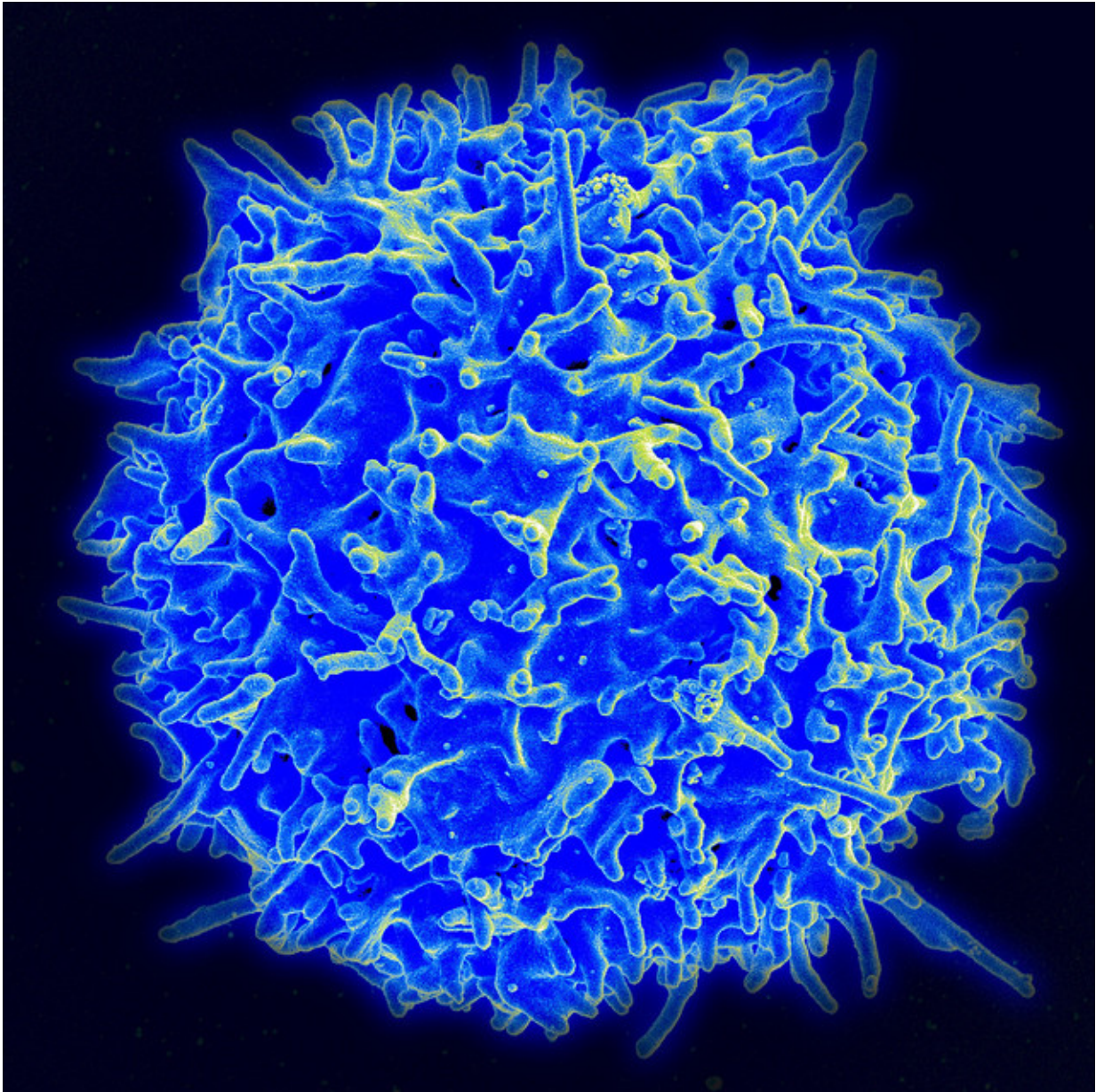# How AI found the words to kill cancer cells

December 8 2022

Scanning electron micrograph of a human T lymphocyte (also called a T cell)

from the immune system of a healthy donor. Credit: NIAID

Using new machine learning techniques, researchers at UC San Francisco (UCSF), in collaboration with a team at IBM Research, have developed a virtual molecular library of thousands of "command sentences" for cells, based on combinations of "words" that guided engineered immune cells to seek out and tirelessly kill cancer cells.

The work, published online Dec. 8, 2022, in *Science*, represents the first time such sophisticated computational approaches have been applied to a field that until now has progressed largely through ad hoc tinkering and engineering cells with existing—rather than synthesized—molecules.

The advance allows scientists to predict which elements—natural or synthesized—they should include in a cell to give it the precise behaviors required to respond effectively to complex diseases.

"This is a vital shift for the field," said Wendell Lim, Ph.D., the Byers Distinguished Professor of Cellular and Molecular Pharmacology, who directs the UCSF Cell Design Institute and led the study. "Only by having that power of prediction can we get to a place where we can rapidly design new cellular therapies that carry out the desired activities."

Meet the molecular words that make cellular command sentences

Much of therapeutic cell engineering involves choosing or creating receptors, that when added to the cell, will enable it to carry out a new function. Receptors are molecules that bridge the cell membrane to sense

the outside environment and provide the cell with instructions on how to respond to environmental conditions.

Putting the right receptor into a type of immune cell called a T cell can reprogram it to recognize and kill cancer cells. These so-called chimeric antigen receptors (CARs) have been effective against some cancers but not others.

Lim and lead author Kyle Daniels, Ph.D., a researcher in Lim's lab, focused on the part of a receptor located inside the cell, containing strings of amino acids, referred to as motifs. Each motif acts as a command "word," directing an action inside the cell. How these words are strung together into a "sentence" determines what commands the cell will execute.

Many of today's CAR-T cells are engineered with receptors instructing them to kill cancer, but also to take a break after a short time, akin to saying, "Knock out some rogue cells and then take a breather." As a result, the cancers can continue growing.

The team believed that by combining these "words" in different ways, they could generate a receptor that would enable the CAR-T cells to finish the job without taking a break. They made a library of nearly 2,400 randomly combined command sentences and tested hundreds of them in T cells to see how effective they were at striking leukemia.

What the grammar of cellular commands can reveal about treating disease

Next, Daniels partnered with computational biologist Simone Bianco, Ph.D., a research manager at IBM Almaden Research Center at the time of the study and now Director of Computational Biology at Altos Labs.

Bianco and his team, researchers Sara Capponi, Ph.D., also at IBM Almeden, and Shangying Wang, Ph.D., who was then a postdoc at IBM and is now at Altos Labs, applied novel machine learning methods to the data to generate entirely new receptor sentences that they predicted would be more effective.

"We changed some of the words of the sentence and gave it a new meaning," said Daniels. "We predictively designed T cells that killed cancer without taking a break because the new sentence told them, 'Knock those rogue tumor cells out, and keep at it.'"

Pairing machine learning with cellular engineering creates a synergistic new research paradigm.

"The whole is definitely greater than the sum of its parts," Bianco said. "It allows us to get a clearer picture of not only how to design cell therapies, but to better understand the rules underlying life itself and how living things do what they do."

Given the success of the work, added Capponi, "We will extend this approach to a diverse set of experimental data and hopefully redefine T-cell design."

The researchers believe this approach will yield cell therapies for autoimmunity, regenerative medicine and other applications. Daniels is interested in designing self-renewing stem cells to eliminate the need for donated blood.

He said the real power of the computational approach extends beyond making command sentences, to understanding the grammar of the molecular instructions.

"That is the key to making cell therapies that do exactly what we want

them to do," Daniels said. "This approach facilitates the leap from understanding the science to engineering its real-life application."

More information: Kyle G. Daniels et al, Decoding CAR T cell phenotype using combinatorial signaling motif libraries and machine learning, *Science* (2022). DOI: 10.1126/science.abq0225. www.science.org/doi/10.1126/science.abq0225

Provided by University of California, San Francisco