Antibiotic discovery effort uses AI to uncover potential new antibiotics in the global microbiome
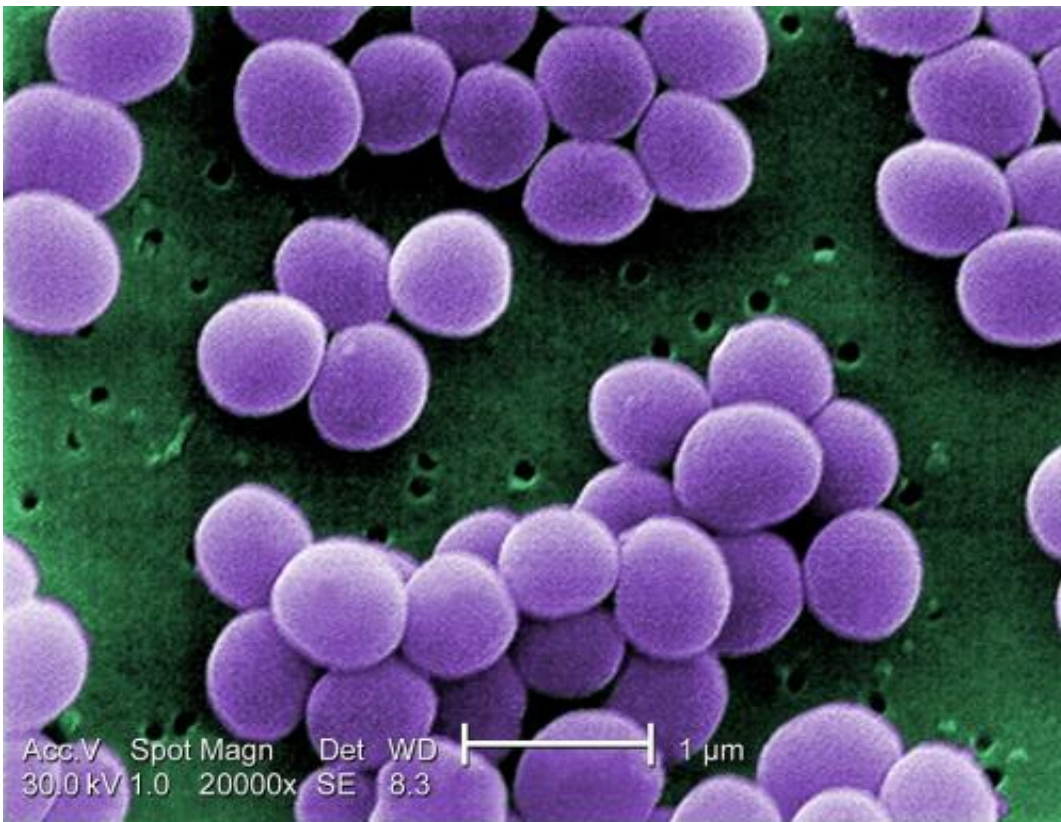
June 5 2024

Scanning electron micrograph of S. aureus; false color added. Credit: CDC

Almost a century ago, the discovery of antibiotics like penicillin revolutionized medicine by harnessing the natural bacteria-killing abilities of microbes. Now, a new study co-led by researchers at the

Perelman School of Medicine at the University of Pennsylvania suggests that natural-product antibiotic discovery is about to accelerate into a new era, powered by artificial intelligence (AI).

The study, "Discovery of antimicrobial peptides in the global microbiome with machine learning" published in *Cell*, details how the researchers used a form of AI called machine learning to search for antibiotics in a vast dataset containing the recorded genomes of tens of thousands of bacteria and other primitive organisms.

This unprecedented effort yielded nearly one million potential antibiotic compounds, with dozens showing promising activity in initial tests against disease-causing bacteria.

"AI in antibiotic discovery is now a reality and has significantly accelerated our ability to discover new candidate drugs. What once took years can now be achieved in hours using computers," said study co-senior author César de la Fuente, Ph.D., a Presidential Assistant Professor in Psychiatry, Microbiology, Chemistry, Chemical and Biomolecular Engineering, and Bioengineering.

Nature has always been a good place to look for new medicines, especially antibiotics. Bacteria, ubiquitous on our planet, have evolved numerous antibacterial defenses, often in the form of short proteins ("peptides") that can disrupt bacterial cell membranes and other critical structures.

While the discovery of penicillin and other natural-product-derived antibiotics revolutionized medicine, the growing threat of antibiotic resistance has underscored the urgent need for new antimicrobial compounds.

In recent years, de la Fuente and colleagues have pioneered AI-powered

searches for antimicrobials. They have identified preclinical candidates in the genomes of contemporary humans, extinct Neanderthals and Denisovans, wooly mammoths, and hundreds of other organisms. One of the lab's primary goals is to mine the world's biological information for useful molecules, including antibiotics.

For this new study, the research team used a machine learning platform to sift through multiple public databases containing microbial genomic data. The analysis covered 87,920 genomes from specific microbes as well as 63,410 mixes of microbial genomes—"metagenomes"—from environmental samples. This comprehensive exploration spanned diverse habitats around the planet.

This extensive exploration succeeded in identifying 863,498 candidate antimicrobial peptides, more than 90 percent of which had never been described before. To validate these findings, the researchers synthesized 100 of these peptides and tested them against 11 disease-causing bacterial strains, including antibiotic-resistant strains of E. coli and Staphylococcus aureus.

"Our initial screening revealed that 63 of these 100 candidates completely eradicated the growth of at least one of the pathogens tested, and often multiple strains," de la Fuente said. "In some cases, these molecules were effective against bacteria at very low doses."

Promising results were also observed in preclinical animal models, where some of the potent compounds successfully stopped infections. Further analysis suggested that many of these candidate molecules destroy bacteria by disrupting their outer protective membranes, effectively popping them like balloons.

The identified compounds originated from microbes living in a wide variety of habitats, including human saliva, pig guts, soil and plants,

corals, and many other terrestrial and marine organisms. This validates the researchers' broad approach to exploring the world's biological data.

Overall, the findings demonstrate the power of AI in discovering new antibiotics, providing multiple new leads for antibiotic developers, and signaling the start of a promising new era in antibiotic discovery.

The team has published their repository of putative antimicrobial sequences, which they call AMPSphere, which is open access and freely available at https://ampsphere.big-data-biology.org/

More information: Discovery of antimicrobial peptides in the global microbiome with machine learning, *Cell* (2024). DOI: 10.1016/j.cell.2024.05.013. www.cell.com/cell/fulltext/S0092-8674(24)00522-1