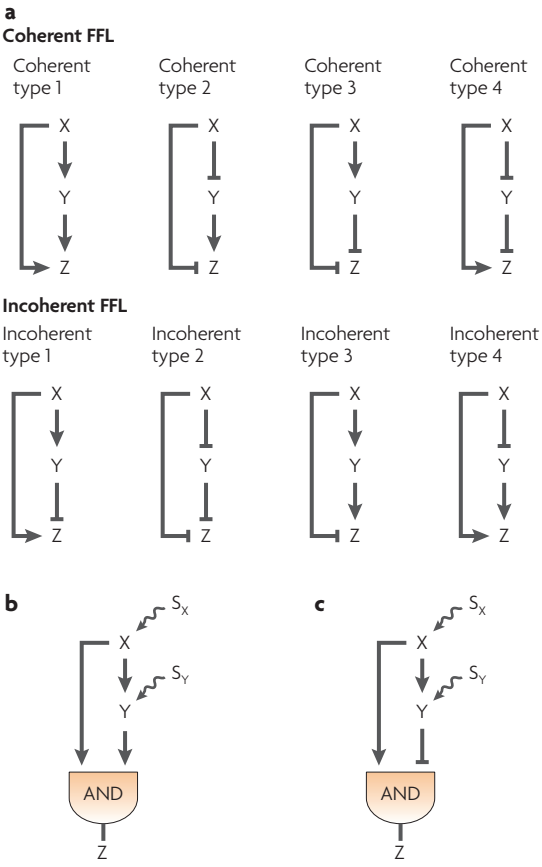


Figure 2 | **Feedforward loops (FFLs). a** | The eight types of feedforward loops (FFLs) are shown. In coherent FFLs, the sign of the direct path from transcription factor X to output Z is the same as the overall sign of the indirect path through transcription factor Y. Incoherent FFLs have opposite signs for the two paths. **b** | The coherent type-1 FFL with an AND input function at the Z promoter. **c** | The incoherent type-1 FFL with an AND input function at the Z promoter. S_X and S_Y are input signals for X and Y.

Feedforward loops

The second family of network motifs is the feedforward loop (FFL). It appears in hundreds of gene systems in *E. coli*[6,9] and yeast[7,10], as well as in other organisms[11–16]. This motif consists of three genes: a regulator, X, which regulates Y, and gene Z, which is regulated by both X and Y. Because each of the three regulatory interactions in the FFL can be either activation or repression, there are eight possible structural types of FFL (FIG. 2a).

To understand the function of the FFLs, we need to understand how X and Y are integrated to regulate the Z promoter[29,30]. Two common 'input functions' are an 'AND gate', in which both X and Y are needed to activate Z, and an 'OR gate', in which binding of either regulator is sufficient. Other input functions are possible, such as the additive input function in the flagella system[24,31] and the hybrid of AND and OR logic in the *lac* promoter[32]. However, much of the essential behaviour of FFLs can be understood by focusing on the stereotypical AND and OR gates. Each of the eight FFL types can thus appear with at least two input functions.

In the best studied transcriptional networks (*E. coli* and yeast), two of the eight FFL types occur much more frequently than the other six types. These common types are the coherent type-1 FFL (C1-FFL) and the incoherent type-1 FFL (I1-FFL)[33,34,36]. Here I discuss their dynamical functions in detail; the functions of all eight FFL types are described in REF. 34.

The C1-FFL is a 'sign-sensitive delay' element and a persistence detector. In the C1-FFL, both X and Y are transcriptional activators (FIG. 2b). I will first consider the behaviour of the FFL when the Z promoter has an AND input function, and then turn to the case of the OR input function.

With an AND input function, the C1-FFL shows a delay after stimulation, but no delay when stimulation stops. To see this, let's follow the behaviour of the FFL. When the signal S_x appears, X becomes active and rapidly binds its downstream promoters. As a result, Y begins to accumulate. However, owing to the AND input function, Z production starts only when Y concentration crosses the activation threshold for the Z promoter. This results in a delay of Z expression following the appearance of S_x (FIG. 3a). In contrast, when the signal S_x is removed, X rapidly becomes inactive. As a result, Z production stops because deactivation of its promoter requires only one arm of the AND gate to be 'shut off'. Hence, there is no delay in deactivation of Z after the signal S_x is removed (FIG. 3a).

This dynamic behaviour is called sign-sensitive delay; that is, delay depends on the sign of the S_x step. An ON step (addition of S_x) causes a delay in Z expression, but an OFF step (removal of S_x) causes no delay.

The duration of the delay is determined by the biochemical parameters of the regulator Y; for example, the

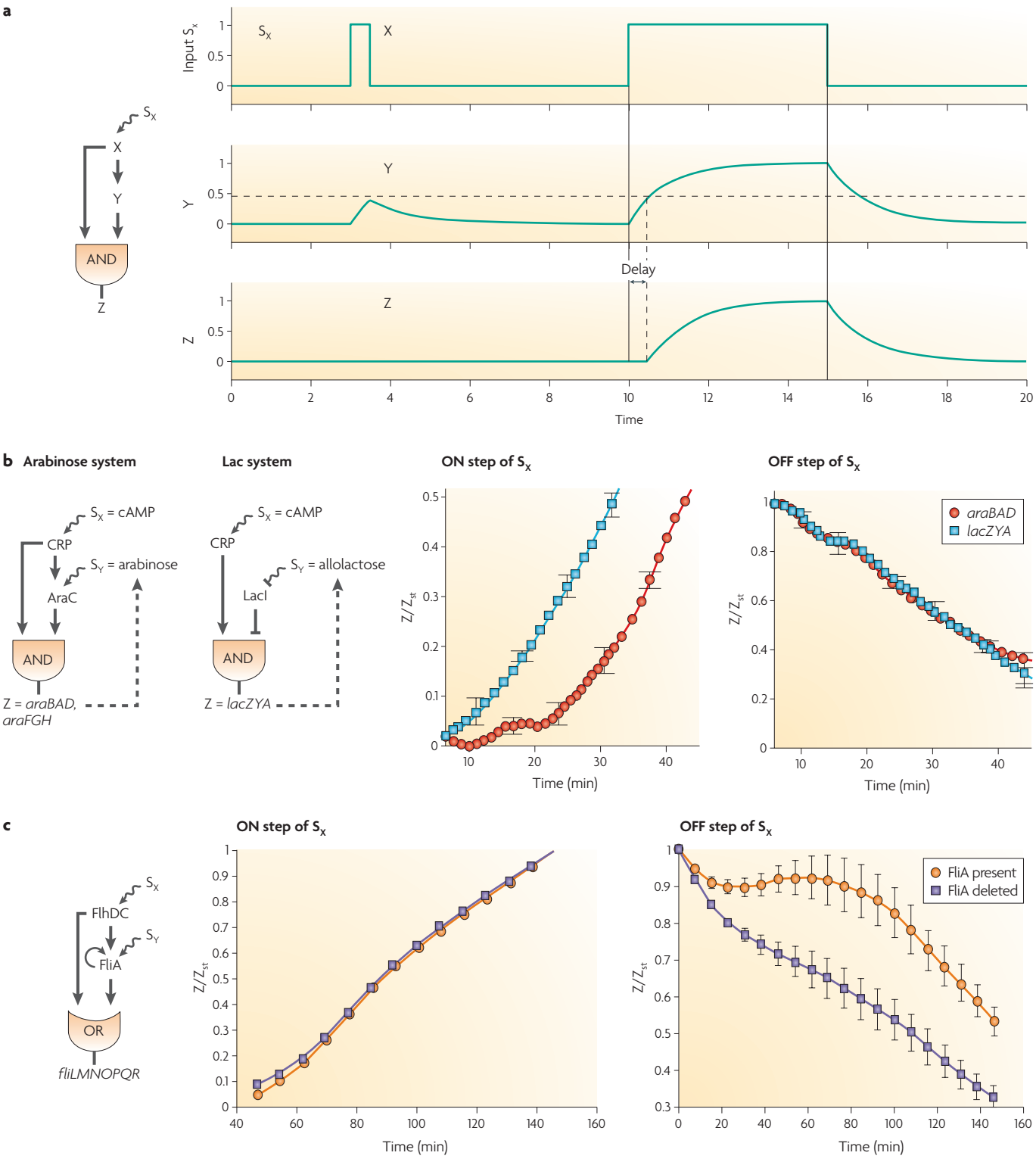


Figure 3 | **The coherent type-1 feedforward loop (C1-FFL) and its dynamics. a** | The C1-FFL with an AND input function shows delay after stimulus (S_X) addition, and no delay after stimulus removal. It thus acts as a sign-sensitive filter, which responds only to persistent stimuli. **b** | An experimental study of the C1-FFL in the arabinose system of *Escherichia coli*, using fluorescent-reporter strains and high-resolution measurements in living cells. This system (represented by red circles) shows a delay after addition of the input signal (cAMP), and no delay after its removal, relative to a simple-regulation system that responds to the same input signal (the *lac* system, represented by blue squares). **c** | The C1-FFL with an OR-like input function in the flagella system of *E. coli* shows a delay after signal removal but not after the onset of signal (represented by orange circles). Deletion of the 'Y' gene (FliA) abolishes this delay (represented by purple squares). Z/Z_{st}, Z concentration relative to the steady state Z_{st}.

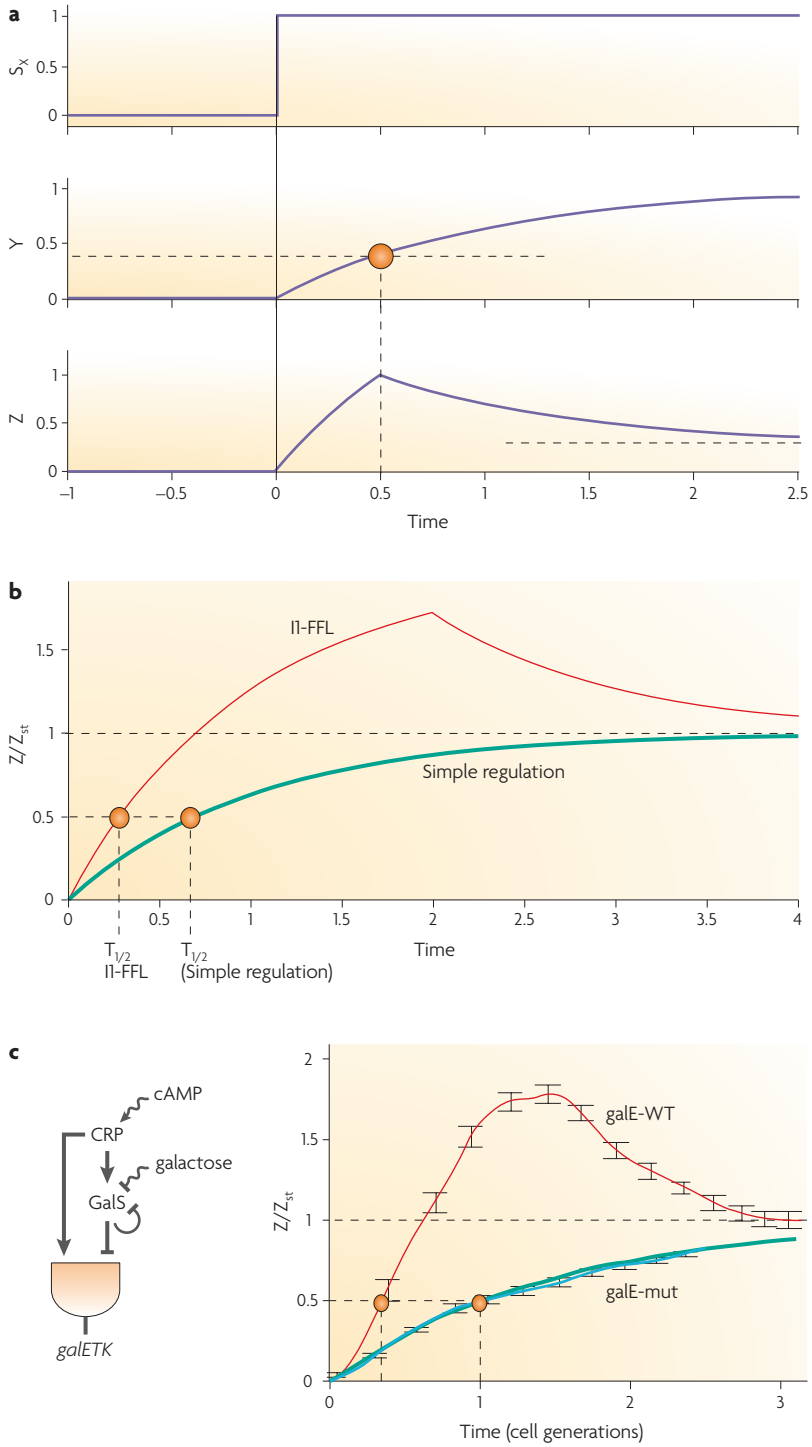


Figure 4 | **The incoherent type-1 feeforward loop (I1-FFL) and its dynamics.** **a** | The I1-FFL can generate a pulse of Z expression in response to a step stimulus of S_x. This occurs because once Y has passed its threshold (indicated by an orange circle) it starts to repress Z. **b** | The I1-FFL shows faster response time for the concentration of protein Z than a simple-regulation circuit with the same steady-state expression level. **c** | An experimental study of the dynamics of the I1-FFL in the galactose system of *E. coli*. Response acceleration in the wild-type system (marked 'galE-WT') is found following steps of the input signal (glucose starvation). The acceleration is disrupted when the effect of the repressor GalS is abolished by mutating its binding site in the promoter of the output gene operon *galETK* (marked 'galE-mut'). $T_{1/2}$, response time; Z/Z_{st}, Z concentration relative to the steady state.

higher the activation threshold for the Z promoter by Y, the longer the delay. The delay that is generated by the FFL can be useful to filter out brief spurious pulses of signal. A signal that appears only briefly does not allow Y to accumulate and cross its threshold, and thus does not induce a Z response. Only persistent signals lead to Z expression (FIG. 3a).

The sign-sensitive delay function of this motif has been experimentally demonstrated in the arabinose-utilization system of *E. coli*[9] (FIG. 3b). A delay occurs after addition of the input signal cAMP, but not after its removal. This delay, of about 20 min, is on the same timescale as spurious pulses of cAMP that occur in the natural environment when *E. coli* transits between growth conditions.

When the Z promoter has OR logic, the FFL has the opposite effect to the AND case we have just discussed: with an OR input function, the C1-FFL shows no delay after stimulation, but does show a delay when stimulation stops. To see this, note that when the signal S_x appears, X alone is sufficient to activate Z because of the OR-gate logic. If the signal suddenly stops after a long period of stimulation, X is no longer active, but the presence of Y is still enough to allow production of Z. Thus, the C1-FFL with OR logic allows continued production in the face of a transient loss of the input signal.

This behaviour was experimentally demonstrated in the flagella system of *E. coli*[24] (FIG. 3c). The flagella motor genes are regulated in an FFL that has input functions that resemble OR gates (additive functions of the two activators FlhDC and FliA). The flagella FFL was found to prolong flagella gene expression after the input signal (active FlhDC) stopped, but no delay occurred when the input signal appeared. Mutations and conditions that inactivate the FliA gene in this FFL lead to a loss of this delay, resulting in immediate shut-off of the flagella genes once the input signal stops. The delay in the flagella system, of about 1 hour, is comparable to the time that is needed for the biogenesis of a complete flagella motor.

The I1-FFL is a pulse generator and response accelerator. In the I1-FFL, the two arms of the FFL act in opposition: X activates Z, but also represses Z by activating the repressor Y (FIG. 2c). As a result, when a signal causes X to assume its active conformation, Z is rapidly produced (FIG. 4a). However, after some time, Y levels accumulate to reach the repression threshold for the Z promoter. As a result, Z production decreases and its concentration drops, resulting in pulse-like dynamics (FIG. 4b). In the extreme case that Y completely represses Z, the pulse drops to zero.

Pulse-like dynamics were experimentally demonstrated in a synthetic I1-FFL that was built of well-characterized bacterial regulators in *E. coli*[35]. In this FFL, the activator LuxR (X) was made to activate both a GFP reporter (Z) and the λ-repressor C1 (Y), which repressed the Z promoter.

In addition to pulse-like dynamics, the I1-FFL can carry out another dynamical function: response acceleration. In cases in which Y does not completely repress

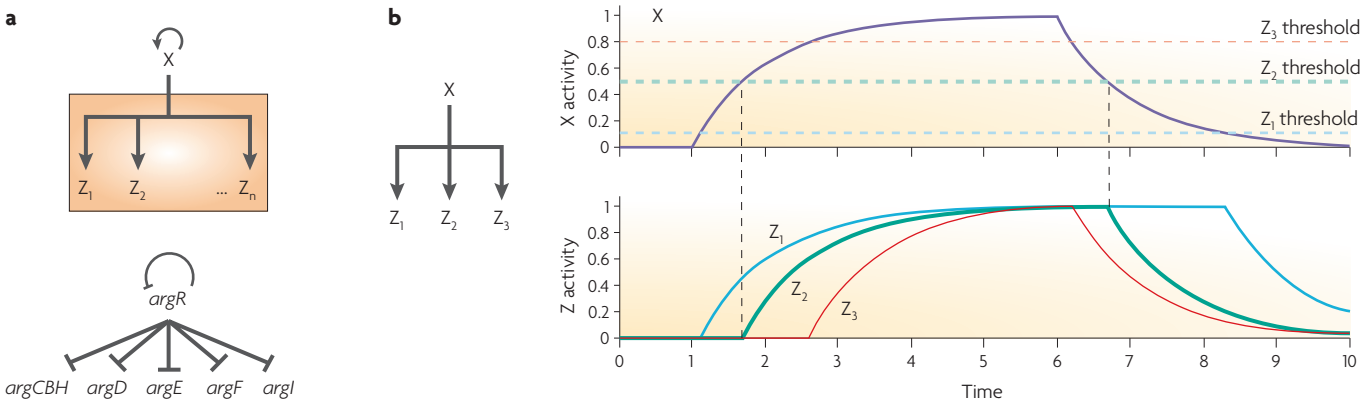


Figure 5 | **The single-input module (SIM) network motif and its dynamics. a** | The single-input module (SIM) network motif, and an example from the arginine-biosynthesis system. **b** | Temporal order of expression in a SIM. As the activity of the master regulator X changes in time, it crosses the different activation threshold of the genes in the SIM at different times, generating a temporal order of expression.

the production of Z, Z concentration reaches a certain non-zero steady-state level. Because of the strong initial production of Z in the time period before Y represses the Z promoter, Z reaches its steady-state rapidly. The response time is shorter than that of a corresponding simple-regulation system (FIG. 4c). Note that, although both NAR and I1-FFLs can speed up responses, NAR works only on transcription factors (or genes that lie on the same operon with transcription factors), whereas the I1-FFL can accelerate any target gene Z.

Such response acceleration was observed experimentally in the galactose utilization system of *E. coli*[36] (FIG. 4c). Here onset of glucose starvation in the absence of galactose leads to a rapid induction of the galactose-utilization genes to a moderate level of expression. The response time of this system is about threefold faster than that of a simple-regulation system that responds to the same signal (the *lac* system). This speed-up was dependent on the I1-FFL: in mutants and conditions in which the motif was disrupted, speed-up was abolished and the dynamics resembled simple regulation (FIG. 4c).

Note that network motifs can utilize not only transcription factor proteins but also microRNAs (miRNAs)[37]. For example, an I1-FFL in mammalian cells involves MYC as activator X, *E2F1* as the target gene Z, and a miRNA in the role of the repressor Y[38]. Diverse FFL motifs with miRNAs have been found in *Caenorhabditis elegans*[39].

The NAR and PAR network motifs are sometimes integrated into FFLs, usually on the regulator Y. These regulatory loops can help to speed up or slow down the response time of Y, enhancing the behaviour of the FFLs.

The dynamical functions of FFLs can be tuned by varying the molecular parameters of the circuit. Changes in parameters such as the production rates or the activation thresholds of the regulators can, as mentioned above, determine the delay in the C1-FFL, or the acceleration factor of the I1-FFL. This tuning can be captured by simple models[1,9,34,36]. Similar functions can, in principle, be accomplished by other circuits that resemble FFLs, but with longer branches that diverge and then merge back. However, such larger circuits are rarely found in known transcription networks. The FFL can potentially perform additional computational functions, as suggested by theoretical analyses[40–43].

Multi-output FFLs. The FFLs in transcription networks tend to combine to form multi-output FFLs[6,44,45], in which X and Y regulate multiple output genes $Z_1, Z_2, \ldots Z_n$. In these configurations, each of the output genes benefits from the dynamical functions that are described above. In addition, the multi-output FFL can generate temporal orders of gene activation and inactivation by means of a hierarchy of regulation thresholds for the different promoters. This was experimentally demonstrated using the flagella genes[31]: mutations in the promoter regions that changed the activation thresholds were able to reprogramme the temporal order of the genes[31]. Further experimental tests of the dynamical behaviour of FFLs in living cells would be of great interest, especially in organisms other than *E. coli*.

Single-input modules (SIM)

Our third family of network motifs have a simple pattern in which a regulator X regulates a group of target genes (FIG. 5a). In the purest form, no other regulator regulates any of these genes, hence the name single-input module. X also typically regulates itself.

The main function of this motif is to allow coordinated expression of a group of genes with shared function. In addition, this motif has a more subtle dynamical property that is similar to that of the multi-output FFLs that are discussed above: it can generate a temporal expression programme, with a defined order of activation of each of the target promoters. X often has different activation thresholds for each gene, owing to variations in the sequence and context of its binding site in each promoter. So, when X activity rises gradually with time, it crosses these thresholds in a defined order, first the lowest threshold, then the next lowest threshold, an so on, resulting in a temporal order of expression (FIG. 5b). Similar reasoning applies when X acts as a repressor.

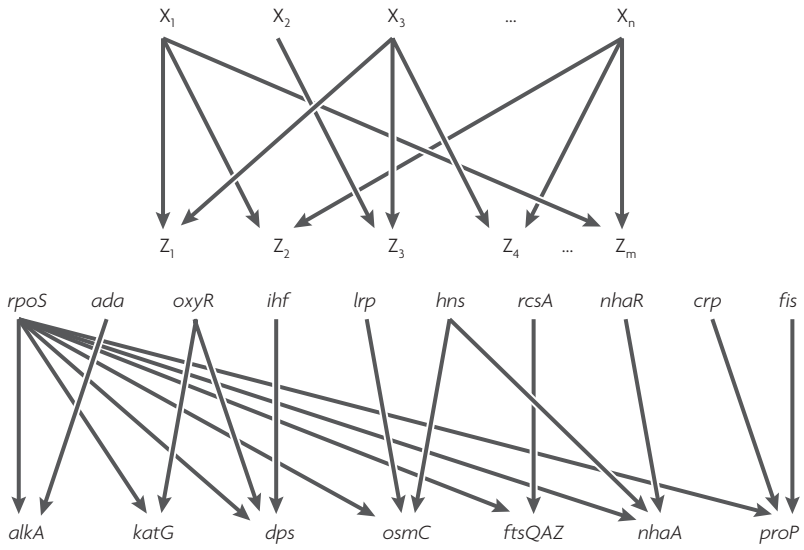


Figure 6 | **The dense overlapping regulon (DOR) network motif.** In this motif, many inputs regulate many outputs (top panel). The bottom panel shows an example from the stress-response system of *Escherichia coli*.

Such a temporal order has been observed experimentally in several *E. coli* systems with SIM architecture that have been studied at high temporal resolution[46,47]. Importantly, the temporal order seems to match the functional order of the genes. The earlier a gene is needed in a multi-gene process, the earlier its promoter is activated. This kind of programme can prevent protein production before it is needed. For example, the arginine-biosynthesis system shows a SIM design in which the repressor ArgR regulates several operons that encode enzymes in the arginine-biosynthesis pathway. When arginine is removed from the medium, these promoters are activated in a temporal order with minutes between promoter activations[47]. The order of activation matches the position of the enzymes in the arginine-biosynthesis pathway. The same principle applies to other linear biosynthesis pathways[47] and stress-response systems such as the SOS DNA repair system[46]. Many other examples of temporal order are known, including the flagella systems of *E. coli*[48] and *Caulobacter crecentus*[49], cell-cycle gene systems in many organisms[50,51] and developmental programmes[4].

Dense overlapping regulons (DOR)

The final family of network motifs that are present in sensory transcription networks consist of a set of regulators that combinatorially control a set of output genes[6] (FIG. 6a). These motifs are referred to as dense overlapping regulons (DORs) or multi-input motifs (MIMs). *E. coli* has several DORs with hundreds of output genes, each responsible for a broad biological function, such as carbon utilization, anaerobic growth, stress response, and so on. Similar patterns are found in yeast[10]. The DOR can be thought of as a gate-array, carrying out a computation by which multiple inputs are translated into multiple outputs. So, to fully understand the function of the DOR, the connectivity arrows are not enough[6,52]: the input functions in the promoter of each output gene must also be specified. Currently, most of the input functions in any organism are unknown[2,53,54]. Once these functions are characterized, for example, by high-resolution mapping using fluorescent-reporter strains[31,32], it will be interesting to study the detailed function of DORs.

The global organization of network motifs

The four motif families that have been discussed seem to cover most of the known interactions in the transcription networks of *E. coli* and yeast. As such, they appear to be the main building blocks of these sensory networks. How do these network motifs combine to form the global structure of the networks?

To answer this question, an image of the network is required. Network motifs can help to portray the network in a compact way, by using symbols to denote SIMs, DORs and FFLs (see REF. 6 for an example). This kind of arrangement shows that FFLs and SIMs are integrated into the DORs. The DORs occur in a single layer: there is no DOR at the output of a second DOR. Thus, most computations are carried out in a single 'cortex' of promoters at the DOR output. Furthermore, long regulatory cascades are rare[55]; most genes are regulated just one step away from their activator (with relatively few exceptions). One possible reason for this shallow architecture is the need for rapid response: as mentioned above, it can take up to one cell-generation time to pass a signal down each step of a cascade. Sensory networks that are designed to respond rapidly to external signals might therefore be limited in their use of long cascades[55].

A view of network-motif behaviour within the global dynamics of gene networks can be gleaned by means of DNA microarrays[56–58]. For example, transcription dynamics of mammalian genes in response to growth-factor stimulation could be related to network motifs[59]. Future work in this direction could help us to refine our understanding of motif dynamics and study the interactions between network motifs.

Network motifs in developmental networks

So far, I have discussed sensory transcription networks that respond rapidly and make reversible decisions. Let us now turn to developmental transcription networks that transduce signals into cell-fate decisions[2,4,14,60,61]. These networks have different constraints: they usually function on the timescale of one or several cell generations, and often need to make irreversible decisions that last even after the input signal has vanished.

Developmental transcription networks use all the network motifs described above. In addition, as a result of their specific requirements, developmental networks use several other network motifs that are not commonly found in sensory networks.

Feedback loops comprising two transcription interactions. Developmental transcription networks often use positive-feedback loops that are made up of two transcription factors that regulate each other. There are two kinds of positive-feedback loops, a double-positive loop and a double-negative loop (FIG. 7a). The double-positive loop, in which two activators activate

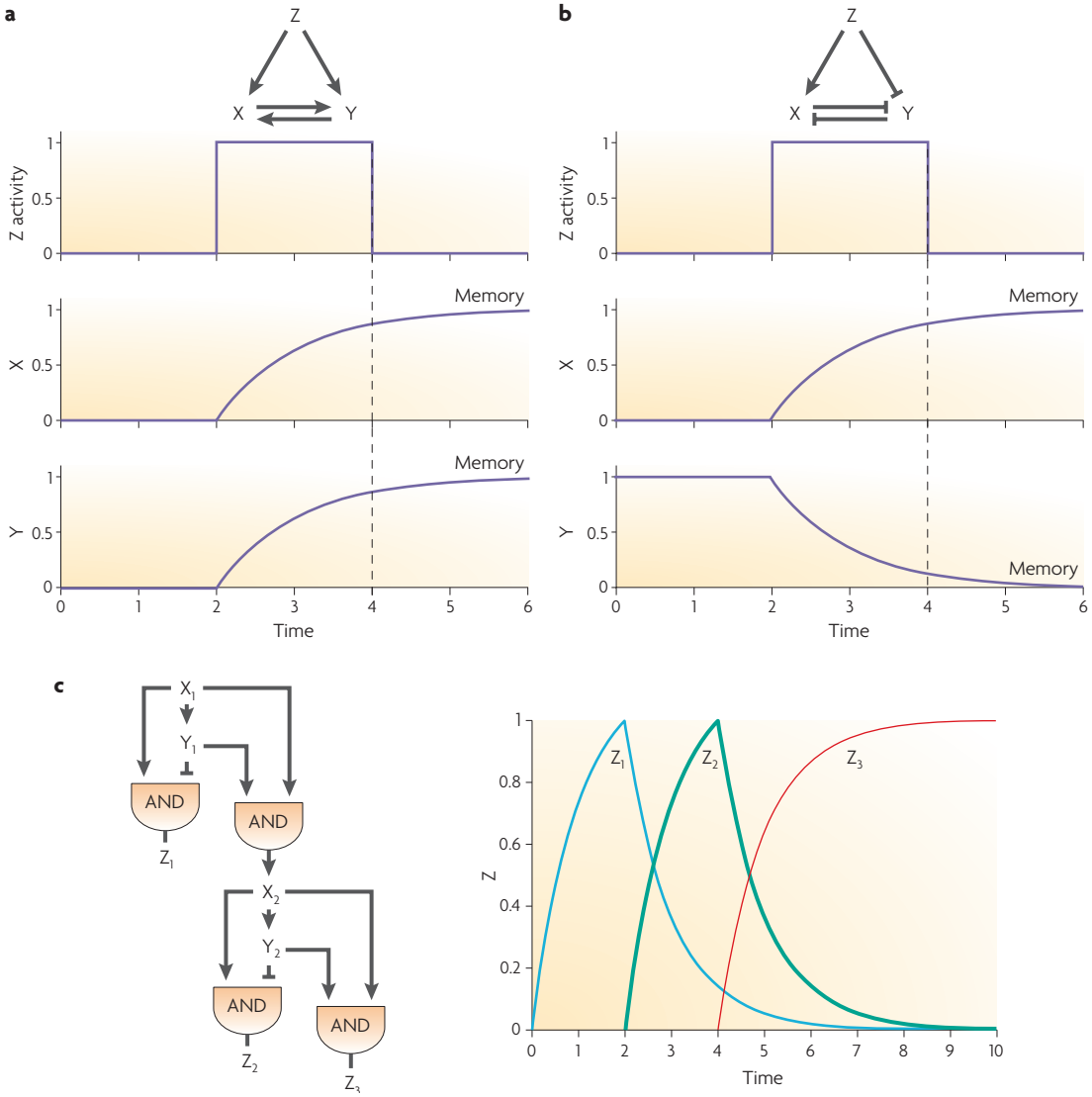
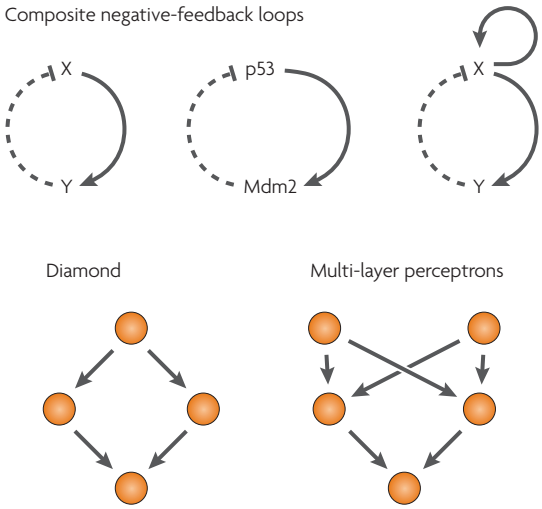


Figure 7 | **Network motifs in developmental transcription networks. a** | Network motifs with a double-positive-feedback loop. When Z is activated, proteins X and Y begin to be produced. They can remain locked ON even when Z is deactivated (at times after the dashed line). **b** | Regulated feedback with a double-negative-feedback loop. Here Z acts to switch the steady states. Initially, Y concentration is high and represses X expression. After Z is activated, X is produced and Y is repressed. This state can persist even after Z is deactivated. Thus, the feedback implements a memory. **c** | A transcription network that guides the development of the *Bacillus subtilis* spore[8]. Z_1, Z_2 and Z_3 represent groups of tens to hundreds of genes. This network is made of two incoherent type-1 feedforward loops (I1-FFLs), which generate pulses of Z_1 and Z_2, and two coherent type-1 feedforward loops (FFLs), one of which generates a delayed Z_3 step.

each other, has two steady states: either both X and Y are OFF, or both are ON. The double-negative loop, in which two repressors repress each other, has different steady states: either X is ON and Y is OFF, or the opposite. In both cases, a transient signal can cause the loop to lock irreversibly into a steady state. In this sense, this network motif can provide memory of an input signal, even after the input signal is gone. Often, X and Y also positively regulate themselves, strengthening the memory effects. The same motif can also comprise miRNAs[62] or post-transcriptional interactions such as phosphorylations[63,64].

Positive-feedback loops can regulate or be regulated by other signals[2,16,60]. In a regulating loop, two regulators X and Y form a feedback loop, and also jointly regulate downstream Z genes. A double-positive loop between X and Y is useful for decisions whereby the cell irreversibly assumes a fate in response to a transient developmental signal. Genes that are specific to the cell fate can be co-activated by X and Y. A double-negative loop (FIG. 7b) is useful in this motif as a toggle switch between two different fates[65], such as lysogeny and lysis in λ-phage[3]. The genes that are activated by X are repressed by Y, and the opposite.

Box 1 | **Network motifs in other biological networks**



In addition to transcription networks, one can seek composite network motifs that include different types of interactions[76,77]. One of the most common composite motifs is a negative-feedback loop between two proteins, in which one arm is a transcriptional interaction (solid arrow) and the other arm is a protein–protein interaction (broken arrow). An example is the p53 and Mdm2 loop involved in monitoring stresses and DNA damage in human cells. Composite negative-feedback loops seem to be much more common than purely transcriptional negative-feedback loops; The separation of timescales between the slow transcription arm and the faster protein-interaction arm might help to stabilize the dynamics of composite loops, avoiding feedback at a delay that promotes instability (as observed in a synthetic three-repressor loop[78]). Experiments on individual living cells have shown that negative-feedback loops, embedded within additional interactions, can sometimes generate oscillations, whereby the levels of X and Y rise and fall[79–82]. Oscillations in biological systems are often generated by a composite negative-feedback loop coupled to a second, positive-feedback loop[83–86] (right hand side of the top panel). The same motif with different parameters can also lead to stochastic, excitable systems that occasionally generate a single large output pulse[87].

Networks of protein modification, notably signal-transduction networks, also seem to display network motifs[88–90]. Here nodes (orange circles) are signalling proteins and edges (arrows) represent modifications such as phosphorylation. Signal-transduction networks show feedforward loops (FFLs), as well as motifs that are not present in transcription networks, such as the diamond pattern. Diamonds combine to form multi-layer perceptron motifs that are composed of three or more layers of signalling proteins[89]. Such patterns can potentially carry out elaborate functions on multiple input signals, including generalization of information from partial signals[91,92]. They also can show graceful degradation of performance upon loss of components[91,92]. Current high-quality data on protein–protein interactions is more limited than data on transcription interactions. It is likely that additional motifs will be discovered once data on protein–interaction networks becomes more complete. Differences in timescale, spatial organization and precision between signalling processes and transcription processes are likely to underlie the differences in the network motifs that are found in these networks.

Networks of synaptic connection between neurons also seem to exhibit network motifs. Cortical circuitry harbours triplets of neurons that are connected as FFLs[93]. In particular, many neuron types are wired as incoherent type-1 feedforward loops (I1-FFLs) with afferent input as X, an inhibitory neuron from which outputs are restricted to a specific brain region as Y, and a relay neuron that sends connection to other regions as Z[94]. The fully mapped synaptic network of *Caenorhabditis elegans*[95] shows motifs including FFLs[7], diamonds and multi-layered perceptrons[89,96], as well as two-neuron feedback loops[97]. Mammalian neuronal networks also display significant network motifs[98], as assayed using electrical measurements of neuron tetrads[99]. It is an interesting question whether neuronal motifs carry out similar computational functions to the motifs that are discussed in the main text.

In a regulated loop, two regulators X and Y form a feedback loop and are both regulated by an upstream regulator Z (FIG. 7a,b). In this motif, an activator Z can be used to lock a double-positive loop into an ON state[2,60]. In the case of a double-negative loop, Z can activate X and repress Y (or the opposite), and thus act to switch the system between its two steady states. Many positive-feedback loops are both regulated and regulating. Variants of this motif include cases in which Z inputs to only one of the two regulators, or each regulator has its own independent input[65].

Transcription cascade. In addition to motifs that use feedback loops, developmental transcription networks tend to have much longer cascades than sensory transcription networks[16,55,66–67]. These cascades pass information on a slow timescale, which can be on the order of one cell generation at each cascade step (or, for degradable regulators, the half-life of the regulator at each step), an appropriate pace for many developmental processes. Development often uses repressor cascades, the timing properties of which can often be more robust to noise in protein-production rates than those of activator cascades[68]. Cascades are also commonly found in signalling networks, which function on a faster timescale than transcription networks (BOX 1).

An interlocked FFL circuit in development

The FFLs in developmental networks often combine into larger and more complex transcription circuits than in sensory networks. Can we still understand the dynamics of such large circuits on the basis of the behaviour of the individual FFLs?

To address this question, I will discuss a well mapped developmental network that is composed of interlocking FFLs. This circuit governs differentiation in the bacterium *Bacillus subtilis*. When starved, *B. subtilis* cells differentiate into durable spores. To produce a spore, *B. subtilis* must

make many proteins that are not found in the growing bacterium. This process, termed sporulation, involves hundreds of genes. These genes are turned ON and OFF in a series of temporal waves, each carrying out specific stages in the formation of the spore. The network that regulates this process[8] includes several transcription factors that are arranged in linked C1-FFLs and I1-FFLs (FIG. 7c).

To initiate the sporulation process, a starvation signal S_x activates X_1 (an activator called *Spo0A*). X_1 acts in an I1-FFL together with Y_1 to control the genes Z_1. This I1-FFL generates a pulse of Z_1 expression. A C1-FFL with AND logic is formed by X_1 and Y_1; that is, both are required to activate X_2. This C1-FFL ensures that X_2 is not activated unless the S_x signal is persistent. Next, X_2 acts in an I1-FFL, by which it generates a pulse of Z_2 genes, timed at a delay relative to the first pulse. Finally, Y_2 and X_2 together join in an AND-gate C1-FFL to activate Z_3 genes, which are turned on last. The result is a three-wave temporal pattern: first a pulse of Z_1 expression, followed by a pulse of Z_2 expression, followed by expression of the 'late' Z_3 genes (FIG. 7c).

The FFLs in this network are combined in a way that utilizes their delay and pulse-generating features to generate a temporal programme of gene expression. The FFLs that control Z_1, Z_2 and Z_3 are actually multi-output FFLs because Z_1,Z_2 and Z_3 each represent groups of genes. This design can generate finer temporal programmes within each group of genes.

The FFLs in this network therefore seem to be linked in a way that allows easy interpretation on the basis of the dynamics of each FFL in isolation. It is interesting to consider whether such modular design applies to network motifs in other systems, a question that can be addressed experimentally.

Convergent evolution of network motifs

How did network motifs evolve? The most common form of evolution for genes is conservative evolution, whereby two genes that have similar functions stem from a common-ancestor gene. This is reflected in a significant degree of sequence similarity between the genes, called gene homology.

Did network motifs such as FFLs evolve in a similar way, in that an ancestral FFL duplicated and gave rise to the present FFLs? In most cases, it seems that the answer is no. For example, homologous genes Z and Z′ in two organisms are often both regulated by FFLs in response to similar environmental stimuli. If the two FFLs had a common-ancestor FFL, the regulators X and Y in these systems would also be homologous. However, this is generally not the case. The sequence of the regulators is sometimes so different that they are classed into completely different transcription factor families. The same applies to SIM and DOR network motifs: similar output genes in different organisms are often regulated by unrelated transcription factors[1]. It therefore seems that, in many cases, evolution has converged independently on the same regulation circuit[56,69].

To understand convergent network-motif evolution, it is important to note that transcription networks seem to rewire rapidly on evolutionary timescales[56,69,70]: it takes only a few mutations to remove the binding site of a regulator in a given promoter, and thereby lose an arrow in a network[70,71]. Hence, even closely related organisms often have different network motifs to regulate a given gene, provided that they live in different environments, as was demonstrated by Babu *et al.*[72]. One hypothesis is that the network motifs are selected according to the computations that are required in the environment of each species. For example, the selective advantage of FFLs in different environments was treated theoretically in REF. 73.

Network motifs might have been 'rediscovered' by evolution because they perform important functions. They seem to be the most robust[74] and use the least number of components of the large set of circuits that can carry out equivalent functions. Intriguingly, network motifs are also found in various other biological networks (BOX 1).

Detection of network motifs

Open-source software that can detect network motifs is available[75] (see the Uri Alon laboratory web site). This software accepts a network as input, and detects network motifs as patterns that occur more often in the network than in random networks with the same size and connectivity properties. The software accepts network data in the form of a list that details the interactions that occur between different nodes (and, optionally, the type of interaction). The software outputs the recurring network motifs and depicts these motifs within the network.

Future directions

Experiments in living cells and mathematical modelling have helped to define some of the functions of network motifs. Much remains to be studied: predictions about the functions of motifs must be tested experimentally in the different systems in which the motifs appear. As most experiments so far have used bacteria, it would be important to test network motifs in eukaryotic organisms. In the examples studied so far, there was a good agreement between the theoretical predictions and experimental tests. The specific ways in which the network motifs in these examples are wired together have allowed us to understand the dynamics of each individual motif, even when it is connected to the rest of the networks of the cell. As more systems are investigated, it is likely that more complicated cases will be found, in which the behaviour of motifs is affected to a much greater extent by its context within the rest of the network: this is an open field for research.

As networks become better characterized, new motifs and new motif functions will doubtless be discovered. Network motifs at the level of signalling networks and neuronal networks are only beginning to be investigated.

If the current findings can be generalized, they suggest that complex biological networks have a degree of structural simplicity, in that they contain a limited set of network motifs. Experimental testing of the functions of each motif might explain why these motifs have been selected again and again in evolution. This raises the hope that the dynamics of large networks can be understood in terms of elementary circuit patterns[100].

1. Alon, U. *Introduction to Systems Biology: Design Principles Of Biological Circuits* (CRC, Boca Raton, 2006).
2. Davidson, E. H. *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*, (Academic, Burlington, 2006).
3. Ptashne, M. & Gann, A. *Genes & Signals* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2002).
4. Levine, M. & Davidson, E. H. Gene regulatory networks for development. *Proc. Natl Acad. Sci. USA* **102**, 4936–4942 (2005).
5. Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* **20**, 433–440 (1998).
6. Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.* **31**, 64–68 (2002).
7. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
8. Eichenberger, P. *et al.* The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol.* **2**, e328 (2004).
9. Mangan, S., Zaslaver, A. & Alon, U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J. Mol. Biol.* **334**, 197–204 (2003).
10. Lee, T. I. *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
11. Odom, D. T. *et al.* Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381 (2004).
12. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
13. Saddic, L. A. *et al.* The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of CAULIFLOWER. *Development* **133**, 1673–1682 (2006).
14. Swiers, G., Patient, R. & Loose, M. Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification. *Dev. Biol.* **294**, 525–540 (2006).
15. Iranfar, N., Fuller, D. & Loomis, W. F. Transcriptional regulation of post-aggregation genes in *Dictyostelium* by a feed-forward loop involving GBF and LagC. *Dev. Biol.* **290**, 460–469 (2006).
16. Milo, R. *et al.* Superfamilies of designed and evolved networks. *Science* **303**, 1538–1542 (2004).
17. Rosenfeld, N., Elowitz, M. B. & Alon, U. Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.* **323**, 785–793 (2002).
18. Savageau, M. A. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* **252**, 546–549 (1974).
19. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
20. Camas, F. M., Blazquez, J. & Poyatos, J. F. Autogenous and nonautogenous control of response in a genetic network. *Proc. Natl Acad. Sci. USA* **103**, 12718–12723 (2006).
21. Zaslaver, A. *et al.* A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nature Methods* **3**, 623–628 (2006).
22. Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nature Rev. Genet.* **6**, 451–464 (2005).
23. Dublanche, Y., Michalodimitrakis, K., Kummerer, N., Foglierini, M. & Serrano, L. Noise in transcription negative feedback loops: simulation and experimental analysis. *Mol. Syst. Biol.* **2**, 41 (2006).
24. Kalir, S., Mangan, S. & Alon, U. The coherent feed-forward loop with a SUM input function prolongs flagella production in *Escherichia coli*. *Mol. Syst. Biol.* **1**, 2005.0006 (2005).
25. Maeda, Y. T. & Sano, M. Regulatory dynamics of synthetic gene networks with positive feedback. *J. Mol. Biol.* **359**, 1107–1124 (2006).
26. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
27. Kramer, B. P. & Fussenegger, M. Hysteresis in a synthetic mammalian gene network. *Proc. Natl Acad. Sci. USA* **102**, 9517–9522 (2005).
28. Wolf, D. M. & Arkin, A. P. Motifs, modules and games in bacteria. *Curr. Opin. Microbiol.* **6**, 125–134 (2003).
29. Yuh, C. H., Bolouri, H. & Davidson, E. H. Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* **279**, 1896–1902 (1998).
30. Buchler, N. E., Gerland, U. & Hwa, T. On schemes of combinatorial transcription logic. *Proc. Natl Acad. Sci. USA* **100**, 5136–5141 (2003).
31. Kalir, S. & Alon, U. Using a quantitative blueprint to reprogram the dynamics of the flagella gene network. *Cell* **117**, 713–720 (2004).
32. Setty, Y., Mayo, A. E., Surette, M. G. & Alon, U. Detailed map of a *cis*-regulatory input function. *Proc. Natl Acad. Sci. USA* **100**, 7702–7707 (2003).
33. Ma, H. W. *et al.* An extended regulatory network of *Escherichia coli* and analysis of its hierarchical structure and network motifs. *Nucleic Acids Res.* **32**, 6643–6649 (2004).
34. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA* **100**, 11980–11985 (2003).
35. Basu, S., Mehreja, R., Thiberge, S., Chen, M. T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl Acad. Sci. USA* **101**, 6355–6360 (2004).
36. Mangan, S., Zaslaver, A. & Alon, U. The incoherent feed-forward loop accelerates the response-time of the *gal* system of *Escherichia coli*. *J. Mol. Biol.* **356**, 1073–1081 (2006).
37. Hornstein, E. & Shomron, N. Canalization of development by microRNAs. *Nature Genet.* **38**, S20–S24 (2006).
38. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate *E2F1* expression. *Nature* **435**, 839–843 (2005).
39. Johnston, R. J. Jr. *et al.* An unusual Zn-finger/FH2 domain protein controls a left/right asymmetric neuronal fate decision in *C. elegans*. *Development* **133**, 3317–3328 (2006).
40. Ghosh, B., Karmakar, R. & Bose, I. Noise characteristics of feed forward loops. *Phys. Biol.* **2**, 36–45 (2005).
41. Wall, M. E., Dunlop, M. J. & Hlavacek, W. S. Multiple functions of a feed-forward-loop gene circuit. *J. Mol. Biol.* **349**, 501–514 (2005).
42. Hayot, F. & Jayaprakash, C. A feedforward loop motif in transcriptional regulation: induction and repression. *J. Theor. Biol.* **234**, 133–143 (2005).
43. Ishihara, S., Fujimoto, K. & Shibata, T. Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes. *Genes Cells* **10**, 1025–1038 (2005).
44. Kashtan, N., Itzkovitz, S., Milo, R. & Alon, U. Topological generalizations of network motifs. *Phys. Rev. E* **70**, 031909 (2004).
45. Dobrin, R., Beg, Q. K., Barabasi, A. L. & Oltvai, Z. N. Aggregation of topological motifs in the *Escherichia coli* transcriptional regulatory network. *BMC Bioinformatics* **5**, 10 (2004).
46. Ronen, M., Rosenberg, R., Shraiman, B. I. & Alon, U. Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics. *Proc. Natl Acad. Sci. USA* **99**, 10555–10560 (2002).
47. Zaslaver, A. *et al.* Just-in-time transcription program in metabolic pathways. *Nature Genet.* **36**, 486–491 (2004).
48. Kalir, S. *et al.* Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* **292**, 2080–2083 (2001).
49. Laub, M. T., McAdams, H. H., Feldblyum, T., Fraser, C. M. & Shapiro, L. Global analysis of the genetic network controlling a bacterial cell cycle. *Science* **290**, 2144–2148 (2000).
50. McAdams, H. H. & Shapiro, L. A bacterial cell-cycle regulatory network operating in time and space. *Science* **301**, 1874–1877 (2003).
51. Spellman, P. T. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–3297 (1998).
52. Ingram, P. J., Stumpf, M. P., Stark, J. Network motifs: structure does not determine function. *BMC Genomics* **5**, 108 (2006).
53. Pilpel, Y., Sudarsanam, P. & Church, G. M. Identifying regulatory networks by combinatorial analysis of promoter elements. *Nature Genet.* **29**, 153–159 (2001).
54. Beer, M. & Tavazoie, S. Predicting gene expression from sequence. *Cell* **117**, 185–198 (2004).
55. Rosenfeld, N. & Alon, U. Response delays and the structure of transcription networks. *J. Mol. Biol.* **329**, 645–654 (2003).
56. Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M. & Teichmann, S. A. Structure and evolution of transcriptional regulatory networks. *Curr. Opin. Struct. Biol.* **14**, 283–291 (2004).
57. Yu, H., Luscombe, N. M., Qian, J. & Gerstein, M. Genomic analysis of gene expression relationships in transcriptional regulatory networks. *Trends Genet.* **19**, 422–427 (2003).
58. Luscombe, N. M. *et al.* Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* **431**, 308–312 (2004).
59. Amit, I. *et al.* A module of negative feedback regulators defines growth factor signaling. *Nature Genet.* **39**, 503–512 (2007).
60. Davidson, E. H. *et al.* A genomic regulatory network for development. *Science* **295**, 1669–1678 (2002).
61. Longabaugh, W. J., Davidson, E. H. & Bolouri, H. Computational representation of developmental genetic regulatory networks. *Dev. Biol.* **283**, 1–16 (2005).
62. Johnston, R. J. Jr, Chang, S., Etchberger, J. F., Ortiz, C. O. & Hobert, O. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc. Natl Acad. Sci. USA* **102**, 12449–12454 (2005).
63. Xiong, W. & Ferrell, J. E. Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460–465 (2003).
64. Brandman, O., Ferrell, J. E. Jr, Li, R. & Meyer, T. Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science* **310**, 496–498 (2005).
65. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
66. Bolouri, H. & Davidson, E. H. Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. *Proc. Natl Acad. Sci. USA* **100**, 9371–9376 (2003).
67. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl Acad. Sci. USA* **102**, 3581–3586 (2005).
68. Rappaport, N., Winter, S. & Barkai, N. The ups and downs of biological timers. *Theor. Biol. Med. Model.* **2**, 22 (2005).
69. Conant, G. C. & Wagner, A. Convergent evolution of gene circuits. *Nature Genet.* **34**, 264–266 (2003).
70. Dekel, E. & Alon, U. Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588–592 (2005).
71. Ihmels, J. *et al.* Rewiring of the yeast transcriptional network through the evolution of motif usage. *Science* **309**, 938–940 (2005).
72. Madan Babu, M., Teichmann, S. A. & Aravind, L. Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J. Mol. Biol.* **358**, 614–633 (2006).
73. Dekel, E., Mangan, S. & Alon, U. Environmental selection of the feed-forward loop circuit in gene-regulation networks. *Phys. Biol.* **2**, 81–88 (2005).
74. Prill, R. J., Iglesias, P. A. & Levchenko, A. Dynamic properties of network motifs contribute to biological network organization. *PLoS Biol.* **3**, e343 (2005).
75. Kashtan, N., Itzkovitz, S., Milo, R. & Alon, U. Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics* **20**, 1746–1758 (2004).
76. Yeger-Lotem, E. *et al.* Network motifs in integrated cellular networks of transcription-regulation and protein–protein interaction. *Proc. Natl Acad. Sci. USA* **101**, 5934–5939 (2004).
77. Zhang, L. V. *et al.* Motifs, themes and thematic maps of an integrated *Saccharomyces cerevisiae* interaction network. *J. Biol.* **4**, 6 (2005).
78. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
79. Lahav, G. *et al.* Dynamics of the p53–Mdm2 feedback loop in individual cells. *Nature Genet.* **36**, 147–150 (2004).
80. Friedman, N., Vardi, S., Ronen, M., Alon, U. & Stavans, J. Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biol.* **3**, e238 (2005).
81. Hoffmann, A., Levchenko, A., Scott, M. L. & Baltimore, D. The IκB–NF-κB signaling module: temporal control and selective gene activation. *Science* **298**, 1241–1245 (2002).

82. Nelson, D. E. *et al.* Oscillations in NF-κB signaling control the dynamics of gene expression. *Science* **306**, 704–708 (2004).
83. Barkai, N. & Leibler, S. Circadian clocks limited by noise. *Nature* **403**, 267–268 (2000).
84. Tyson, J. J., Chen, K. C. & Novak, B. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **15**, 221–231 (2003).
85. Tyson, J. J., Csikasz-Nagy, A. & Novak, B. The dynamics of cell cycle regulation. *Bioessays* **24**, 1095–1109 (2002).
86. Pomerening, J. R., Sontag, E. D. & Ferrell, J. E. Jr. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nature Cell Biol.* **5**, 346–351 (2003).
87. Suel, G. M., Garcia-Ojalvo, J., Liberman, L. M. & Elowitz, M. B. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**, 545–550 (2006).
88. Ptacek, J. *et al.* Global analysis of protein phosphorylation in yeast. *Nature* **438**, 679–684 (2005).
89. Itzkovitz, S. *et al.* Coarse-graining and self-dissimilarity of complex networks. *Phys. Rev. E* **71**, 016127 (2005).
90. Ma'ayan, A. *et al.* Formation of regulatory patterns during signal propagation in a mammalian cellular network. *Science* **309**, 1078–1083 (2005).
91. Hertz, J., Krogh, A. & Palmer, R. G. *Introduction to the Theory of Neural Computation* (Perseus Books, Boulder, 1991).
92. Bray, D. Protein molecules as computational elements in living cells. *Nature* **376**, 307–312 (1995).
93. Lund, R. D. Synaptic patterns of the superficial layers of the superior colliculus of the rat. *J. Comp. Neurol.* **135**, 179–208 (1969).
94. White, E. L. *Cortical Circuits* (Birkhauser, Boston, 1989).
95. White, J., Southgate, E., Thomson, J. & Brenner, S. The nervous system of *Caenorhabditis elegans*. *Philos.Trans. R. Soc. London B Biol. Sci.* **314**, 1 (1986).
96. Itzkovitz, S. & Alon, U. Subgraphs and network motifs in geometric networks. *Phys. Rev. E* **71**, 026117 (2005).
97. Sporns, O. & Kotter, R. Motifs in brain networks. *PLoS Biol.* **2**, e369 (2004).
98. Sakata, S., Komatsu, Y. & Yamamori, T. Local design principles of mammalian cortical networks. *Neurosci. Res.* **51**, 309–315 (2005).
99. Song, S., Sjostrom, P. J., Reigl, M., Nelson, S. & Chklovskii, D. B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, e68 (2005).
100. Alon, U. Biological networks: the tinkerer as an engineer. *Science* **301**, 1866–1867 (2003).

Acknowledgements
I thank my laboratory members for stimulating discussions, especially R. Milo, S. Itzkovitz, S. Mangan, S. Shen-Orr and N. Kashtan. I thank M. Elowitz, M. Surette, S. Leibler and H. Westerhoff for discussions, and the Israel Science Foundation, the Human Frontiers Science Foundation, Minerva, National Institutes of Health (USA) and the Kahn Family Foundation for support

Competing interests statement
The author declares no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
E2F1 | FlhDC | FliA
UniProtKB: http://ca.expasy.org/sprot
ArgR | LexA | MYC | Spo0A

FURTHER INFORMATION
Uri Alon laboratory: http: www.weizmann.ac.il/mcb/UriAlon
Access to this links box is available online.