Living cells respond to information from their environment on the basis of the interactions of a large yet limited number of molecular species that are arranged in complex cellular networks (1, 2). A classic example of such biochemical computation is the chemotaxis behavior of Escherichia coli, which is mediated by a well-characterized signal transduction network (3). However, despite growing knowledge about the molecular components of the cell, the dynamics of even simple cellular networks are not well understood. For instance, a quantitative explanation of the high sensitivity and exact adaptation observed in bacterial chemotaxis is still lacking (3). Similarly, many other cellular networks, such as the ones responsible for signal transduction, regulation of gene expression, or metabolism, are poorly understood from a quantitative point of view. Thus, simple and modular experimental systems are needed to study how the genetic structure and connectivity of cellular networks are related to their function. To this end, we devised an in vivo synthetic system that enables the generation of combinatorial libraries of genetic networks.

We have generated a combinatorial library composed of a small set of transcriptional regulatory genes and their corresponding promoters with varying connectivity (Fig. 1). We chose genes of three well-characterized prokaryotic transcriptional regulators: LacI, TetR, and lambda CI (4). The binding state of LacI and TetR can be changed with the small molecule inducers, isopropyl β-D-thiogalactopyranoside (IPTG) and anhydrotetracycline (aTc), respectively. We also chose five promoters regulated by these proteins, which cover a broad range of regulatory characteristics such as repression, activation, leakiness, and strength. Two of the promoters are

from the Alfred P. Sloan Foundation (to M.K.C.). Work was done partially at Stanford Synchrotron Radiation Laboratory (SSRL), which is operated by the Department of Energy, Office of Basic Energy Sciences. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR07707. The beamline X4A at the National Synchrotron Light Source, a Department of Energy facility, is supported by the Howard Hughes Medical Institute. The authors thank K. Grenchuch at the Ohio State University Chemical Instrumentation Center for her assistance with mass spectrometric analysis. The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 1L2Q (NaCl form) and 1L2R ([NaCl]SO₄ form).

Supporting Online Material
www.sciencemag.org/cgi/content/full/296/5572/1462/DC1
Materials and Methods
Figs. S1 to S3
4 January 2002; accepted 27 March 2002
Fig. 2. Detailed analysis of two binary logical circuits (D038 and D052). (A) Cells of both E. coli host strains, lac– strain CMW101 and lac+ strain DH10B, transformed with each of two networks. "Logical circuit" behavior can be observed directly on agar plates (top), where fluorescence of colonies defines "on" and "off" states. Cells containing the indicated network were patched onto minimal agar media containing all four combinations of the two inducers in separate wells, as indicated. To increase fluorescent signal and to show that reporter expression is not cis-dependent, cells contained plasmids deleted for gfp and were cotransformed with compatible plasmids (~15 copies/cell) containing an equivalent P一般-yfp transcriptional unit. Cells were grown also in liquid culture and populations were analyzed with FACS for distributions of GFP expression (bottom). In each set of histograms, the blue curve shows the fluorescence distribution without inducers, the green curve shows when IPTG alone was present, the red curve indicates aTC alone, and the cyan curve shows the distribution when both inducers were present simultaneously. Single-peaked histograms differ most significantly by the presence or absence of both inducers. Mean fluorescence was used to define the "on" and "off" states. Cells of DH10B containing the indicated network were patched onto minimal plates containing the indicated inducers in separate wells. The output (fluorescence) is indicated in the lower rows by "On" or "Off." We do not distinguish here between the two inputs and, thus, between two different types of NOT IF logic functions. Colored bars act as legends for (A), (B), and (C), where the horizontal bars indicate the presence or absence of each inducer. The output (fluorescence) is indicated in the lower rows by "On" or "Off." The output (fluorescence) threshold parameter. A single universal threshold value was applied simultaneously to all networks. Sequencing was used to determine the connectivity of each of the two networks. Despite the fact that their logical behavior is different, both networks have the same connectivity ("topology"), as can be inferred from the corresponding diagrams of the interactions between the repressors and promoters. The schematic connectivity or topology diagrams shown at the bottom are identical for the two networks.

Fig. 3. Distribution of logical phenotypes in the two strains. (A) Definition of the logic operations performed by the circuits. In the top row, + and − indicate the presence or absence of each inducer input. The output (fluorescence) is indicated in the lower rows by "On" or "Off." The output (fluorescence) threshold parameter. A single universal threshold value was applied simultaneously to all networks. Besides being greater than this threshold, the minimal "on" value in each particular network was also required to be at least fourfold greater than the maximal "off" value. The difference in magnitudes on the x-axis for the two strains is due to different instrument settings.)

### Notes

- The processing of multiple chemical signals is a fundamental task performed by many natural biochemical networks. In order to screen for such behaviors, we incorporated in each plasmid a fourth transcriptional unit, in which green fluorescent protein (GFP) expression was controlled by the λ cI repressible promoter. The fluorescent signal acts as the network “output,” whereas the levels of the three chemical inputs were used as “inputs.” The plasmid library was transformed into two different host strains of E. coli (6), which differed most significantly by the presence of a wild-type copy of lacI at a chromosomal locus. Each clone was grown under four conditions, with and without IPTG and with or without aTC. GFP fluorescence was monitored simultaneously during cell growth (7). In this way, we searched the library for circuits in which the output is a binary logical function of both inducers. Examples of such “logical circuits” are NAND, NOR, or NOT IF (Figs. 2A and 3).

- We found that, in many cases, the output fluorescence levels of an individual culture in our library were sufficiently distinct for different inputs that an unambiguous binary output value for each input state could be assigned at a glance. In other cases, the designation was somewhat arbitrary. However, as shown in Fig. 3, it is quite remarkable that one can impose a single universal output-fluorescence threshold on the entire library and still obtain a large number of logical circuits for which “on” and “off” states differ significantly (8). Note that the
spectrum of logical behaviors differed in the two hosts (Fig. 3). This is not surprising, because the chromosomal lacI gene present in one strain (DH10B) acts as a network component and thus may drastically change the resulting phenotype (Fig. 2A).

Phenotypic variation in organisms often arises through mutations in the protein coding regions of the DNA. Another important contribution to phenotypic variability comes from changes in the cis-regulatory connections of existing genetic elements (9–10). To determine the origin of the phenotypic variability observed in the library, 30 clones with a variety of different behaviors were retransformed into both hosts, rescreened (to verify their phenotypic behavior), and sequenced (Fig. 4). We found a low level of point mutations, which, in some cases, modify the logical behavior of the networks (11). More interestingly, however, a large variety of behaviors remains among the many networks that do not have mutations in their regulatory regions. Sequencing allowed us to identify the three promoters incorporated in each plasmid; connectivity between different genetic elements varies from network to network so that 13 different “topologies” can be distinguished among the sequenced networks (Fig. 4). This variety of network connectivity is evidently the major source of phenotypic diversity in the library. In fact, the sequence data show that single step changes to the network connections, in which one promoter replaces another, frequently converted network operation from one logical function to another. For example, by replacing a single promoter in a network (D133 in Fig. 5A), which is always in the “on” state, one obtains a network (D038), which acts as a NOT IF in one strain and as a NAND in the other. Alternatively, by performing a different promoter replacement, one obtains network D016, which acts as a NOR circuit in both strains.

The fact that our restricted and conserved set of genes and promoters has the potential to switch among a variety of different computational functions by one-step changes in connectivity is noteworthy. From an evolutionary point of view, this observation suggests that once a simple set of genes and cis-regulatory elements is in place, it should be possible to jump from one functional phenotype to another using the same “toolkit” of genes (10) by modifying the regulatory connections. Such discontinuous changes, differ-
ent from the more gradual effects driven by successive point mutations, may be achieved in evolution by natural combinatorial mechanisms like transposition, recombination, or gene duplication.

Connectivity of a network does not uniquely determine its behavior. We found examples of networks that share the same connectivity but perform different logical operations. For instance, two networks depicted in Fig. 2 have the same connectivity (“topology,” see Fig. 2B) but show very different phenotypic behavior (Fig. 2A). We also found examples of networks with different connectivity that exhibited qualitatively similar behaviors. For instance, two networks shown in Fig. 5B (D016 and D052) both perform NOR operations despite their different connectivity. Thus, as shown by these two examples, the behavior of even simple networks built out of a few, well-characterized components cannot always be inferred from connectivity diagrams alone.

Can one predict the behavior of the logical circuits obtained here? “Boolean-type” models of gene regulation are often used to intuitively understand the operation of genetic networks. In this simplified description, one considers only discrete values of the biochemical variables and parameters (leading to commonly used reasoning, such as: “gene product A is produced, it inhibits the expression of gene product B, which is thus absent, etc. . .”) (12, 13). This description, though adequate for some of the present networks, seems not to apply to others (14). For instance, for the circuits in the lac- strain, shown in Fig. 2, the Boolean description is consistent with the NOT IF behavior of network D038 but not the NOR behavior of network D052 (15). It is possible that even for these well-studied transcriptional regulators subtle additional regulation may be at work among the plasmid-encoded elements (16). Genetic networks are nonlinear, stochastic systems in which the unknown details of interactions between components might be of crucial importance. Combinatorial libraries of simple networks should be useful in the future to uncover the existence of such additional regulation mechanisms and to explore the limits of quantitative modeling of cellular systems (17). For instance, it would be interesting to see whether the behavior of all the networks in the library could be described within a single theoretical model, a model defined by a unique set of parameters characterizing the interactions between the genetic components.

Combinatorial techniques inspired by recombination, such as DNA shuffling, have often proven successful in enhancing or changing the enzymatic activities of proteins and pathways (18, 19) without requiring an understanding of the mechanisms by which they work. However, combinatorial methods in simple and well-controlled systems, such as the one presented here, can and should also be used to gain better understanding of system-level properties of cellular networks. This is particularly important before using these powerful techniques more widely, e.g., in any practical applications.

The present results show that a handful of interacting genetic elements can generate a surprisingly large diversity of complex behaviors. Although the current system uses a small number of building blocks restricted to a single type of interaction (transcriptional regulation), both the number of elements and the range of biochemical interactions can be extended by including other modular genetic elements. The approach can be taken beyond the intracellular level by linking input and output through cell-cell signaling molecules, such as those involved in quorum sensing. Lastly, this combinatorial strategy can be used to search for other dynamic behaviors such as switches, sensors, oscillators, and amplifiers, as well as for high-level structural properties, such as robustness or noise-resistance (20).

References and Notes
4. In order to avoid toxicity effects due to overexpression and reductions in the dynamic regulatory range, we used the reduced-stability variants of these transcriptional regulators described previously (21).
5. The following promoters were used. Podd0705 (designated P23) from plasmid pO D3 70.5 (O2); P23 (P1) and PA (P2) were obtained from lambda phage DNA, and P2lacO1 (P1) and P2tetO1 (P2) from their original plasmids (23). Each of these promoters was amplified by polymerase chain reaction (PCR). In the case of P2lacs the 43 nt mutation, D41 which eliminates repression of P23 at highLI concentrations, was introduced by PCR.
6. Host strains were CMW101 (lacT, tetR) and DH10B (lacT, tetR) (24). Recombinant plasmids were constructed in CMW101 and 262 in DH10B. CMW101 was constructed by P1 transduction of a replica from CLC90 (a gift from T. J. Silhavy) to MC1061.
7. Individual colonies were grown overnight in 96-well round-bottom plates (Sarstedt 82.158.001, Newto, NC) on a microplate shaker at 30°C (180 μl well). Defined medium was used: 0.5 g (NH4)2SO4, 5.25 g KH2PO4, 0.225g MgSO4-7H2O, 19 mg EDTA, 2.5 mg FeSO4 per 500 ml of H2O, pH 6.8, with 0.5% glycerol and 0.5% Casamino acids. IPTG was used at 1 mM, and atC at 100 ng/ml. When required, ampicillin (100 μg/ml) and chloramphenicol (30 μg/ml) were added. All experiments were conducted in triplicate and in the dark. Overnight cultures were diluted 1:360 into fresh media with each of four inducer combinations. We measured GFP fluorescence and optical density (OD) four to five times along the growth curve, using a Wallac Victor2 multiwell-fluorimeter (Turku, Finland). Absorbance measurements were made at 600 nm (10 nm bandwidth, 0.1 integration time). Fluorescence readings for strain CMW101 used the settings as follows: CW-lamp filter f485 and emission filter f535; integration time, 0.1 s; measured GFP-CW-lamp energy, 7000; emission side, above. In the screen of the DH10B library, the settings were changed as follows: emission aperture, band; CW-lamp energy, 2.1673; emission side, above.
8. Furthermore, the value of this threshold can be changed many times without much affecting the distribution of logic functions obtained. These results indicate that in general the “on” and “off” output states of the networks are well separated. The detailed distributions of logical circuits behaviors obtained from the library may depend on aspects of the construction process, including mutations during PCR (11).
11. We used sequencing primers located about 60 base pairs into the coding regions of the genes to check the integrity of the upstream regions that contain the important gene control elements: ribosome-binding sites (RBS), promoters, terminators, and translational stop codons. Sequencing revealed that the construction method had introduced a number of mutations (we detected up to three) in some of the final plasmids. In at least three cases, this altered the behavior of the networks. On the basis of phenotypic behavior, 30 networks were chosen for sequencing. From this set, eight pairs were identified as having identical behavior, such five of these pairs exhibited matching behaviors, three differed. Detailed sequence analysis revealed several mutations that may explain the difference in behavior: one insertion in the proximal promoters, point mutations within terminator sequences, and point mutations that affected amino acid coding. Among the networks described in Fig. 5, only D016 contained a point mutation in the tet terminator; none were found in the others. Nevertheless, in order to check whether such mutations have a significant impact on network function, we deliberately reconstructed [by ligation] independent clones of selected networks. In each set, most of the clones shared the same behavior in strain DH10B. For example, the results for networks shown in Fig. 5 are as follows: five out of five reconstructed D016 clones exhibited a NOR phenotype, five out of seven D038 clones behaved as NAND and six out of seven D052 clones behaved as NOR. These values are typical. Thus the observed phenotypic variability is due to both permutation of promoters and a low level of mutation in some of the networks.
14. Boolean-type models neglect many potentially important intracellular phenomena, including stochastic fluctuations in the levels of components and the details of the intrinsic biochemistry and interactions. Nevertheless within this framework, we enumerated all Boolean network structures theoretically possible in the library (for the lac- strain CMW101). Only three network structures depend on both inputs, and they all behaved as NOT IF. However, experimentally we found in addition NORs.
15. We can easily make sense of the NOT IF behavior of the first network (D038) in the lac- strain. Negative autoregulation of TetR maintains LacI levels below the activity threshold of the P2 promoter. Therefore, in order to turn off α CI, and hence turn on the expression of gfp, expression of lacI must be induced by atc and not blocked by IPTG. The use of reporter plasmids for each of the promoters in the network showed that their activities are all consistent with this interpretation. Data not shown; our set of reporter plasmids was based on the p15A origin of replication and contained a chloramphenicol resistance gene, along with a promoter from the set driven by p2 expression.) When this plasmid was transformed into the lac- strain, and the ‘NAND’ type behavior, which could not be explained by similar arguments, was observed. The behavior of the second network (D052) in Fig. 2B, expressed in the lac- strain, is not deducible from simple Boolean analysis despite its structural similarity to the first network. A crucial difference is that the threshold for repression of P2 is known to be lower than that for P1. Therefore, the feedback loop leads to sufficient repression of αC to block production. However, when IPTG is added to the system, αCI levels evidently increase. 16. RNA polymerase read-through, DNA looping, little-studied effects of low levels of antibiotics on repression levels (25), and interactions between wild-type and destabilized alleles of lacI could have significant effects in this system.
17. Similarly, a recent quantitative study of gene expres-
An LRR Receptor Kinase Involved in Perception of a Peptide Plant Hormone, Phytosulfokine

Yoshikatsu Matsubayashi,* Mari Ogawa, Akiko Morita, Youji Sakagami

The sulfated peptide phytosulfokine (PSK) is an intercellular signal that plays a key role in cellular dedifferentiation and proliferation in plants. Using ligand-based affinity chromatography, we purified a 120-kilodalton membrane protein, specifically interacting with PSK, from carrot microsomal fractions. The corresponding complementary DNA encodes a 1021-amino acid receptor kinase that contains extracellular leucine-rich repeats, a single transmembrane domain, and a cytoplasmic kinase domain. Overexpression of this receptor kinase in carrot cells caused enhanced callus growth in response to PSK and a substantial increase in the number of tritium-labeled PSK binding sites, suggesting that PSK and this receptor kinase act as a ligand-receptor pair.

Evidence for the existence of high-affinity binding sites for PSK has been provided by binding assays with radiolabeled PSKs (6, 7). The observed binding is saturable, reversible, and localized in plasma membrane fractions. In addition, photoaffinity cross-linking analysis has shown that the putative receptors for PSK in rice plasma membrane are 120- and 160-kD glycosylated proteins (8). In the present study, we performed purification, molecular cloning, and functional expression of the PSK receptor to gain further insight into the molecular basis of signal transduction triggered by PSK.

For the purification of membrane proteins that specifically interact with PSK, we used microsomal fractions derived from the carrot cell line NC, which has been found to contain a relatively high concentration of high-affinity PSK-binding proteins: ~150 fmol per mg of microsomal proteins, with a dissociation constant ($K_d$) of $4.2 \pm 0.4$ nM, as determined by a $[^{35}S]$PSK binding assay (9) (Fig. 1A and B). This $K_d$ value is consistent with physiological concentrations of PSK in carrot suspension cultures (10), and it is also consistent with the PSK concentration that induces a 50% cell division of dispersed mesophyll cells (3). Photoaffinity labeling of NC membrane proteins with a photoactivatable PSK analog (8) indicated that a 120-kD protein and a minor 150-kD protein specifically interact with PSK (Fig. 1C). Both proteins contain ~10 kD of N-linked oligosaccharide chains that can be cleaved by treatment with peptide N-glycosidase F (PNGase F) (Fig. 1C).

We purified these PSK-binding proteins from the microsomal fractions of NC cells by Triton X-100 solubilization and specific ligand-based affinity chromatography using a [Lys$^5$]PSK-Sepharose column containing a long spacer chain between the ligand and matrix (9) (fig. S1). Elongation of the Lys$^5$ side chain of [Lys$^5$]PSK does not interfere with its binding affinity and specificity (11). Proteins specifically eluted by PSK were further purified by hydroxyapatite column chromatography and concentrated by ultrafiltration (9). SDS–polyacrylamide gel electrophoresis (PAGE) and Nile red staining of the proteins in the fractions eluted by PSK showed specific recovery of a major 120-kD protein and a minor 150-kD protein (Fig. 1D). Both of these proteins were absent in the fractions eluted by [2-5]PSK, a synthetic analog of PSK with no biological or binding activities (12) (Fig. 1D). PNGase F treatment of these two proteins decreased their apparent sizes to 110 and 140 kD, respectively, suggesting that they are identical to the proteins we detected in photoaffinity cross-linking experiments (Fig. 1D; see also Fig. 1C).

Four independent purifications were performed, yielding 50 μg of the major 120-kD protein from 4800 mg of microsomal proteins, with an overall recovery rate of 40%. The protein was digested with TPCK-trypsin (TPCK, tosyl phenylalanyl chloromethyl ketone), and peptide fragments thus generated were separated by reversed-phase high-performance liquid chromatography (HPLC) (9) (Fig. E). We analyzed the fragments of the 120-kD protein contained in 15 independent peaks, using a protein sequencer and MALDI-TOF MS (matrix-assisted laser desorption/ionization time-of-flight mass spec-