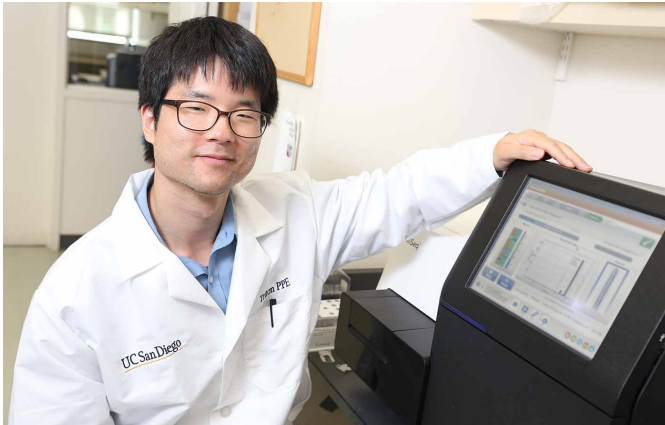


August 10, 2015 | By Liezel Labios

Bioengineers Identify the Key Genes and Functions for Sustaining Microbial Life



Laurence Yang PNAS paleome paper. Caption: Laurence Yang, a postdoctoral researcher in Palsson's Systems Biology Research Group at UC San Diego and a co-first author of the paper.

A new study led by bioengineers at the University of California, San Diego defines the core set of genes and functions that a bacterial cell needs to sustain life. The research, which answers the fundamental question of what minimum set of functions bacterial cells require to survive, could lead to new cell engineering approaches for *E. coli* and other microorganisms, the researchers said.

The findings are published online in the early edition of *Proceedings of the National Academy of Sciences* the week of August 10,

2015.

This core set of genes is “the smallest common denominator that microbes need to have to become functional,” said Bernhard Palsson, the Galetti Professor of Bioengineering at UC San Diego and corresponding author on the paper. “If the cell lacks any of the genes from this set, the cell can neither function nor survive.”

According to the researchers, these findings could open up new avenues for cell engineering applications. Consider, for example, the genetic engineering of microbes to make value-added chemicals. This engineering process is typically done by making changes to the genetic makeup of a cell, which can end up toying with the cell’s core genes and functions, resulting in a “sick” cell.

Rather than risk compromising the cell's core genes and functions, a new engineering approach could involve building the cell starting with the core set and adding on the extra desired functions, like chemical production. The *PNAS* paper presents the minimum core components that are absolutely necessary to include in the blueprints of an engineered cell.

“By defining the vital set of genes and functions that need to always be present in a cell to sustain life, we can begin to realize new ways to engineer a cell to optimize production of a desired product without sacrificing the cell's health,” said Laurence Yang, a postdoctoral researcher in Palsson's Systems Biology Research Group at UC San Diego and a co-first author of the paper.

The work, led by Palsson's research group at UC San Diego Jacobs School of Engineering, is a collaborative effort with numerical and statistical experts from Stanford University.

Defining the core set of genes and functions for cellular life

In this study, the researchers defined the core set of genes and functions as the “paleome,” referring to the ancestral genes and proteins that are at the heart of sustaining life for microbial cells.

“Other approaches have tried to define the paleome by comparing genome sequences and finding the gene portfolio that seemed to be similar in all of these sequences. This just defines the minimal genome. Our definition of the paleome takes a more comprehensive approach. It is a systems-biology-based definition that takes into account not just the minimum set of genes, but also the minimum set of functions, reactions and processes needed to build a cell,” said Palsson.

The team's approach to define the paleome is based on a genome-scale computational model for cellular growth in *E. coli*. The researchers developed this model to account for all the metabolic and gene expression processes in the cell. Using this model, the researchers simulated the growth of a well-studied strain of *E. coli* across 333 different growth conditions. In each simulated growth condition, the main nutrient source of the growth medium (carbon, nitrogen, phosphorus, or sulfur source) was varied. The team observed which set of genes was consistently expressed throughout all the different growth environments and used this set to construct the paleome. In total, the team identified 356 genes that were expressed in all of these simulations.

“Our paleome definition is representative of core function not only in the well-studied strain of *E. coli*, but also in another strain of *E. coli* and three other microorganisms. We are hoping to use this paleome as a starter kit to rapidly build a new generation of genome-scale cellular growth models for other organisms,” said Yang.

“Big Data to Knowledge”

“This study is an example of what’s called a ‘Big Data to Knowledge’ study,” added Palsson.

“We are demonstrating that we can take large data sets, integrate them together and analyze them to generate knowledge. In this case, we have used large amounts of experimental data and integrated them in the form of a computational model to arrive at our systems biology definition of the paleome.”

UC San Diego’s [Studio Ten 300](#) offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit <https://ucsdnews.ucsd.edu/media-resources/faculty-experts>.