Media Monitoring Report
(Updated on 15 July 2022)

Title: HKUMed research team combines artificial intelligence and protein engineering technology to enhance gene editing efficiency of CRISPR-Cas9

Date: 29 June 2022

Participants:

Dr Alan Wong Siu-lun
Assistant Professor
School of Biomedical Sciences, HKUMed

Ms Dawn Thean Gek-Lian
Research Assistant
School of Biomedical Sciences, HKUMed

Dr Athena Chu Hoi-yee
Postdoctoral Fellow
School of Biomedical Sciences, HKUMed

Mr Fong Hoi-chun
PhD student
School of Biomedical Sciences, HKUMed

Ms Becky Chan Ka-ching
PhD student
School of Biomedical Sciences, HKUMed

Dr Zhou Peng
Postdoctoral Fellow
School of Biomedical Sciences, HKUMed

Ms Cynthia Kwok Chui-shan
Research Assistant
School of Biomedical Sciences, HKUMed

Dr Gigi Choi Ching-gee
Postdoctoral Fellow
School of Biomedical Sciences, HKUMed

Dr Joshua Ho Wing-kei
Associate Professor
School of Biomedical Sciences, HKUMed

Nature

Date: 29 June 2022
Topic: First reported case of a person getting COVID from a cat
Scientists in Thailand have established that a tabby passed SARS-CoV-2 to a veterinary surgeon — although such cases of cat-to-human transmission are probably rare.
Link: https://www.nature.com/articles/d41586-022-01792-y

First there were sneezing hamsters, now sneezing cats. A team in Thailand reports the first solid evidence of a pet cat infecting a person with SARS-CoV-2 — adding felines to the list of animals that can transmit the virus to people.

Researchers say the results are convincing. They are surprised that it has taken this long to establish that transmission can occur, given the scale of the pandemic, the virus's ability to jump between animal species, and the close contact between cats and people. "We've known this was a possibility for two years," says Angela Bosco-Lauth, an infectious-disease researcher at Colorado State University in Fort Collins.

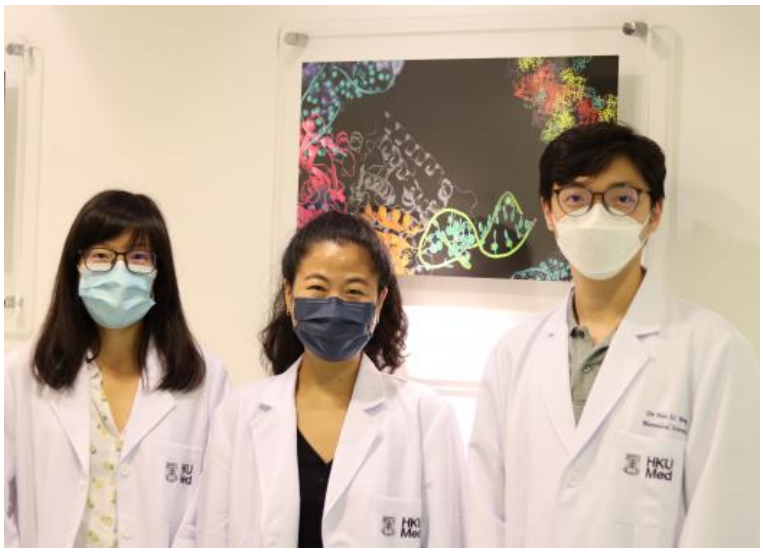
Studies early in the pandemic found that cats shed infectious virus particles and can infect other cats. And over the course of the pandemic, countries have reported SARS-CoV-2 infections in dozens of pet cats. But establishing the direction of viral spread — from cat to person or from person to cat — is tricky. The Thai study "is an interesting case report, and a great example of what good contact tracing can do", says Marion Koopmans, a virologist at the Erasmus University Medical Center in Rotterdam, the Netherlands.

The feline finding, published in Emerging Infectious Diseases1 on 6 June, came about by accident, says co-author Sarunyou Chusri, an infectious-disease researcher and physician at Prince of Songkla University in Hat Yai, southern Thailand. In August, a father and son who had tested positive for SARS-CoV-2 were transferred to an isolation ward at the university's hospital. Their ten-year-old cat was also swabbed and tested positive. While being swabbed, the cat sneezed in the face of a veterinary surgeon, who was wearing a mask and gloves but no eye protection.

Three days later, the vet developed a fever, sniffles and a cough, and later tested positive for SARS-CoV-2, but none of her close contacts developed COVID-19, suggesting that she had been infected by the cat. Genetic analysis also confirmed that the vet was infected with the same variant as the cat and its owners, and the viral genomic sequences were identical.

Mirage News

Date: 29 June 2022
Topic: HKUMed research team combines artificial intelligence and protein engineering technology to enhance gene editing efficiency
Link: https://www.miragenews.com/hkumed-research-team-combines-artificial-810081/



A research team from the LKS Faculty of Medicine, The University of Hong Kong (HKUMed) discovered more efficient CRISPR-Cas9 variants that could be useful for gene therapy applications. By establishing a new pipeline methodology that implements machine learning on high-throughput screening to accurately predict the activity of protein variants, the team expands the capacity to analyse up to 20 times more variants at once without the need for acquiring additional experimental data, which vastly accelerates the speed in protein engineering. The research team has successfully applied the pipeline in several Cas9 optimisations and engineered new Staphylococcus aureus Cas9 (SaCas9) variants with enhanced gene editing efficiency. The findings are now published in Nature Communications (link to the publication) and a patent application has been filed based on this work.

Background

Staphylococcus aureus Cas9 (SaCas9) is a great candidate for in vivo gene therapy due to its small size allowing packaging into adeno-associated viral vectors to be delivered into human cells for therapeutic applications. However, its gene editing activity could be insufficient for some specific disease loci. Further optimisations of SaCas9 are crucial in precision medicine before it can be used as a reliable tool to treat human diseases. Such optimisations consist of boosting its efficiency and precision by altering the Cas9 protein. Standard protocol for modifying the protein entails saturation mutagenesis, where the

number of possible modifications that could be introduced to the protein far exceeds the experimental screening capacity of even the state-of-art high-throughput platforms by orders of magnitudes.

In this work, the research team explored if combining machine learning with structure-guided mutagenesis library screening could enable the virtual screening of many more modifications to accurately identify the rare and better performing variants for further in-depth validations.

Research findings

The research team tested the machine learning framework on several previously published mutagenesis screens on Cas9 variants and illustrated that machine learning could robustly identify the best performing variants by using merely 5-20% of the experimentally determined data.

The Cas9 protein contains several parts, including protospacer adjacent motif (PAM)-interacting (PI) and Wedge (WED) domains to facilitate its interaction with the target DNA duplex. The research team coupled the machine learning and high-throughput screening platforms to design activity-enhanced SaCas9 protein by combining mutations in its PI and WED domains surrounding the DNA duplex bearing a (PAM). PAM is essential for Cas9 to edit the target DNA and the idea was to reduce the PAM constraint for wider genome targeting whilst securing the protein structure by reinforcing the interaction with the PAM-containing DNA duplex via the WED domain.

In the screen and subsequent validations, the researchers identified new variants, including one named KKH-SaCas9-plus, with enhanced activity by up to 33% at specific genomic loci. The subsequent protein modelling analysis revealed the new interactions created between the WED and PI domains at multiple locations within the PAM-containing DNA duplex, attributing to KKH-SaCas9-plus's enhanced efficiency.

Research significance

Structure-guided design has been dominating the field of Cas9 engineering; however, it only explores a small number of sites, amino-acid residues, and combinations. In this study, the research team showed that screening with larger scale and less experimental efforts, time and cost can be conducted using the machine learning-coupled multi-domain combinatorial mutagenesis screening approach, which led them to identify a new high-efficiency variant KKH-SaCas9-plus.

'This approach will greatly accelerate the optimisation of Cas9 proteins, which could allow genome editing to be applied in treating genetic diseases in a more efficient way,' said Dr Alan Wong Siu-lun, Assistant Professor of the School of Biomedical Sciences, HKUMed.

About the research team

This research was led by Dr Alan Wong Siu-lun, Assistant Professor of the School of Biomedical Sciences, HKUMed, as the corresponding author. Ms Dawn Thean Gek-lian, Research Assistant; Dr Athena Chu Hoi-yee, Postdoctoral Fellow, School of Biomedical Sciences, HKUMed, were co-first authors, with assistance from Mr Fong Hoi-chun, PhD student; Ms Becky Chan Ka-ching, PhD student; Dr Zhou Peng, Postdoctoral Fellow; Ms Cynthia Kwok Chui-shan, Research Assistant, and Dr Gigi Choi Ching-gee, Postdoctoral Fellow, School of Biomedical Sciences, HKUMed. Other collaborators included Dr Joshua Ho Wing-kei, Associate Professor of the School of Biomedical Sciences, HKUMed; Dr Zheng Zongli and his team from Ming Wai Lau Centre for Reparative Medicine, Karolinska Institutet, Hong Kong node.

Acknowledgement

This work was supported by the Excellent Young Scientists Fund, National Natural Science Foundation of China (32022089), the Hong Kong Research Grants Council (17104619), and the Centre for Oncology and Immunology Limited under the Health@InnoHK Programme launched by the Innovation and Technology Commission, the Government of Hong Kong Special Administrative Region (HKSAR). This work was also in part supported by the Ming Wai Lau Centre for Reparative Medicine Associate Member Programme and the AIR@InnoHK launched by the Innovation and Technology Commission, the Government of HKSAR.

Sing Tao Daily Online

日期：2022 年 6 月 29 日
主題：港大研結合人工智能與蛋白質工程技術 提升基因編輯效率
連結：https://std.stheadline.com/realtime/article/1850567/%E5%8D%B3%E6%99%82-%E6%B8%AF%E8%81%9E-%E6%B8%AF%E5%A4%A7%E7%A0%94%E7%B5%90%E5%90%88%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD%E8%88%87%E8%9B%8B%E7%99%BD%E8%B3%AA%E5%B7%A5%E7%A8%8B%E6%8A%80%E8%A1%93-%E6%8F%90%E5%8D%87%E5%9F%BA%E5%9B%A0%E7%B7%A8%E8%BC%AF%E6%95%88%E7%8E%87

香港大學李嘉誠醫學院研究團隊發現更有效的 CRISPR-Cas9 變體，可應用於基因治療。是項研究將人工智能中的「機器學習」應用於大型蛋白質篩選，從有限的蛋白變體實驗數據，拓展出更龐大的虛擬數據作分析，令變體數據可增至原有的 20 倍，加快篩選的速度。團隊將此方法成功地應用於改良多個 Cas9 蛋白，並設計出具有增強基因編輯效率的新金黃葡萄球菌 Cas9（SaCas9）的變體。相關研究成果現已在國際科學期刊《自然-通訊》，並就此提交專利申請。

是項研究探索如何結合「機器學習」方法去進行多點突變的蛋白篩選，將結構導向的突變數據庫中的實證數據，再以「機器學習」方法進行虛擬篩選，以準確識別稀有和性能更好的變體，以作進一步的深入驗證。團隊亦根據早前發表的 Cas9 變體篩選之數據作為測試機器學習的框架，印證機器學習只需憑藉 5%至 20%的實驗數據，即可識別性能最佳的變體。

挑選定點突變的位置均集中於 Cas9 蛋白內的間隔序列前體臨近基序（PAM）相互作用（PI）結構域和 WED 結構域，因為這兩個結構域靠近目標脫氧核醣核酸（DNA），及圍繞帶有 PAM 的 DNA 雙鏈體的位置。研究團隊將「機器學習」融合至高通量篩選平台，結合 PI 和 WED 結構域中的多點突變以設計活性更強的 SaCas9 蛋白。PAM 對於 Cas9 的可編輯目標相當重要，透過減弱 PAM 與 DNA 之間的相互作用，從而減少 PAM 所帶來的編輯限制，

這樣可以讓 Cas9 編輯更廣泛的基因目標，為了彌補與 DNA 之間被減弱的相互作用，需要同時於 WED 結構域加強與 DNA 的相互作用，以增強 Cas9 的編輯能力。

在篩選和隨後的驗證中，研究人員辨別新的變體，包括當中名為 KKH-SaCas9-plus 的變體，它在特定基因組位點的活性增強高達 33%。而蛋白質建模分析亦預測新改良的變體，有機會增加在 WED 和 PI 結構域與帶有 PAM DNA 雙鏈之間的新相互作用，這亦可解釋多點突變如何增強新變體的編輯效率。

結構導向設計一直主導著 Cas9 改良工程的領域，然而只探索少數位點、氨基酸殘基突變和多位點的組合突變。是項研究發現在結合「機器學習」到多點組合突變的篩選研究中，有助將實驗數據產出最大化，降低實驗的篩選時間和成本，並從更多的變體中尋找到更高效的變體 KKH-SaCas9-plus。領導是項研究的港大醫學院生物醫學學院助理教授黃兆麟表示，這方法將加速 Cas9 蛋白的改良，使基因組編輯技術得以更有效地應用於治療遺傳疾病。

Headline Daily Online

日期：2022 年 6 月 29 日
主題：港大研結合人工智能與蛋白質工程技術 提升基因編輯效率
連結：https://hd.stheadline.com/news/realtime/hk/2350567/%E5%8D%B3%E6%99%82-%E6%B8%AF%E8%81%9E-%E6%B8%AF%E5%A4%A7%E7%A0%94%E7%B5%90%E5%90%88%E4%BA%BA%E5%B7%A5%E6%99%BA%E8%83%BD%E8%88%87%E8%9B%8B%E7%99%BD%E8%B3%AA%E5%B7%A5%E7%A8%8B%E6%8A%80%E8%A1%93-%E6%8F%90%E5%8D%87%E5%9F%BA%E5%9B%A0%E7%B7%A8%E8%BC%AF%E6%95%88%E7%8E%87

香港大學李嘉誠醫學院研究團隊發現更有效的 CRISPR-Cas9 變體，可應用於基因治療。是項研究將人工智能中的「機器學習」應用於大型蛋白質篩選，從有限的蛋白變體實驗數據，拓展出更龐大的虛擬數據作分析，令變體數據可增至原有的 20 倍，加快篩選的速度。團隊將此方法成功地應用於改良多個 Cas9 蛋白，並設計出具有增強基因編輯效率的新金黃葡萄球菌 Cas9（SaCas9）的變體。相關研究成果現已在國際科學期刊《自然-通訊》，並就此提交專利申請。

是項研究探索如何結合「機器學習」方法去進行多點突變的蛋白篩選，將結構導向的突變數據庫中的實證數據，再以「機器學習」方法進行虛擬篩選，以準確識別稀有和性能更好的變體，以作進一步的深入驗證。團隊亦根據早前發表的 Cas9 變體篩選之數據作為測試機器學習的框架，印證機器學習只需憑藉 5%至 20%的實驗數據，即可識別性能最佳的變體。

挑選定點突變的位置均集中於 Cas9 蛋白內的間隔序列前體臨近基序（PAM）相互作用（PI）結構域和 WED 結構域，因為這兩個結構域靠近目標脫氧核醣核酸（DNA），及圍繞帶有 PAM 的 DNA 雙鏈體的位置。研究團隊將「機器學習」融合至高通量篩選平台，結合 PI 和 WED 結構域中的多點突變以設計活性更強的 SaCas9 蛋白。PAM 對於 Cas9 的可編輯目

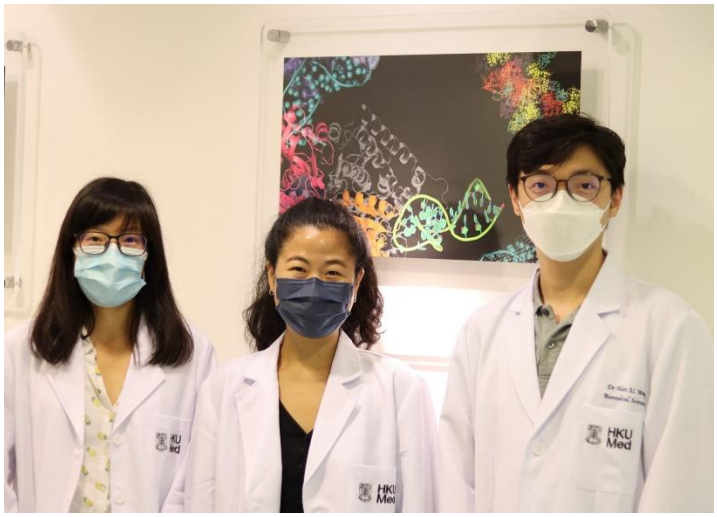
標相當重要，透過減弱 PAM 與 DNA 之間的相互作用，從而減少 PAM 所帶來的編輯限制，這樣可以讓 Cas9 編輯更廣泛的基因目標，為了彌補與 DNA 之間被減弱的相互作用，需要同時於 WED 結構域加強與 DNA 的相互作用，以增強 Cas9 的編輯能力。

在篩選和隨後的驗證中，研究人員辨別新的變體，包括當中名為 KKH-SaCas9-plus 的變體，它在特定基因組位點的活性增強高達 33%。而蛋白質建模分析亦預測新改良的變體，有機會增加在 WED 和 PI 結構域與帶有 PAM DNA 雙鏈之間的新相互作用，這亦可解釋多點突變如何增強新變體的編輯效率。

結構導向設計一直主導著 Cas9 改良工程的領域，然而只探索少數位點、氨基酸殘基突變和多位點的組合突變。是項研究發現在結合「機器學習」到多點組合突變的篩選研究中，有助將實驗數據產出最大化，降低實驗的篩選時間和成本，並從更多的變體中尋找到更高效的變體 KKH-SaCas9-plus。領導是項研究的港大醫學院生物醫學學院助理教授黃兆麟表示，這方法將加速 Cas9 蛋白的改良，使基因組編輯技術得以更有效地應用於治療遺傳疾病。

HK01

日期：2022 年 6 月 29 日
主題：港大研結合 AI 及蛋白質工程技術　提升基因編輯效率　治療遺傳疾病
連結：https://www.hk01.com/sns/article/786979



港大醫學院研究團隊結合人工智能與蛋白質工程技術，發現更有效的 CRISPR-Cas9 變體，可應用於基因治療。研究拓展出更龐大的虛擬數據作分析，令變體數據可大增至原有的 20 倍，並加快篩選的速度。研究團隊更將方法應用於改良多個 Cas9 蛋白，並設計出具有增強基因編輯效率的新金黃葡萄球菌 Cas9（SaCas9）的變體。

由港大醫學院生物醫學學院系助理教授黃兆麟領導的研究團隊，將「機器學習」融合至高通量篩選平台，結合 PI 和 WED 結構域中的多點突變以設計活性更強的 SaCas9 蛋白。PAM 對於 Cas9 的可編輯目標相當重要，透過減弱 PAM 與 DNA 之間的相互作用，從而減少 PAM 所帶來的編輯限制，可讓 Cas9 編輯更廣泛的基因目標，為了彌補與 DNA 之間被減弱的相互作用，需要同時於 WED 結構域加強與 DNA 的相互作用，以增強 Cas9 的編輯能力。

蛋白質建模分析亦預測新改良變體

在篩選和隨後的驗證中，團隊辨別了新的變體，包括當中名為 KKH-SaCas9-plus 的變體，它在特定基因組位點的活性增強了高達 33%。而蛋白質建模分析亦預測新改良的變體，有機會增加在 WED 和 PI 結構域與帶有 PAM DNA 雙鏈之間的新相互作用。

結構導向設計一直主導著 Cas9 改良工程的領域，然而，它只探索了少數位點、氨基酸殘基突變和多位點的組合突變。研究發現在結合「機器學習」到多點組合突變的篩選研究中，有助將實驗數據產出最大化，大大降低實驗的篩選時間和成本，並從更多的變體中尋找到更高效的變體 KKH-SaCas9-plus。

已就研究成果提交了專利申請

黃兆麟表示，此方法將大大加快 Cas9 蛋白的改良，使基因組編輯技術得以更有效地應用於治療遺傳疾病。相關研究成果現已在國際科學期刊《自然—通訊》，並就此提交了專利申請。

TOPick.hket.com

日期：2022 年 6 月 29 日
主題：【醫療科技】港大醫學院團隊結合 AI 及蛋白質工程技術　提升基因編輯效率可更有效治療遺傳疾病
連結：https://topick.hket.com/article/3288925

港大醫學院研究團隊結合人工智能與蛋白質工程技術，發現更有效的 CRISPR-Cas9 變體，可應用於基因治療。相關研究成果已在國際科學期刊《自然—通訊》刊登，並已就此提交專利申請。

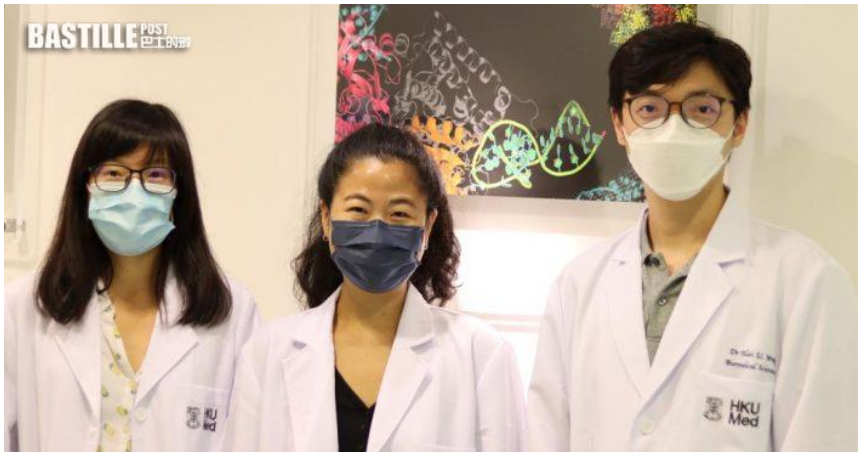
是項研究將人工智能中的「機器學習」應用於大型蛋白質篩選，從有限的蛋白變體實驗數據，拓展出更龐大的虛擬數據作分析，令變體數據可大增至原有的 20 倍，大大加快篩選的速度。研究團隊將此方法成功應用於改良多個 Cas9 蛋白，並設計出具有增強基因編輯效率的新金黃葡萄球菌 Cas9（SaCas9）的變體。

結構導向設計主導着 Cas9 改良工程的領域，然而，它只探索了少數位點、氨基酸殘基突變和多位點的組合突變。是項研究發現在結合「機器學習」到多點組合突變的篩選研究中，有助將實驗數據最大化，大大降低實驗的篩選時間和成本，並從更多變體中尋找到更高效的變體 KKH-SaCas9-plus。

領導研究的港大醫學院生物醫學學院助理教授黃兆麟表示，這方法將大大加速 Cas9 蛋白的改良，使基因組編輯技術得以更有效地應用於治療遺傳疾病。

Bastille Post

日期：2022 年 6 月 29 日
主題：港大研究結合人工智能與蛋白質工程技術 增強基因編輯效率
連結：
https://www.bastillepost.com/hongkong/article/10934683-%e6%b8%af%e5%a4%a7%e7%a0%94%e7%b5%90%e5%90%88%e4%ba%ba%e5%b7%a5%e6%99%ba%e8%83%bd%e8%88%87%e8%9b%8b%e7%99%bd%e8%b3%aa%e5%b7%a5%e7%a8%8b%e6%8a%80%e8%a1%93-%e6%8f%90%e5%8d%87%e5%9f%ba%e5%9b%a0%e7%b7%a8



香港大學李嘉誠醫學院研究團隊發現更有效的 CRISPR-Cas9 變體，可應用於基因治療。是項研究將人工智能中的「機器學習」應用於大型蛋白質篩選，從有限的蛋白變體實驗數據，拓展出更龐大的虛擬數據作分析，令變體數據可增至原有的 20 倍，加快篩選的速度。團隊將此方法成功地應用於改良多個 Cas9 蛋白，並設計出具有增強基因編輯效率的新金黃葡萄球菌 Cas9（SaCas9）的變體。相關研究成果現已在國際科學期刊《自然-通訊》，並就此提交專利申請。

是項研究探索如何結合「機器學習」方法去進行多點突變的蛋白篩選，將結構導向的突變數據庫中的實證數據，再以「機器學習」方法進行虛擬篩選，以準確識別稀有和性能更好的變體，以作進一步的深入驗證。團隊亦根據早前發表的 Cas9 變體篩選之數據作為測試機器學習的框架，印證機器學習只需憑藉 5%至 20%的實驗數據，即可識別性能最佳的變體。

挑選定點突變的位置均集中於 Cas9 蛋白內的間隔序列前體臨近基序（PAM）相互作用（PI）結構域和 WED 結構域，因為這兩個結構域靠近目標脫氧核醣核酸（DNA），及圍繞帶有 PAM 的 DNA 雙鏈體的位置。研究團隊將「機器學習」融合至高通量篩選平台，結合 PI 和 WED 結構域中的多點突變以設計活性更強的 SaCas9 蛋白。PAM 對於 Cas9 的可編輯目標相當重要，透過減弱 PAM 與 DNA 之間的相互作用，從而減少 PAM 所帶來的編輯限制，

這樣可以讓 Cas9 編輯更廣泛的基因目標，為了彌補與 DNA 之間被減弱的相互作用，需要同時於 WED 結構域加強與 DNA 的相互作用，以增強 Cas9 的編輯能力。

在篩選和隨後的驗證中，研究人員辨別新的變體，包括當中名為 KKH-SaCas9-plus 的變體，它在特定基因組位點的活性增強高達 33%。而蛋白質建模分析亦預測新改良的變體，有機會增加在 WED 和 PI 結構域與帶有 PAM DNA 雙鏈之間的新相互作用，這亦可解釋多點突變如何增強新變體的編輯效率。

結構導向設計一直主導著 Cas9 改良工程的領域，然而只探索少數位點、氨基酸殘基突變和多位點的組合突變。是項研究發現在結合「機器學習」到多點組合突變的篩選研究中，有助將實驗數據產出最大化，降低實驗的篩選時間和成本，並從更多的變體中尋找到更高效的變體 KKH-SaCas9-plus。領導是項研究的港大醫學院生物醫學學院助理教授黃兆麟表示，這方法將加速 Cas9 蛋白的改良，使基因組編輯技術得以更有效地應用於治療遺傳疾病。

News-medical.net
Date: Jun 30, 2022
Topic: HKUMed researchers discover more efficient CRISPR-Cas9 variants for gene therapy applications
Link:
https://www.news-medical.net/news/20220630/HKUMed-researchers-discover-more-efficient-CRISPR-Cas9-variants-for-gene-therapy-applications.aspx

A research team from the LKS Faculty of Medicine, The University of Hong Kong (HKUMed) discovered more efficient CRISPR-Cas9 variants that could be useful for gene therapy applications. By establishing a new pipeline methodology that implements machine learning on high-throughput screening to accurately predict the activity of protein variants, the team expands the capacity to analyze up to 20 times more variants at once without the need for acquiring additional experimental data, which vastly accelerates the speed in protein engineering. The research team has successfully applied the pipeline in several Cas9 optimizations and engineered newStaphylococcus aureusCas9 (SaCas9) variants with enhanced gene editing efficiency. The findings are now published in Nature Communications and a patent application has been filed based on this work.

Background
Staphylococcus aureusCas9 (SaCas9) is a great candidate for in vivo gene therapy due to its small size allowing packaging into adeno-associated viral vectors to be delivered into human cells for therapeutic applications. However, its gene-editing activity could be insufficient for some specific disease loci. Further optimizations of SaCas9 are crucial in precision medicine before it can be used as a reliable tool to treat human diseases. Such optimizations consist of boosting its efficiency and precision by altering the Cas9 protein. Standard protocol for modifying the protein entails saturation mutagenesis, where the number of possible modifications that could be introduced to the protein far exceeds the experimental screening capacity of even the state-of-art high-throughput platforms by orders of magnitudes.

In this work, the research team explored if combining machine learning with structure-guided mutagenesis library screening could enable the virtual screening of many more modifications to accurately identify the rare and better performing variants for further in-depth validations.

Research findings
The research team tested the machine learning framework on several previously published mutagenesis screens on Cas9 variants and illustrated that machine learning

could robustly identify the best performing variants by using merely 5-20% of the experimentally determined data.

The Cas9 protein contains several parts, including protospacer adjacent motif (PAM)-interacting (PI) and Wedge (WED) domains to facilitate its interaction with the target DNA duplex. The research team coupled the machine learning and high-throughput screening platforms to design activity-enhanced SaCas9 protein by combining mutations in its PI and WED domains surrounding the DNA duplex bearing a (PAM). PAM is essential for Cas9 to edit the target DNA and the idea was to reduce the PAM constraint for wider genome targeting whilst securing the protein structure by reinforcing the interaction with the PAM-containing DNA duplex via the WED domain.

In the screen and subsequent validations, the researchers identified new variants, including one named KKH-SaCas9-plus, with enhanced activity by up to 33% at specific genomic loci. The subsequent protein modeling analysis revealed the new interactions created between the WED and PI domains at multiple locations within the PAM-containing DNA duplex, attributing to KKH-SaCas9-plus's enhanced efficiency.

Research significance

Structure-guided design has been dominating the field of Cas9 engineering; however, it only explores a small number of sites, amino-acid residues, and combinations. In this study, the research team showed that screening with larger scale and less experimental efforts, time and cost can be conducted using the machine learning-coupled multi-domain combinatorial mutagenesis screening approach, which led them to identify a new high-efficiency variant KKH-SaCas9-plus.

'This approach will greatly accelerate the optimization of Cas9 proteins, which could allow genome editing to be applied in treating genetic diseases in a more efficient way,' said Dr Alan Wong Siu-lun, Assistant Professor of the School of Biomedical Sciences, HKUMed.

Newsfounded.com
Date: Jun 30, 2022
Topic: HKUMed researchers discovered better variants of CRISPR-Cas9 for gene therapy applications
Link:
https://newsfounded.com/singapore/hkumed-researchers-discovered-better-variants-of-crispr-cas9-for-gene-therapy-applications/

A research team from LKS Faculty of Medicine, The University of Hong Kong (HKUMed) has discovered better variants of CRISPR-Cas9 that could be useful for gene therapy applications. By establishing a new pipeline methodology that implements machine learning in high-throughput screening to accurately predict the activity of protein variants, the team expands the capacity to test up to 20 times more variants simultaneously. does not require additional experimental data, which greatly accelerates the speed of protein engineering. The research team successfully applied the pipeline to several Cas9 optimizations and created a new oneStaphylococcus aureusCas9 (SaCas9) variants with enhanced gene editing efficiency. The findings are now published in Communication in Nature and a patent application was filed based on this work.

Background
Staphylococcus aureusCas9 (SaCas9) is an excellent candidate for in vivo gene therapy due to its small size that allows the packaging of adeno -associated viral vectors to be delivered to human cells for therapeutic applications. However, its gene editing activity may not be sufficient for some specific disease areas. Further optimizations of SaCas9 are essential to precision medicine before it can be used as a reliable tool to treat human diseases. Such optimizations consist of boosting its efficiency and accuracy by modifying the Cas9 protein. The standard protocol for protein modification requires saturation mutagenesis, in which the number of possible changes that can be introduced into the protein far exceeds the experimental screening capacity of even the state-of-the-art high-throughput platform by orders of magnitude.

In this work, the research team explored whether combining machine learning with structure-guided mutagenesis library screening could enable virtual screening of many more changes to accurately identify rare and better performing variants for more in -depth validation.

Research findings
The research team tested the machine learning framework on several previously published mutagenesis screens on Cas9 variants and described that machine learning

could robustly determine the best performing variants by using only 5-20% of the data determined by the experiment.

The Cas9 protein contains several components, including protospacer adjacent motif (PAM) -interacting (PI) and Wedge (WED) domains to facilitate its interaction with the target DNA duplex. The research team combined machine learning and high-throughput screening platforms to design an activity-enhanced SaCas9 protein by combining mutations in its PI and WED domains surrounding the DNA duplex carrying (PAM). PAM is important for Cas9 to edit target DNA and the idea is to reduce the PAM constraint for broader genome targeting while securing the protein structure by strengthening interaction with PAM containing of DNA duplex through the WED domain.

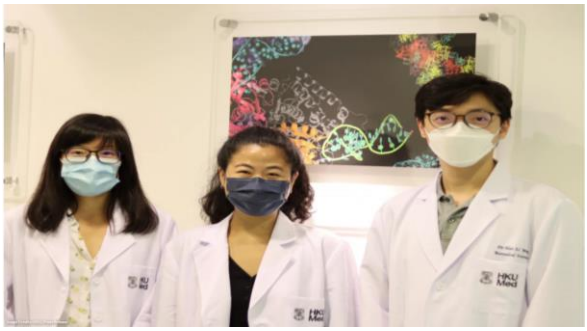
On screen and subsequent validations, the researchers identified new variants, including one named KKH-SaCas9-plus, with enhanced activity of up to 33% at specific genomic loci. Subsequent protein modeling analysis revealed new interactions created between WED and PI domains at multiple locations within PAM containing duplex DNA, associated with enhanced KKH-SaCas9 efficiency. -plus.

Importance of research

Structure-guided design is dominant in the field of Cas9 engineering; however, it only explored a small number of sites, amino-acid residues, and combinations. In this study, the research team demonstrated that screening with larger scale and less effort, time, and cost could be performed using the machine learning-coupled multi-domain combinatorial mutagenesis screening approach, which led to them to define a new high-efficiency variant KKH -SaCas9-plus.

‘This approach will greatly accelerate the optimization of Cas9 proteins, which could allow genome editing to be applied to treating genetic diseases in a better way,’ said Dr Alan Wong Siu-lun, Assistant Professor of the School of Biomedical Sciences, HKUMed.

Opengovasia.com
Date: Jun 30, 2022
Topic: HKUMed Research Team Develops Novel Gene Editing Variants
Link:
https://opengovasia.com/hkumed-research-team-develops-novel-gene-editing-variants/



A research team from the LKS Faculty of Medicine at The University of Hong Kong (HKUMed) has developed more efficient CRISPR-Cas9 variants that could be useful for gene therapy applications. By establishing a new pipeline methodology that implements machine learning on high-throughput screening to accurately predict the activity of protein variants, the team has expanded the capacity to analyse up to 20 times more variants at once without needing to acquire additional experimental data, which vastly accelerates the speed in protein engineering.

The pipeline has been successfully applied in several Cas9 optimisations and engineered new Staphylococcus aureus Cas9 (SaCas9) variants with enhanced gene editing efficiency. The findings are now published in Nature Communications and a patent application has been filed based on this work.

Staphylococcus aureus Cas9 (SaCas9) is an ideal candidate for in vivo gene therapy owing to its small size that allows packaging into adeno-associated viral vectors to be delivered into human cells for therapeutic applications. However, its gene-editing activity could be insufficient for some specific disease loci.

Before it can be used as a reliable tool for the treatment of human diseases, further optimisations of SaCas9 are vital within precision medicine. These optimisations must comprise the boosting of its efficiency and precision by altering the Cas9 protein.

The standard protocol for modifying the protein involves saturation mutagenesis, where the number of possible modifications that could be introduced to the protein far exceeds the experimental screening capacity of even the state-of-art high-throughput platforms by

order of magnitude.

In their work, the team explored whether combining machine learning with structure-guided mutagenesis library screening could enable the virtual screening of many more modifications to accurately identify the rare and better-performing variants for further in-depth validations.

The machine learning framework was tested on several previously published mutagenesis screens on Cas9 variants and the team was able to show that machine learning could robustly identify the best performing variants by using merely 5-20% of the experimentally determined data.

The Cas9 protein contains several parts, including protospacer adjacent motif (PAM)-interacting (PI) and Wedge (WED) domains to facilitate its interaction with the target DNA duplex. The research team married the machine learning and high-throughput screening platforms to design activity-enhanced SaCas9 protein by combining mutations in its PI and WED domains surrounding the DNA duplex bearing a (PAM). PAM is crucial for Cas9 to edit the target DNA and the aim was to reduce the PAM constraint for wider genome targeting whilst securing the protein structure by reinforcing the interaction with the PAM-containing DNA duplex via the WED domain.

In the screen and subsequent validations, the researchers identified new variants, including one named KKH-SaCas9-plus, with enhanced activity by up to 33% at specific genomic loci. The subsequent protein modelling analysis revealed the new interactions created between the WED and PI domains at multiple locations within the PAM-containing DNA duplex, attributing to KKH-SaCas9-plus's enhanced efficiency.

Until recently, structure-guided design has dominated the field of Cas9 engineering. However, it only explores a small number of sites, amino-acid residues, and combinations. In this study, the research team was able to illustrate that screening with a larger scale and less experimental efforts, time and cost can be conducted using the machine learning-coupled multi-domain combinatorial mutagenesis screening approach, which led them to identify a new high-efficiency variant KKH-SaCas9-plus.

The Assistant Professor of the School of Biomedical Sciences, HKUMed stated that this approach will greatly accelerate the optimisation of Cas9 proteins, which could allow genome editing to be applied in treating genetic diseases more efficiently.