

SEPTEMBER 20, 2009

Technique Identifies New Drug and Vaccine Targets in Record Time

Howard Hughes Medical Institute (HHMI) researchers have devised a new method called Tn-seq for rapidly identifying genes that are crucial for survival of bacteria – cutting the time it takes to pinpoint promising new vaccine and drug targets from years to weeks.

In an article published in the September 20, 2009, issue of the journal *Nature Methods*, HHMI investigator Andrew Camilli and colleagues Tim van Opijnen and Kip L. Bodi report that they have used their Tn-seq method to identify genes essential for the survival of *Streptococcus pneumoniae*, the bacteria that cause pneumonia and meningitis. In addition, they report the use of Tn-seq to rapidly map gene interaction networks for essential cellular processes. The research provides a better understanding of the physiology of that pathogen, and should aid in developing better vaccines and antibiotics against pathogens that kill millions worldwide, says Camilli, who is at Tufts University.

For years, researchers have tried to predict which genes are most important for a pathogen's survival. They reasoned that if they could identify and then target the most important genes, they could identify new ways to kill the pathogen. Too often, however, those efforts were thwarted by a built-in redundancy, or "genetic robustness," in the genes, proteins, and biochemical pathways present in pathogens. This robustness helps ensure that the organism can function if one gene or protein is altered or disabled.

"It can get quite complicated -- there are multiple interactions on many levels that are virtually impossible to predict. That is where our method comes in, to actually map the gene networks that will then inform future antibiotic and vaccine studies."

- Andrew Camilli

“If you can identify and interfere with the function of an essential protein or process, or alternatively, target them with a vaccine, you could destroy the pathogen in an infected patient,” says Camilli. “In the pathogenesis field we have seen a number of incremental improvements in that direction, but I think this study represents a technological leap.”

A comprehensive survey of a microbe’s complete genome would provide a better picture -- and this has been done in several cases -- but usually at great cost in terms of time and money. With Camilli’s approach, a single researcher can now scan the entire genome of a microorganism and quickly locate the genes most likely to be vulnerable to attack by new drugs or therapies.

“It can get quite complicated -- there are multiple interactions on many levels that are virtually impossible to predict,” says Camilli. “That is where our method comes in, to actually map the gene networks that will then inform future antibiotic and vaccine studies.”

Other researchers have attempted to identify essential genes by deleting each gene individually and then looking to see how loss of the gene affects the viability of the microorganism. Camilli’s group came up with a less laborious method, taking advantage of the sequences of DNA called jumping genes, or transposons. Transposons have a habit of flitting about the genome and inserting themselves at random locations, where they almost always interfere with the function of the gene into which they have inserted.

Camilli and his colleagues used a jumping gene from the Mariner family of transposons for their analysis. When Mariner hops into the genome, it lands in random locations. In one instance a transposon might insert itself at the beginning of the gene. The next time it might land at the end of the gene, or in a regulatory region nearby. The beauty of the new approach, Camilli says, is that in a large population of bacteria with single Mariner insertions, the same gene will be interrupted in many different locations. Thus the researchers can accumulate copious data that addresses the same question – what is the fitness cost of disrupting the gene.

The researchers allowed a Mariner transposon to hop into the DNA of the *S. pneumoniae* bacteria. After the bacteria with the transposon insertions grew for a set period of time, the team analyzed the DNA in the bacteria at the beginning and end of the growth period. To pinpoint exactly where each transposon had inserted in the genome, they sequenced the junctions between the transposon and the genomic DNA using a high-throughput sequencing machine. Based on the frequency with which each of these junction sequences appeared in the overall population, they could calculate the effects of each transposon-induced mutation on the growth, or “fitness,” of the pathogen. If a particular gene was important for aiding *S. pneumoniae*’s growth, insertion of a transposon in that gene would interfere with the

bacteria's survival – so that type of mutation would be decreased in frequency.

Camilli's group found that only about 16 percent of the pathogen's approximately 2,000 genes were critical. Disruption of those genes stopped or severely reduced growth. The majority of the *S. pneumoniae* genes were dispensable, presumably because they either played no role in the growth condition used or were part of a genetically robust process. A much smaller percentage of genes were labeled as advantageous (6 percent) or disadvantageous (2 percent).

The researchers then used Tn-seq to further analyze five “query” genes that regulate how *S. pneumoniae* feeds off sugars from host cells. To identify genes that work in conjunction with these five query genes, they evaluated the effects of Mariner hops in each of the pathogen's genes in the context of five genetic backgrounds, in which each of the five query genes had been deleted. They measured whether each gene's contribution to fitness changed depending on the presence or absence of each of the five query genes. This allowed them to identify genetic interactions on a genome-wide scale.

From these experiments, they identified 97 gene interactions that affect *S. pneumoniae*'s ability to efficiently utilize sugars. Among these were a number of previously uncharacterized genes encoding secreted and surface proteins implicated in clipping sugars off of host cells and transporting the products into the bacterial cell.

Camilli hopes that the information garnered from such studies could lead to more effective vaccines that target not just one protein, but several proteins that comprise a genetically robust essential process. Working in collaboration with colleagues at other institutions, he has already tested vaccine antigens based on individual proteins in a mouse model, and will soon begin to combine multiple protein antigens from the same network into a single vaccine.

“Our goal is to develop a protein-based vaccine for *S. pneumoniae* that is affordable to the developing world,” says Camilli. “Though much work remains to be done, hopefully we can develop a vaccine that targets an essential network of genes and can disable the pathogen.”