

MICROBIOLOGY

When bacteria experience a tremor

A trackable multiplex recombineering method in *Escherichia coli* allows for the identification of genes that affect a selectable trait.

Microbes have the potential for impressive genomic diversity. Take the microbial laboratory pet *Escherichia coli*. With a genome size of 4.6 million base pairs and just under 4,300 protein-coding genes, the sequence space for combinatorial diversity is huge, and exploring which combinations affect interesting traits is consequently challenging. As Ryan Gill from the University of Colorado puts it, “how do you search the genome for a collection of genes that do something you care about?” He came up with an answer by combining synthetic DNA synthesis, recombineering and barcoding technology in a method he named TRMR—for ‘trackable multiplex recombineering’ and pronounced tremor—to map the specific effect of each gene on a trait that can be selected for.

The plan was to target each of the ~4,000 genes in *E. coli* with an oligonucleotide cassette that either enhances expression or shuts it down. The ‘up cassettes’ introduce a strong promoter in front of the gene of interest, and the ‘down cassettes’ replace a ribosomal binding site with an inert sequence that no longer initiates translation.

The length of each of these cassettes was around 200 nucleotides, roughly twice as long as commercial providers can synthesize on high-throughput arrays. Consequently, obtaining these reagents was not as straightforward as placing an order with a company. It took the ingenuity of Joseph Warner, a postdoc on Gill’s team, to develop a way around the length restrictions.

He saw the cassettes as a module made up of three features: a targeting region that homes in on each gene, a functional region, either the artificial promoter or the ribosomal binding site, and a barcode region for tracking. Only the targeting and tracking regions are unique to each cassette, the functional regions are the same in all up or down cassettes. Warner had a company synthesize only the unique parts—which were well within the size

limits of 120 nucleotides—on an array; then he cloned the functional regions to the unique oligonucleotides, amplified them and then separated each cassette via strategically placed restriction sites.

TRMR was then ready to go. The Gill team first tested it by growing *E. coli* cells transformed with the cassettes in four different growth conditions. They characterized the colonies on DNA microarrays and ranked each allele for its contribution to the fitness in each condition. Simultaneously they sequenced the barcodes of individual colonies and found the results to agree.

Gill sees TRMR as a first screening step to find trait-associated genes, and he warns of two pitfalls: “it may fail, [owing] to either not having a good library or not having a good selection or screening strategy.” His team was careful to ascertain the cassette library complexity by sequencing barcode tags from almost 400 colonies, and in addition, they measured the concentration of each barcode in cell mixtures on DNA microarrays. To Gill’s surprise they did not need too much optimization; in their first attempt they already covered 95% of all genes.

Once identified by TRMR, genes can be subjected to a more thorough mutational analysis. A method that would lend itself to this task is multiplex automated genome engineering (MAGE), developed in George Church’s lab, that targets a subset of genes with synthetic oligonucleotides that introduce various mutations, thus allowing more detailed follow-up on the function of genes known to contribute to a specific trait.

Neither TRMR nor MAGE are restricted to *E. coli*. All one needs is an organism that is amenable to recombineering and has a sequenced genome. Gill plans to apply TRMR “to traits that industry cares about,” invoking applications such as efficient biofuel production.

Nicole Rusk

RESEARCH PAPERS

Warner, J.R. *et al.* Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat. Biotechnol.* **28**, 856–862 (2010).