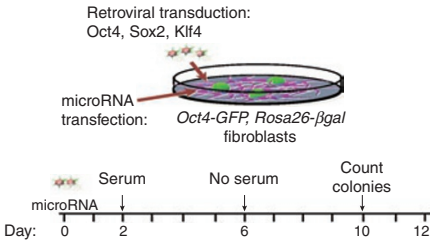
STEM CELLS

Micro-reprogramming

Researchers use microRNAs to more efficiently generate induced pluripotent stem cells in the mouse.

The search for methods to reprogram somatic cells to pluripotency without permanent modification of the genome is ongoing. In a recent report, Robert Blelloch and colleagues at the University of California, San Francisco, add microRNAs to the growing list of factors that can increase the efficiency of this process (Judson *et al.*, 2009).

Blelloch and colleagues came to this discovery based on their prior work on mouse embryonic stem cells (mESCs). These cells are known to proliferate rapidly, which they achieve largely by bypassing a checkpoint at the G1-S transition of the cell cycle. In work on *Dgcr8* knockout mESCs, in which all canonical microRNAs are lost, Blelloch and colleagues had noted that microRNAs are essential for this process. They identified a key set of microRNAs, which they named the embryonic stem cell–specific cell cycle–regulating (ESCC) microRNAs, as promoters of the unique cell cycle of mESCs.



Using microRNAs to reprogram cells. Reprogramming factors and microRNA mimics are delivered to fibroblasts expressing a fluorescent Oct-4 reporter; GFP-positive colonies are scored. Image courtesy of Robert Judson.

"We think that the cell cycle and differentiation are linked," says Blelloch. "As an [embryonic stem] cell differentiates, it simultaneously extends the G1 phase, and this extension may be absolutely required for differentiation. Conversely, if you could induce an [embryonic stem cell]-like cell cycle in a differentiated cell, you could possibly promote de-differentiation. So we thought it would make a lot of sense if these microRNAs could promote reprogramming of a somatic cell."

Indeed, this is what the researchers saw when they examined the activity of the ESCC microRNAs in a fibroblast reprogramming assay. miR-294, as well as several other microRNAs that share the same seed sequence, increased the efficiency with which mouse embryonic fibroblasts could be reprogrammed with the transcription factors Oct-4, Sox-2 and Klf-4, by 10- to 50-fold, depending on the specific microRNA and the concentration at which it was expressed. Notably, however, microRNAs had no enhancing effect on four-factor reprogramming efficiency, when Myc was added back to the reprogramming cocktail. In addition, microRNAs could not replace either Oct-4, Sox-2 or Klf-4 in this process.

This led the researchers to think that the microRNAs might be acting downstream of Myc in promoting reprogramming, and

SYSTEMS BIOLOGY

A FUNCTIONAL BLUEPRINT OF *E. COLI*

Researchers integrate proteomics data with genomic-context analysis and develop a protein-function prediction tool to annotate functional orphans in *Escherichia coli*.

In this age of 'omics' technologies, it is perhaps a bit surprising that one-third of the 4,225 proteins found in *E. coli*, the most studied bacterium, are functionally uncharacterized. Andrew Emili of the University of Toronto, Gabriel Moreno-Hagelsieb of Wilfrid Laurier University and their colleagues wondered whether this was because these proteins had unusual properties or low expression that made them difficult to study, or whether these proteins, which are generally not major network nodes, were just not deemed 'popular' enough to focus research efforts on understanding what they do in the cell.

One way to infer molecular function is to use sequence alignments. But sequence alignments, Emili notes, "don't give you a lot of biological context, for instance, the process or pathway the protein works in." However, looking at 'whom' a protein 'hangs out' with in its natural environment can lead to clues about its function, which is why Emili and his colleagues decided to begin the process of annotating *E. coli* proteins with a proteome-wide analysis of protein-protein interactions. "Knowledge of the components in a complex is a stage towards understanding how that protein is positioned in terms of a biological pathway," he explains.

The researchers used a tandem affinity purification method to systematically discover protein-protein interactions in *E. coli*, a project on which Emili has been working with Jack Greenblatt, also at the University of Toronto, for many years. With this method, a purification tag is expressed as a fusion to a bait protein, which naturally associates with its interaction partners, the 'prey', in the cell. The tagged bait and any prey proteins are then isolated via a gentle, two-step purification, which maintains the endogenous interactions. The proteins in the complex can then be identified by mass spectrometry.

Emili and his colleagues performed such purifications with more than a thousand tagged bait proteins. In the end they identified nearly 6,000 pairwise interactions, about half of which were novel findings, including findings for 451 functional orphans. A clustering algorithm assigned many of the orphans to multiprotein complexes. However, they were unable to detect 469 of the orphan proteins, which were likely membrane-associated or present at very low abundance. Thus they used a complementary approach to look at the natural chromosomal clustering of *E. coli* genes, consisting of four different computational genomic-context profiling methods. These included looking at gene fusions, intergenic distances, the similarity between phylogenetic profiles and the evolutionary conservation of gene order. With this approach, they predicted pairwise interactions for most of the orphans.

indeed, bioinformatic analysis of existing sequence data for both Myc binding and for epigenetic marks at the promoter of the miR-290 cluster, supports this idea. Experimental testing of the possible mechanisms of microRNA function is ongoing, but Blelloch cautions that it will be difficult to tease out the function of individual microRNAs. "There are three microRNAs in the miR-290 cluster that share the same seed sequence, and that is just the beginning; there are other clusters as well, miR-17, for instance. And we know from our previous work that the system is incredibly redundant. In embryonic stem cells, we identified 11 microRNAs with the same seed sequence that were all interchangeable in their function," he says.

Could microRNAs replace all transcription factors in the reprogramming process? Blelloch is doubtful, but points out that even replacement of some of them would be beneficial, and screens for such factors are in progress in his laboratory. Others have also reported the use of both cell-penetrating recombinant proteins—in recent work from the laboratory of Sheng Ding at the Scripps Institute (Zhou *et al.*, 2009)—and small molecules, to achieve reprogramming. "Ultimately the goal is to use some mix of factors to replace DNA elements altogether," Blelloch states, "and I think that will happen, but what the final cocktail will be remains to be seen."

Natalie de Souza

RESEARCH PAPERS

Judson, R.L. *et al.* Embryonic stem cell–specific microRNAs promote induced pluripotency. *Nat. Biotechnol.* **27**, 459–461 (2009).

Zhou, H. *et al.* Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* **4**, 381–384 (2009).

Integrating the results from the proteomics experiments and genomic context analysis, the researchers generated a dataset of high-confidence pairwise interactions for 99% of the annotated and 96% of the unannotated genome. "Is protein-complex information or genomic context enough to tell you about a protein's role in the cell? No, but it certainly gives you hints," says Emili. They created a function prediction tool called StepPLR to assign putative functions to the orphans, using the information about 'whom' the orphans interacted with, directly and indirectly. Notably, they predicted that many of the orphans are actually involved in core cellular processes. Emili hopes that *E. coli* researchers with proteins or genes or pathways of interest will follow up on their functional predictions.

The researchers have set up a public resource called eNet to host their data, and they plan to keep adding to it and refining it. "It's not a complete story; we'd like to fill in the gaps. Certainly what's missing [from the proteomics data] is the membrane proteins," says Emili. Although similar resources exist for other model species, such as yeast, worm and fly, Emili acknowledges that bacteria have been largely understudied by genomics researchers. He hopes that this resource will help "bring in the bacterial community, making them aware of the things we can do with omics approaches."

Allison Doerr

RESEARCH PAPERS

Hu, P. *et al.* Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol.* **7**, e100096 (2009).

NEWS IN BRIEF

CHEMICAL BIOLOGY

Designing specificity

Designing molecules that specifically interact with only the intended biological target is a major challenge, especially in therapeutic applications. Grigoryan *et al.* describe a computational approach to design protein interaction specificity by maximizing the tradeoff between affinity and specificity. They used their approach to design highly selective peptide partners for 19 of 20 families of closely related human basic-region leucine zipper transcription factors.

Grigoryan, G. *et al. Nature* **458**, 859–864 (2009).

SENSORS

Fluorescent metabolite sensors

Brun *et al.* describe a general, modular approach for constructing fluorescence resonance energy transfer (FRET)-based metabolite-sensor proteins. The sensor consists of a Snap tag, a fluorescent protein, a metabolite-binding protein and a synthetic connector that contains both a fluorophore and a ligand that binds to the metabolite-binding protein. In the presence of a target metabolite, which displaces the connector ligand, the 'closed' sensor springs open and results in a change in the FRET efficiency.

Brun, M.A. *et al. J. Am. Chem. Soc.* **131**, 5873–5884 (2009).

GENOMICS

Identifying protein folding genes

Jonikas *et al.* describe a strategy to identify *Saccharomyces cerevisiae* genes involved in protein folding. They harnessed the transcription factor Hac1p, which activates the unfolded protein response, to drive expression of a GFP reporter. They introduced the reporter into ~4,500 deletion mutant strains and used flow cytometry to monitor single-cell fluorescence, thus identifying genes that either up- or downregulated the expression of the unfolded protein response reporter.

Jonikas, M.C. *et al. Science* **323**, 1693–1697 (2009).

PROTEIN BIOCHEMISTRY

A function for GFP

GFP, a protein found in the humble jellyfish, *Aequorea victoria*, and its fluorescent protein cousins have had a major impact on biological imaging. However, the biological functions of fluorescent proteins are not well-understood. Bogdanov *et al.* now report that GFPs can act as light-induced electron donors for various electron acceptors and suggest that they may play a role in cellular processes such as light sensing.

Bogdanov, A.M. *et al. Nat. Chem. Biol.* advance online publication (26 April 2009).

NANOTECHNOLOGY

Monitoring enzyme activity in real time

Orosco *et al.* describe a two-layer porous silicon nanoreactor as a label-free tool to monitor protease activity. The upper layer contains large pores, which trap the protease. The smaller reaction products filter down into the lower layer, which contains smaller pores. This causes a change in optical reflectivity of the silicon nanoreactor, allowing enzyme kinetics to be quantitatively observed in real time.

Orosco, M.M. *et al. Nat. Nanotechnol.* **4**, 255–258 (2009).