Daou et al.[2] also showed that nocifensive behaviors could be reduced by treatment with the analgesics morphine and pregabalin, suggesting that the strategy developed by the two groups will be useful for medium-throughput screening of compounds designed to modulate peripheral mechanisms of pain signaling and associated changes in the central nervous system. For drug screening applications, the relatively stable gene expression afforded by genetically engineered mice has advantages over viral delivery methods. Although viruses permit rapid introduction of a gene, the resulting expression levels are often variable between animals and over time. Indeed, Iyer et al.[1] observed transgene shutdown five weeks after virus injection.

The new strategy will also help researchers identify the particular neuronal subsets responsible for transmitting different pain signals. Many mouse strains expressing Cre recombinase under the control of promoters specific for particular neuronal subsets are available, and many more are under development. By combining these Cre strains with the methods of Iyer et al.[1] or Daou et al.[2], one could conditionally express transgenic opsins in particular neuronal subsets and stimulate the subsets optogenetically to analyze their involvement in acute and chronic pain. A related area of investigation is the downstream circuitry of central pain pathways. For example, using c-Fos expression as a readout of activity, Daou et al.[2] detected connectivity between $Na_V1.8^+$ neurons and second-order spinal neurons in the superficial lamina of the dorsal horn.

Looking beyond the use of microbial opsins, it should also be possible to confer optical control on endogenous proteins involved in pain signaling with other tools, such as light-gated P2X receptor ion channels or acid-sensing ion channels[7]. Optogenetic control of peripheral

neurons could also enable the study of cells and pathways involved in sensory modalities other than pain, such as temperature, touch and itch, and in the control of autonomic and somatic motor axons.

The therapeutic relevance of optogenetics is difficult to gauge at this early stage of technology development. Neuromodulatory approaches that use electrical impulses are already widely used to treat disease, but all involve untargeted stimulation of nerves rather than selective modulation of particular neuronal subsets. The capacity to therapeutically hyperpolarize axons in a targeted fashion, thereby shutting down conduction of action potentials, is undeniably exciting. That said, clinical development of optogenetic proteins would face many technical, practical and regulatory barriers[8]. Iyer et al.[1] did consider some of these challenges in their experimental design, and their study provides at least a proof of principle for optogenetic modulation as a pain therapy. In particular, they showed that optogenetic silencing of peripheral nerve terminals attenuated sensitivity to mechanical and thermal stimuli in a model of chronic constriction injury. The viral vector they chose was AAV6, which is currently being used in clinical trials.

Nevertheless, establishing the efficacy and safety of an optogenetic gene therapy would surely encounter daunting obstacles. Some of these are common to gene therapy as a whole, such as achieving therapeutically relevant levels and duration of gene expression and circumventing the immunogenicity of the viral vector and exogenous proteins. Other important considerations are whether non-native opsins have adverse effects in human cells and whether viral delivery is completely restricted to nociceptors, which would be necessary to avoid optogenetic interference with other sensory modalities and motor functions.

In recent years, interest in optogenetics as a means of studying neuronal circuitry and as a neuromodulatory therapy has moved beyond the brain to the peripheral nervous system[8]. The approach developed by Iyer et al.[1] and Daou et al.[2] should provide an exciting way of evaluating the function of optogenetic tools in vivo because nocifensive behavior provides a robust readout. The number of tools for optical control of biological processes is quickly increasing, and each tool differs in properties such as activation spectra, light sensitivity, kinetics, cellular localization and downstream effects[9,10]. For rapid evaluation of new optogenetic tools, the viral delivery approach established by Iyer et al.[1] can be accomplished in a few weeks as opposed to the months needed to generate a new transgenic mouse line—a great advantage.

In the near term, the new optogenetic approaches will likely have the most impact in preclinical research aimed at understanding the biology of pain, such as which fibers are responsible for maintaining different pain states and the specific contribution of peripheral input to ongoing pain. By shining light on peripheral pain pathways, we will now be able to unravel exactly how pain begins and how we might shut it off.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Iyer, S.M. et al. Nat. Biotechnol. **32**, 274–278 (2014).
2. Daou, I. et al. J. Neurosci. **33**, 18631–18640 (2013).
3. Towne, C., Pertin, M., Beggah, A.T., Aebischer, P. & Decosterd, I. Mol. Pain **5**, 52 (2009).
4. Stirling, L.C. et al. Pain **113**, 27–36 (2005).
5. Madisen, L. et al. Nat. Neurosci. **15**, 793–802 (2012).
6. Woolf, C.J. Nature **306**, 686–688 (1983).
7. Browne, L.E. et al. Proc. Natl. Acad. Sci. USA **111**, 521–526 (2014).
8. Williams, J.C. & Denison, T. Sci. Transl. Med. **5**, 177ps176 (2013).
9. Yizhar, O., Fenno, L.E., Davidson, T.J., Mogri, M. & Deisseroth, K. Neuron **71**, 9–34 (2011).
10. Szobota, S. & Isacoff, E.Y. Optical control of neuronal activity. Annu. Rev. Biophys. **39**, 329–348 (2010).

E. coli network upgrade

Ulrich Stelzl

Combining yeast two-hybrid data with mass-spectrometry and structural data provides a detailed view of the E. coli interactome.

Protein interaction networks are crucial for understanding complex cellular systems, from how signals are processed and lead to cellular

Ulrich Stelzl is at the Max Planck Institute for Molecular Genetics (MPIMG), Berlin, Germany.
e-mail: stelzl@molgen.mpg.de

phenotypes to the function of individual network components under specific conditions. High-quality, binary protein-protein interaction maps have been generated for human cells and for many model organisms, including yeast, worm, fly and Arabidopsis thaliana. Now, Escherichia coli has joined this group

with a study by Rajagopala et al.[1] in this issue. Using yeast two-hybrid analysis, the authors construct a binary interaction network for E. coli that contains 2,234 protein-protein interactions between 1,850 proteins. These data, which include both transient interactions and more stable interactions within protein complexes, are then integrated with previous data on protein complexes and protein structure to produce a high-resolution E. coli interactome map—the most detailed map of its kind for any organism.

There are two main techniques for systematic protein interaction mapping—yeast two-hybrid and affinity purification or co-fractionation

Figure 1 The path toward a detailed, high-resolution *E. coli* interactome map starts by combining binary protein-protein interaction data from yeast two-hybrid analysis and data from affinity purification–MS[5–7]. This provides a combined set of *E. coli* interactions and internal topology for some of the complexes that had been identified by affinity purification–MS. The complexes are annotated with previous knowledge of protein structures, using Interactome3D[10], and the annotated complexes are added to the combined interactome to provide a high-resolution *E. coli* interactome map.

coupled to mass spectrometry (MS). The methods yield different types of information: yeast two-hybrid analysis reveals binary interactions, including transient interactions, whereas the purification–MS approaches report multiple interactions connecting all the proteins in fairly stable complexes. Because of these and other differences, data sets from the two methods may have little overlap. For example, in recent studies of human cells, a yeast two-hybrid study implicated methyltransferase proteins in several cytoplasmic functions[2] whereas a large-scale co-fractionation–MS analysis showed these proteins to be involved in transcription or chromatin-remodeling complexes[3]. Such discrepancies suggest that a more comprehensive picture of cellular networks can be obtained by merging data from the two methods.

Protein-protein interaction networks have found application in both basic and translational research. In the study of disease, human protein-protein interaction data can be used to identify functionally important genetic variants because, for a particular phenotype, these tend to be found in proteins that are closely connected to each other in an interaction network. Differential network analysis—that is, measuring or inferring changes in networks—can reveal the effects of protein sequence variation or environmental changes. For model organisms such as yeast, protein-protein interaction networks have led to insights into the principles of molecular wiring and relationships among cellular processes. For bacteria, however, progress in developing interactome maps has been slower. Several groups have produced large-scale affinity purification–MS data for *Mycoplasma pneumoniae*[4] and *E. coli*[5–7], but apart from early yeast two-hybrid maps for *Helicobacter pylori*[8] and *Campylobacter jejuni*[9], there has been a lack of bacterial, binary protein-protein interaction networks.

Rajagopala *et al.*[1] make several important advances. First, the authors screened 3,600 *E. coli* proteins (~70% of the proteome) for pair-wise, protein-protein interactions, resulting in 2,234 interactions involving about half of the proteins screened. To obtain a more comprehensive binary interactome for *E. coli*,

they merged their yeast two-hybrid data with interactions from previous studies, including affinity purification–MS data sets (**Fig. 1**). The combined interactome contains 3,946 binary interactions, which they estimate is ~37% of the *E. coli* binary interactome.

Second, the authors exploited the complementary information that affinity purification–MS and yeast two-hybrid data provide about protein complexes. Affinity purification–MS data identify proteins bound together in complexes but usually do not resolve individual subunit-subunit interactions within these complexes. By contrast, binary interaction data from yeast two-hybrid analysis identifies the subunit-subunit interactions but not necessarily all the members of a complex. By combining binary interactions (from their work and the literature) with information on complex composition from previous affinity purification–MS studies, the authors deduced at least part of the internal complex topology for more than 100 multiprotein complexes (**Fig. 1**). Finally, the authors incorporated existing protein structural data into their

interactome map. By feeding the binary interaction data into a computational tool, Interactome 3D[10], they were able to propose three-dimensional structural models for some complexes lacking co-crystal structures. The result is the first high-resolution view of the *E. coli* interactome (**Fig. 1**).

The field of protein-protein interaction mapping has struggled with standardized quality measures for interaction data because of the different characteristics of the data generated by different approaches. The performance of a method can be assessed by measures of sensitivity (how many true interactions do you find?) and specificity (how many false interactions do you find?). Following a concept put forward by Vidal and colleagues[11], sensitivity can be assessed empirically by examining a set of 'true' interactions (i.e., previously well-characterized interactions) and specificity can be assessed by examining a set of randomly generated interactions (i.e., likely not genuine interactions). However, unlike 'average' interactions, the interactions in the true reference set have been reported in the literature many
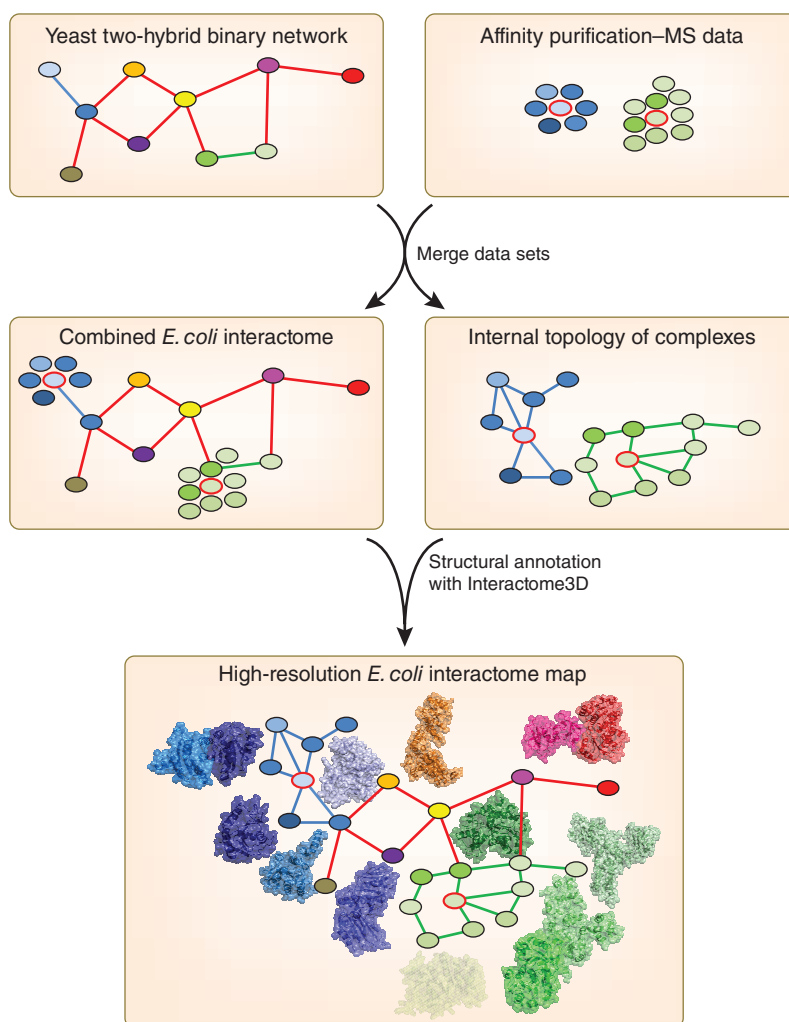
times; because these interactions are relatively easy to find, the sensitivity of a method is likely to be overestimated. A potential pitfall of testing random interactions is related to the fact that biological networks are scale free—that is, many proteins have one or few interactions, and few proteins are 'hubs' with many interactions. A protein that has a very large number of interactors might have a much higher probability of giving an interaction in a set of randomly selected set of proteins.

Rajagopala et al.[1] were able to overcome some of these limitations by testing sensitivity and specificity in their full interactome data set in addition to doing smaller benchmarking experiments. None of 500 random interactions ended up in the final data, so the false-positive rate was below 0.2%, and the sensitivity was estimated as <21%. These numbers are typical for current interactome mapping approaches. Sensitivities below 20% emphasize the need to develop more efficient methods and highlight why large-scale interaction screens should combine different experimental methods. Such methods might include luciferase-based co-immunoprecipitation (LUMIER) approaches and mammalian interaction trap assays (MAPPIT) or the development of large-scale cross-linking mass spectrometry. Clearly, the problem of obtaining comprehensive interactome coverage grows with increasing proteome complexity and dynamics in higher organisms.

The new data from Rajagopala et al.[1] will support attempts to delineate detailed molecular mechanisms for core biological machines that are typically studied in E. coli. Notably, a large number of E. coli proteins remain uncharacterized, and protein-protein interaction data can assist in linking them to fundamental cellular processes. For example, systematically generated interaction data implicated RsfA in ribosomal protein translation, leading to its subsequent characterization as a nutrition-dependent ribosomal silencing factor[12]. Furthermore, E. coli is crucial in synthetic biology, both as a host to test novel circuitries and as a source for reconstituting minimal systems that reproduce essential functions. Information from the protein-protein interaction network might help system design to avoid side reactions or suggest novel connections for new synthetic systems. The E. coli network upgrade may also stimulate large-scale modeling approaches for this model organism, providing a chance to link genome-scale models of metabolism to other cellular processes.

Three-dimensional high-resolution network views, as presented here for E. coli, are an important step on the way to interactome maps that are comprehensive and capture condition-specific properties. It will be easier to test hypotheses based on this resource, thanks to the availability of three-dimensional models of protein-protein interfaces; interfaces can be targeted—for example, through mutation—without affecting other functions of the interacting proteins. As has been shown by structural annotation of human protein networks, assessing the function of interactions rather than of proteins is crucial for understanding specific phenotypes in health and disease.

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

1. Rajagopala, S.V. et al. Nat. Biotechnol. 32, 285–290 (2014).
2. Weimann, M. et al. Nat. Methods 10, 339–342 (2013).
3. Havugimana, P.C. et al. Cell 150, 1068–1081 (2012).
4. Kühner, S. et al. Science 326, 1235–1240 (2009).
5. Butland, G. et al. Nature 433, 531–537 (2005).
6. Arifuzzaman, M. et al. Genome Res. 16, 686–691 (2006).
7. Hu, P. et al. PLoS Biol. 7, e96 (2009).
8. Rain, J.C. et al. Nature 409, 211–215 (2001).
9. Parrish, J.R. et al. Genome Biol. 8, R130 (2007).
10. Mosca, R., Ceol, A. & Aloy, P. Nat. Methods 10, 47–53 (2013).
11. Venkatesan, K. et al. Nat. Methods 6, 83–90 (2009).
12. Hauser, R. et al. PLoS Genet. 8, e1002815 (2012).

You are what you eat

Harald Brüssow & Scott J Parkinson

Eating foods rich in plant fiber promotes health by changing the composition and metabolic products of gut bacteria.

The health benefits of eating plant fiber have long been appreciated. Epidemiological studies show an inverse relationship between dietary plant fiber and the risk of heart disease, obesity and type-2 diabetes, and consumption of dietary fiber is recommended by the American Dietetic Association and other health organizations. But the biological mechanisms underlying the health effects of dietary fiber have been hard to pin down. Gut microbes are presumed to be involved because they contribute to fiber digestion. Two recent papers in Nature[1] and Cell[2] shed light on their roles, demonstrating

Harald Brüssow is at the Nestlé Research Center, Lausanne, Switzerland, and
Scott J. Parkinson is at the Nestlé Institute of Health Sciences, Lausanne, Switzerland.
e-mail: harald.bruessow@rdls.nestle.com or scottjames.parkinson@rd.nestle.com

that ingestion of plant fiber induces rapid shifts in the composition and function of the gut microbiota and that metabolites produced by the microbiota support metabolic health by regulating glucose control in the host. These findings may guide efforts to design 'prebiotics'—dietary supplements capable of manipulating the gut microbiota to improve human health.

Fermentation of soluble fibers such as fructo-oligosaccharides and galacto-oligosaccharides by gut bacteria produce the short-chain fatty acids acetate, propionate and butyrate. Although chemically similar, short-chain fatty acids are metabolized differently and exert very different effects on host physiology (**Fig. 1**). Acetate (the most abundant short-chain fatty acid) is a substrate for hepatic de novo lipogenesis and cholesterol biosynthesis, propionate is a substrate for hepatic gluconeogenesis

and butyrate acts as an energy substrate for enterocytes lining the colon. Propionate and butyrate also act as signaling molecules by binding to and activating the G protein–coupled receptors FFAR2 and FFAR3 (free fatty acid receptors).

It is well established that changing the diet changes the composition of the gut microbiota[3]. David et al.[1] explored this phenomenon by measuring the rates at which changes in dietary fiber intake translate into changes in the composition and transcriptional profile of the gut microbiota in humans. To this end they fed ten healthy human volunteers either a plant-based or an animal-based diet for five days. At various points in the study, they analyzed fecal samples by sequencing 16S ribosomal RNA (to determine the relative abundance of different microbes) and by RNA-seq (to determine the relative expression of microbial genes).

Changes in microbial gene expression and community structure were seen within days of starting the diets. In subjects fed the animal-based diet, these changes correlated with a reversible physiological response (weight loss) and an increase in the abundance of bile-tolerant bacteria and of short-chain fatty acids indicative of amino acid fermentation.