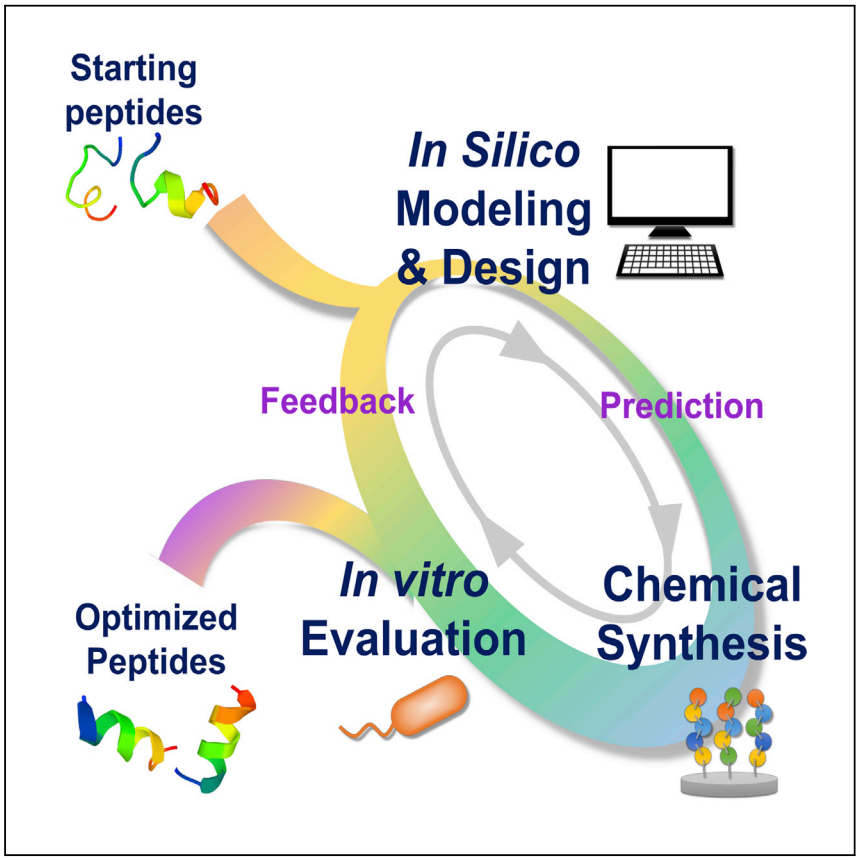
Chem

CellPress

Article

Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides



Here, we use a closed-loop discovery and optimization approach for searching the peptide sequence space. Combining an evolutionary algorithm with machine learning and *in vitro* assay allowed for rapid development of new antimicrobial peptides.

Mari Yoshida, Trevor Hinkley, Soichiro Tsuda, ..., Sabrina Galiñanes Reyes, Maria D. Castro, Leroy Cronin

lee.cronin@glasgow.ac.uk

HIGHLIGHTS

Presents a general closed-loop approach for evolution of functional molecules

Couples machine learning and artificial intelligence with *in vitro* assay

Demonstrates quick identification of a number of potent antimicrobial peptides

Selects for α-helical conformation, a common motif of potent antimicrobial peptides

Yoshida et al., Chem 4, 1–11
March 8, 2018
https://doi.org/10.1016/j.chempr.2018.01.005

Please cite this article in press as: Yoshida et al., Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides, Chem (2018), https://doi.org/10.1016/j.chempr.2018.01.005

Chem

CellPress

Article

Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides

Mari Yoshida,[1] Trevor Hinkley,[1] Soichiro Tsuda,[1] Yousef M. Abul-Haija,[1] Roy T. McBurney,[1] Vladislav Kulikov,[1] Jennifer S. Mathieson,[1] Sabrina Galiñanes Reyes,[1] Maria D. Castro,[1] and Leroy Cronin[1,2,*]

SUMMARY

We present a proof-of-concept methodology for efficiently optimizing a chemical trait by using an artificial evolutionary workflow. We demonstrate this by optimizing the efficacy of antimicrobial peptides (AMPs). In particular, we used a closed-loop approach that combines a genetic algorithm, machine learning, and *in vitro* evaluation to improve the antimicrobial activity of peptides against *Escherichia coli*. Starting with a 13-mer natural AMP, we identified 44 highly potent peptides, achieving up to a ca. 160-fold increase in antimicrobial activity within just three rounds of experiments. During these experiments, the conformation of the peptides selected was changed from a random coil to an α-helical form. This strategy not only establishes the potential of *in vitro* molecule evolution using an algorithmic genetic system but also accelerates the discovery of antimicrobial peptides and other functional molecules within a relatively small number of experiments, allowing the exploration of broad sequence and structural space.

INTRODUCTION

Natural biological polymers, such as peptides and RNAs, play a crucial role to maintain cellular functions. It is widely known that short RNA molecules interfere with gene expression within eukaryotic cells.[1] RNA molecules with enzymatic functions (ribozymes) are involved in various intracellular processes, such as RNA self-splicing.[2] Peptides are known to function as signaling molecules, such as hormones and neurotransmitters.[3] Antimicrobial peptides (AMPs) are yet another example of such natural functional biopolymers.[4] They are essential to the immune systems of all multicellular organisms as they have evolved to cope with bacterial invasion and infection. AMPs have been investigated as new antibiotic agents,[5] and more than 2,600 peptides that display antimicrobial activity from a broad range of organisms, including bacteria and mammalian cells, have been isolated.[6] Although the main mode of action causing bacterial death is disruption of the integrity of the bacterial membrane, AMPs have a wide variety of effective antimicrobial mechanisms, including the inhibition of DNA, RNA, and protein synthesis, to increase their efficacy to combat invading pathogens.[7,8] Methodologies to artificially discover such functional biopolymer sequences have been challenging because of the massive size of possible sequence space.[9] Conventionally, *in vitro* evolutionary methods, such as mRNA display and liposome display,[10,11] are used to tackle this combinatorial problem by generating billions of variants and screening them in a high-throughput fashion. However, there are limitations in these methodologies,[12]

The Bigger Picture

Biological evolution is a powerful way to produce new form and function but requires a fully functional biological organism. Here, we developed a closed-loop artificial evolution system and applied it to the exploration of antimicrobial peptides (AMPs). AMPs are a promising class of antibiotics to combat this issue because of their diverse mechanisms of action. However, discovery of antimicrobial peptides has been difficult because of a massive number of possible peptide sequences. Using this approach, we identified AMPs with improved potency. This method employs a genetic algorithm with peptide sequence as the "gene" and *in vitro* bacterial assay as "fitness." In addition, efficient predictions by machine learning further accelerated the process, showing up to a 160-fold potency increase within only three optimization rounds. This demonstrates a possibility of optimizing peptides without relying on an existing physicochemical database.