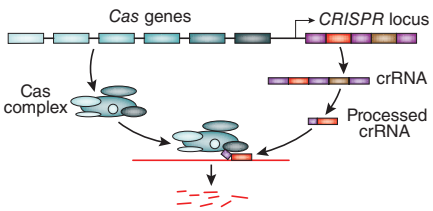
MICROBIOLOGY

Prokaryotic RNAi

Rationally designed RNAs that work with the bacterial immune response allow targeted gene silencing in prokaryotes.



Schematic of bacterial immune response.

Remember the time when the only way to alter expression of an endogenous gene in eukaryotes was to knock it out? Then came the discovery by Craig Mello and Andrew Fire in 1998 that a double-stranded RNA can silence a gene with a complementary sequence. This triggered the development of RNA interference as a tool, and with it came a myriad of investigations into the function of genes. For researchers working with prokaryotes it is now 1998.

Bacteria do not have a directly homologous RNAi machinery, and targeted modification of gene expression is thus not easily possible. Bacteria do, however, have their own immune system by which they recognize invading DNA and RNA and eliminate them. Rebecca and Michael Terns at the University of Georgia sought to harness members of this pathway as tools to target any RNA of choice.

Prokaryotic genomes incorporate short fragments of invading sequences into their clusters of regularly interspaced short palindromic repeat (*CRISPR*) locus. Long *CRISPR* RNA (crRNA) transcripts containing these invader-derived guide sequences are processed into shorter fragments that then interact with CRISPR-associated (Cas) proteins to form an effector complex. This Cas complex recognizes and silences a target complementary to the guide sequence.

Various modules of Cas proteins exist, one of them, the Cas RAMP (Cmr) complex consisting of six proteins, exclusively targets and cleaves RNA.

SYSTEMS BIOLOGY

A BIRD'S-EYE VIEW OF DISEASE

The integration of protein-protein interaction networks with structural information about the interacting protein partners creates a three-dimensional scaffold on which are mapped mutations involved in human disease.

The list of genetic mutations associated with human disease is long and continues to grow, and with it the challenges of deciphering what actually goes wrong in the cells, tissues and bodies of sick people. The problem is that the relationship between genotype and phenotype, in health or in disease, is typically very complex. It is the rare exception when a single mutation in a single gene causes disease.

One way to begin to probe this problem is to ask how disease mutations affect the way in which proteins interact. This is what Haiyuan Yu set out to do when he moved to Cornell University to start his laboratory. Yu and his colleagues began with a carefully curated map of directly interacting human proteins, using information both from the literature and from well-verified high-throughput yeast two-hybrid datasets. But just physical interaction data, Yu thought, was not enough. "When people talk about protein networks they have a sort of a mathematical view where proteins are dots, mathematical dots; there is no shape or structure," he explains. "But we know that's not true; in the cell the structure of the protein is fundamentally important in determining its function."

For those interacting human proteins for which crystal structures of the protein pair ('co-crystal' structures) have been solved, the researchers therefore added the interface structures to the network. But such information is not yet available for most putatively interacting human proteins. So Yu reached back to his bioinformatics-heavy past and used a homology modeling approach to predict the structures of interacting interfaces.

If for a particular human protein interaction there is a co-crystal structure of the orthologous yeast proteins, for example, Yu and colleagues used that information and the sequence homologies to predict the interface of the human proteins; notably,

To design crRNAs against a transcript of choice the Ternses and their colleagues first analyzed the features essential for crRNA/Cmr–mediated target recognition in *Pyrococcus furiosus*.

They immunoprecipitated the Cmr complex proteins and deep sequenced the short RNAs bound to them. The team found two predominant features on all crRNAs, an eight-nucleotide tag at the 5′ end, derived from the conserved bacterial *CRISPR* repeat, and ~30-nucleotide guide sequence, derived from the invading DNA. "The tag is critical," says Michael Terns; "just base-pairing of the crRNA is not sufficient." When the researchers removed the tag or even just mutated two of the eight nucleotides they no longer saw target cleavage.

The researchers tested whether they could target a sequence of choice. In *in vitro* assays they combined the six proteins of the Cmr complex with crRNAs complementary to a target mRNA and showed the predicted cleavage products.

The simplicity of the system makes it an attractive tool for prokaryotic gene silencing *in vivo*. Many prokaryotes already express the Cmr complex, and all one will need to do is introduce a plasmid encoding an engineered crRNA with the conserved tag and a 30-base-pair sequence complementary to any target.

The six Cmr proteins are expressed from the same operon and can be easily introduced into prokaryotes that do not have them, together with the crRNA of choice. "Because we have identified the components, it is straightforward to introduce it into another organism," says Rebecca Terns. Such *in vivo* experiments will explore the efficacy of crRNAs as analogs of small interfering RNAs and will also highlight any problems with off-target effects.

The potential of such a tool is hard to overstate. It opens applications from basic research into gene functions to metabolic engineering and the customization of microorganisms for the synthesis of a product of choice.

Nicole Rusk

RESEARCH PAPERS

Hale, C.R. *et al.* Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNA. *Mol. Cell* **45**, 292–302 (2012).

they first tested this strategy on existing co-crystal data to determine that it indeed makes reliable predictions. Combining both the experimental and predicted structural information with the binary interaction network produced the human structural interaction network (hSIN). Of its 4,222 binary interactions (between 2,816 proteins), about two-thirds of the interface structures are predicted by homology modeling. Then, with their scaffold built, Yu and colleagues combed through the Online Mendelian Inheritance in Man database and Human Gene Mutation Database and found more than 20,000 disease-associated mutations in about 600 genes that they could map to hSIN.

What did the scientists see in their bird's-eye view of disease? To mention just some of their observations, they note that missense mutations in disease-associated genes are enriched in protein interaction interfaces, they investigate the basis for pleiotropy and for locus heterogeneity effects, and they propose molecular hypotheses for the effects of disease-associated mutations. It should be noted that their observations apply to the subset of mutations that maps onto hSIN; the extent to which this analysis applies to other disease mutations remains to be seen. Finally, Yu and colleagues used hSIN to predict 292 new disease-associated genes, in 182 different diseases.

In the present work, the researchers validated experimentally only a few of their observations, using the yeast two-hybrid system, but Yu hopes to take this to an entirely different scale in the future. He envisions hSIN-derived predictions about the effects of disease mutations on protein interactions combined with rapid experimental testing of these predictions in high throughput.

There is a very long way to go before human disease can be understood mechanistically—for one thing, interaction networks such as hSIN do not take cell or tissue type into account—but at least the global effects of disease mutations on the generic protein interaction map are becoming a little clearer.

Natalie de Souza

RESEARCH PAPERS

Wang, X. *et al.* Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat. Biotechnol.* **30**, 159–164 (2012).