

Harnessing Artificial Intelligence to Enhance Microbial Cell Factories in Microalgae and Cyanobacteria Part 1: The Potential of Machine Learning

การใช้ปัญญาประดิษฐ์เพื่อเสริมศักยภาพโรงงานชีวภาพในสาหร่ายขนาดเล็กและไซยาโนแบคทีเรีย ตอนที่ 1 : ศักยภาพของ Machine Learning

Cyanobacteria and microalgae are photosynthetic microorganisms found in aquatic environments. Their ability to capture carbon dioxide from industrial emissions, assist in wastewater treatment, and produce bio-based substances positions them as valuable tools for sustainable production and environmental protection.

ไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กเป็นจุลินทรีย์อาศัยในน้ำที่สังเคราะห์แสงได้ ด้วยความสามารถในการใช้ก๊าซคาร์บอนไดออกไซด์จากโรงงานอุตสาหกรรม บำบัดน้ำเสีย และผลิตสารชีวภาพที่มีประโยชน์ทำให้ไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กเป็นโซลูชันที่สำคัญสำหรับการผลิตอย่างยั่งยืนและการรักษาสิ่งแวดล้อม

In recent years, researchers have sought to transform these organisms into microbial cell factories capable of generating high-value compounds. For instance, in *Synechococcus elongatus* PCC 7942, scientists successfully increased the production of 2,3-butanediol — an important industrial chemical — to 12.6 grams per liter. Meanwhile, in *Nannochloropsis gaditana*, modifying a single transcriptional regulator doubled the organism's lipid yield.

ในช่วงไม่กี่ปีที่ผ่านมา นักวิทยาศาสตร์ได้ทำการวิจัยเพื่อพัฒนาไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กให้กลายเป็น "โรงงานชีวภาพ" (Microbial Cell Factory) ที่สามารถผลิตสารชีวภาพมูลค่าสูง เช่น ในแบคทีเรีย *Synechococcus elongatus* PCC 7942 นักวิจัยสามารถเพิ่มการผลิต 2,3-บิวทานไดออล (สารที่ใช้ในอุตสาหกรรม) ได้ถึง 12.6 กรัมต่อลิตร และในสาหร่าย *Nannochloropsis gaditana* พบว่าการดัดแปลงตัวควบคุมการถอดรหัสพันธุกรรมเพียงตัวเดียวสามารถเพิ่มปริมาณไขมันได้เป็นสองเท่า

Despite these successes, enhancing the cell factory performance of cyanobacteria and microalgae remains challenging due to limited genetic engineering tools, incomplete knowledge of metabolic regulatory networks, and relatively low production efficiency. This is where machine learning (ML) offers game-changing possibilities. By enabling computers to learn from data and make predictions, ML can address several bottlenecks, from simple regression analysis to sophisticated deep learning models and neural network.

Researchers are now applying ML to accelerate improvements in four key areas:

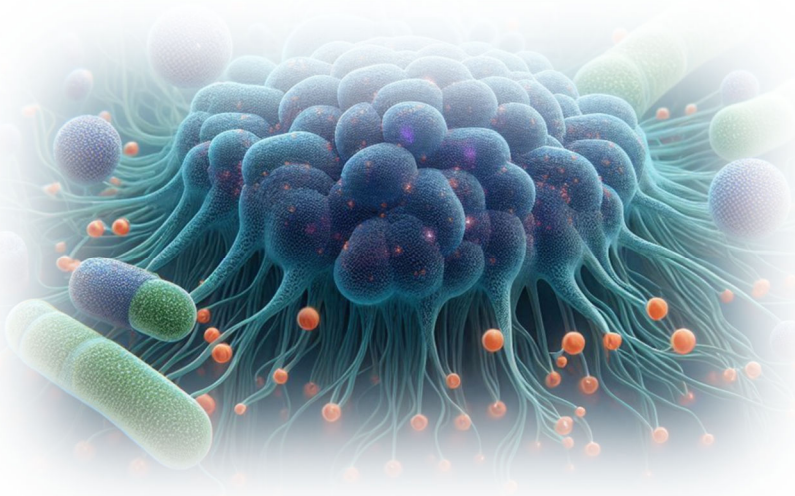
แม้จะมีศักยภาพมหาศาล แต่การใช้ไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กเป็นโรงงานชีวภาพยังมีความท้าทายหลายประการ เช่น เครื่องมือด้านพันธุวิศวกรรมที่ยังมีจำกัด ความเข้าใจเกี่ยวกับเครือข่ายการควบคุมเมแทบอลิซึมที่ยังไม่สมบูรณ์ และประสิทธิภาพการผลิตที่ยังต่ำอยู่ Machine Learning (ML) เป็นเทคโนโลยีที่ช่วยให้คอมพิวเตอร์สามารถเรียนรู้จากข้อมูลและพัฒนาความสามารถในการทำนายหรือตัดสินใจ สามารถนำมาใช้แก้ปัญหาค่าความท้าทายในการพัฒนาและประยุกต์ใช้ไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กได้หลากหลาย ตั้งแต่การวิเคราะห์การถดถอย (regression) ไปจนถึงแบบจำลองเชิงซ้อนอย่างโครงข่ายประสาทเทียม (neural network) และการเรียนรู้เชิงลึก (deep learning)

นักวิจัยได้นำ ML มาประยุกต์ใช้ในการพัฒนาไซยาโนแบคทีเรียและสาหร่ายขนาดเล็ก ดังนี้:





V. 002-2568



1. Enhancing Genetic Data Analysis

Genomic information on cyanobacteria and microalgae is often incomplete. ML aids in identifying critical genes and regulatory elements:

- In *Synechocystis* sp. PCC 6803, Support Vector Machine (SVM) and Lasso algorithms identified seven key genes involved in salinity, iron, and light stress responses.
- A machine learning model called mSRFR (microalgae SMOTE Random Forest Relief) was developed to detect non-coding RNAs (Ribonucleic Acids) — essential for gene expression control — in microalgae with up to 97% accuracy.

1. การวิเคราะห์ข้อมูลพันธุกรรมให้สมบูรณ์ยิ่งขึ้น

ปัจจุบันข้อมูลพันธุกรรมของไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กโดยส่วนใหญ่ยังขาดความสมบูรณ์ ML สามารถช่วยในการวิเคราะห์และการคาดการณ์การแสดงผล:

- ในแบคทีเรีย *Synechocystis* sp. PCC 6803 นักวิจัยใช้อัลกอริทึม Support Vector Machine และ Lasso เพื่อค้นหายีนโดยสามารถระบุยีนสำคัญ 7 ตัวที่มีบทบาทในการตอบสนองต่อความเครียดจากสิ่งแวดล้อม เช่น ความเค็ม ธาตุเหล็ก และแสง
- นักวิจัยได้พัฒนาโมเดล mSRFR (microalgae SMOTE Random Forest Relief) เพื่อระบุ RNA ที่ไม่ได้ถูกแปลรหัสเป็นโปรตีน (non-coding RNA) ในสาหร่ายขนาดเล็ก ซึ่งเป็นส่วนสำคัญในการควบคุมการแสดงผลของยีน โดยโมเดลนี้มีความแม่นยำสูงถึง 97%

2. Improving Genetic Engineering Tools

ML is advancing the development of genetic modification tools:

- Deep Neural Networks (DNN) have been used to design and predict promoter strength in *Synechocystis* PCC 6803, resulting in highly effective synthetic promoters.

2. การพัฒนาเครื่องมือตัดแต่งพันธุกรรม

การตัดแต่งพันธุกรรมในไซยาโนแบคทีเรียและสาหร่ายขนาดเล็กยังมีข้อจำกัด ML ช่วยพัฒนาเครื่องมือเหล่านี้ให้มีประสิทธิภาพมากขึ้น

- Deep Neural Network (DNN) ได้ถูกนำมาใช้ออกแบบและทำนายความแรงของโปรโมเตอร์ (ตัวควบคุมการแสดงผลของยีน) ในแบคทีเรีย *Synechocystis* PCC 6803 ทำให้สามารถสร้างโปรโมเตอร์สังเคราะห์ที่มีประสิทธิภาพสูง





V. 002-2568

3. Understanding Metabolic Networks

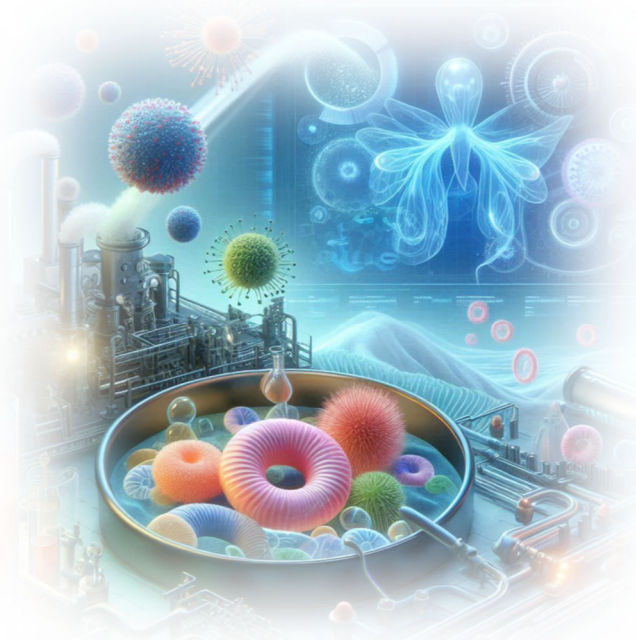
Improving production efficiency requires detailed insight into cellular metabolism:

- ARCTICA, an integrated platform combining Flux Balance Analysis with ML, was developed to identify key metabolic reactions in *Synechocystis* sp. PCC 6803.
- In *Dunaliella*, transcriptomic data combined with supervised machine learning revealed core components of its salt stress response system.

3. สร้างความเข้าใจในเครือข่ายเมแทบอลิซึม

เพื่อให้สามารถเพิ่มการผลิตสารที่ต้องการ นักวิจัยต้องเข้าใจกระบวนการเมแทบอลิซึมภายในเซลล์อย่างลึกซึ้ง:

- แบบจำลอง ARCTICA ซึ่งเป็นการผสมผสานระหว่างการวิเคราะห์กระแสเมแทบอลิซึม (Flux Balance Analysis) กับ Machine Learning ได้ถูกพัฒนาขึ้นเพื่อระบุปฏิกิริยาสำคัญที่ควบคุมการผลิตในแบคทีเรีย *Synechocystis* sp. PCC 6803
- ในสาหร่าย *Dunaliella* การวิเคราะห์ข้อมูลทรานสคริปโตม (transcriptome) ถูกนำมาใช้ร่วมกับ Supervised Machine Learning เพื่อศึกษาจากข้อมูลการตอบสนองต่อความเค็ม



4. Optimizing Production Processes

ML helps optimize culture conditions and bioprocess parameters:

- A combination of Convolutional Neural Networks (CNN) and Genetic Algorithm increased phycocyanin production by 61.76% and biomass by nearly 90% in *Nostoc* sp. CCC-403.
- Artificial Neural Networks (ANN) combined with Deep Reinforcement Learning improved phycocyanin levels in *Plectonema* sp. UTEX 1541 by 20.1%.
- In *Chlorella*, integrating ANN with Genetic Algorithm boosted biomass by 78% and carbon dioxide capture by 77%.
- A Back-Propagation Neural Network predicted optimal conditions for producing astaxanthin — a high-value antioxidant — in *Haematococcus pluvialis*, achieving concentrations of 197.64 mg/L.

4. การเพิ่มประสิทธิภาพกระบวนการผลิต

ML ช่วยปรับสภาวะการเลี้ยงและกระบวนการผลิตให้ได้ผลผลิตสูงสุด:

- การใช้ CNN (Convolutional Neural Network) ร่วมกับ Genetic Algorithm ในการเพิ่มการผลิตไฟโคไซยานิน ในแบคทีเรีย *Nostoc* sp. CCC-403 ได้มากถึง 61.76% และเพิ่มชีวมวลได้เกือบ 90%
- การใช้ Artificial Neural Network และ Deep Reinforcement Learning สามารถเพิ่มความเข้มข้นของไฟโคไซยานินในแบคทีเรีย *Plectonema* sp. UTEX 1541 ได้ 20.1%

- ในสาหร่าย *Chlorella* นักวิจัยใช้ Artificial Neural Network ร่วมกับ Genetic Algorithm เพื่อเพิ่มผลผลิตชีวมวลได้ถึง 78% และเพิ่มประสิทธิภาพการดักจับคาร์บอนไดออกไซด์ได้ 77%
- การใช้ Back-propagation Neural Network ช่วยกำหนดสภาวะที่เหมาะสมในการผลิตสามารถแอสตาแซนทินในสาหร่าย *Haematococcus pluvialis* สามารถเพิ่มความเข้มข้นของแอสตาแซนทินได้ถึง 197.64 มิลลิกรัมต่อลิตร

These advancements illustrate how ML is transforming the development of bio-based industries and supporting sustainable technologies.

ความก้าวหน้าเหล่านี้สะท้อนให้เห็นศักยภาพของ ML ในการพัฒนาอุตสาหกรรมชีวภาพ และสนับสนุนเทคโนโลยีที่ยั่งยืน

