

# From bench to biofactory: high-throughput technologies and automated workflows to accelerate biomanufacturing

## Background/Objective

- Recent advances in enzyme engineering, DNA synthesis, and genome editing have led to notable progress in biomanufacturing
- Yet, a major obstacle to effective biomanufacturing is optimizing the combination of feedstock, host, and product together, as these decisions are interdependent

## Approach

- This paper focuses on the integration of automation, high-throughput technologies, self-driving and cloud labs, and sophisticated data management to enable Artificial Intelligence/Machine Learning and mechanistic models

## Significance/Impacts

- These advancements are expected to speed up the development of engineered microbes, offering more renewable, scalable, and cost-effective alternatives to petroleum-derived production and bolstering supply chain resilience
- The emergence of cloud labs and biofoundries, enabled by these approaches, will reshape biomanufacturing research by providing high levels of automation, reproducibility, and accessibility, democratizing access to state-of-the-art infrastructure and fostering collaborative research

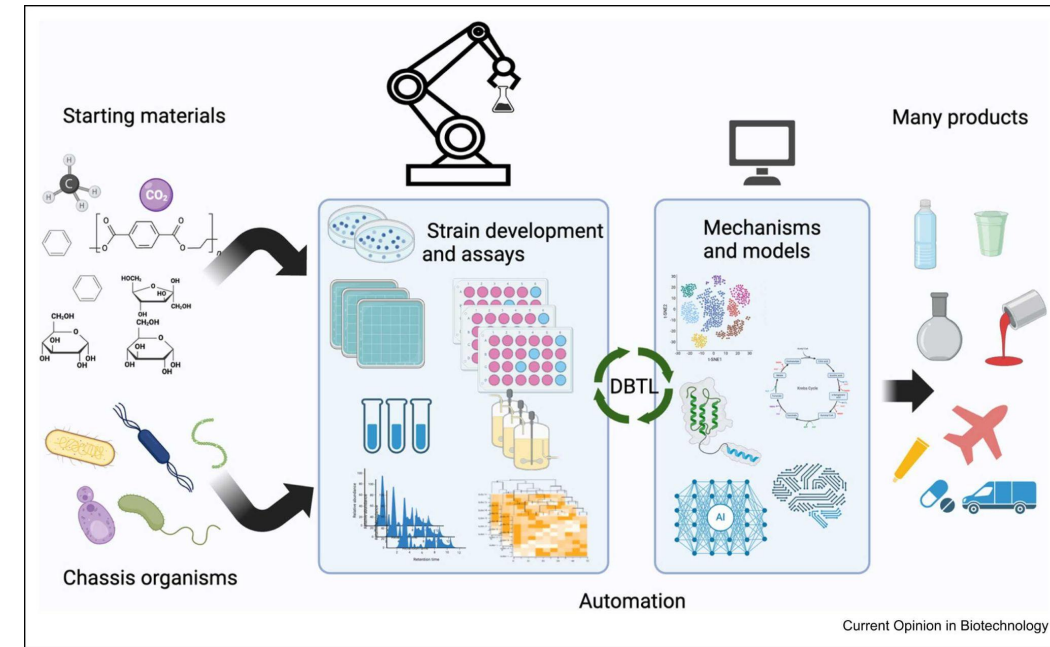


Figure caption: Many sources of starting materials can be combined with a range of microbial chassis for conversion but represent a vast design space. The vast space can be navigated using high throughput workflows that require automation and can lead to production of many valuable products for a robust bioeconomy.

# Engineering controllable alteration of malonyl-CoA levels to enhance polyketide production

## Background/Objective

- Expression of polyketide synthase (PKS) genes in *E. coli* has enabled the production of various valuable products. However, the limited availability of malonyl-CoA (M-CoA) in *E. coli* remains a substantial impediment to high-titer polyketide production

## Approach

- Disrupted the native malonyl-CoA (M-CoA) biosynthetic pathway and introduced an orthogonal pathway to enhance M-CoA production
- Applied adaptive laboratory evolution to identify mutations that further improved M-CoA availability and increased polyketide production

## Results

- Achieved significantly increased titers from four distinct PKS constructs
- Identified mutations from adaptive laboratory evolution that included both previously known M-CoA-enhancing variants and novel targets

## Significance/Impacts

- This work highlights the power of combining orthogonal control of metabolic pathways with ALE to enhance PKS precursor supply and polyketide titers while also deepening our understanding of M-CoA metabolism in microbial systems

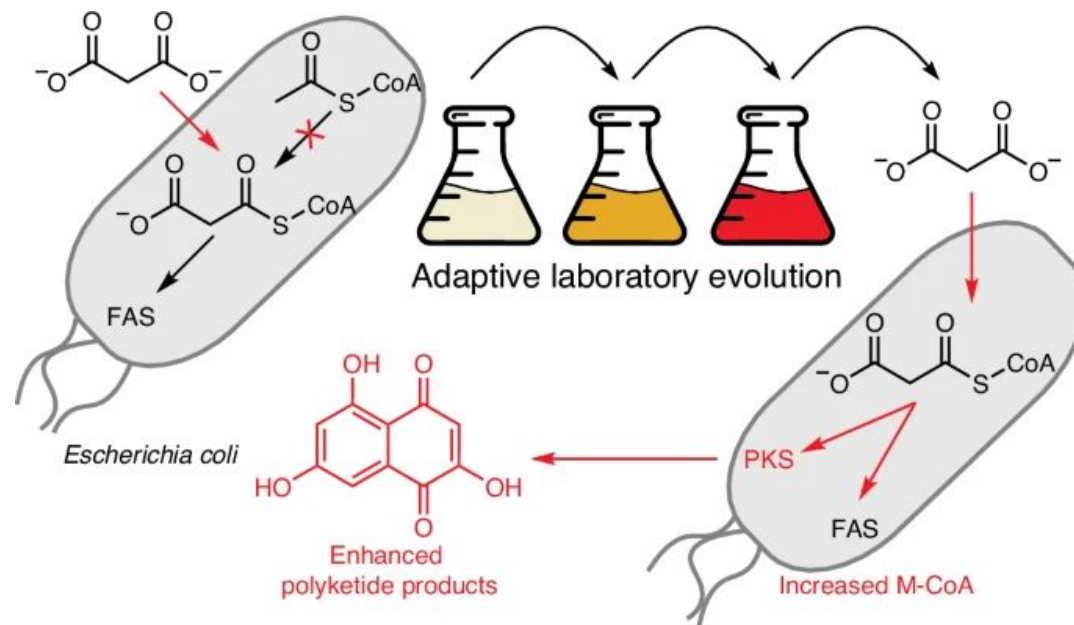


Figure caption: Disruption of the native M-CoA biosynthetic pathway and introduction of an orthogonal pathway consisting of a malonate transporter and an M-CoA ligase enables efficient M-CoA biosynthesis under malonate supplementation, resulting in tunable and enhanced M-CoA levels. Subsequent adaptive laboratory evolution of these strains identifies mutations that further elevate M-CoA levels and boost production of M-CoA-derived compounds, such as polyketides.

# Comparison between two different approaches for the deconstruction of lignocellulosic feedstocks using alkanolamine-based solvents

## Background/Objective

- This study aims to compare the deconstruction efficiency of various lignocellulosic feedstocks using ethanolamine and its protic ionic liquid form, ethanolammonium acetate, while also evaluating two different solvent recovery methods: water washing and vacuum distillation.

## Approach

- Alkanolamines have the potential for recovery and reuse that contribute to their appeal as sustainable agents in biomass pretreatment, positioning them as valuable tools in the development of renewable biofuels.

## Results

### Successfully demonstrated

- The effective use of alkanolamines as biomass pretreatment solvents relevant to a commercial biorefinery setting.
- That vacuum-based solvent removal is a better strategy for improved release of fermentable sugars that also enables facile solvent removal.

## Significance/Impacts

- This work makes significant contributions to advancing sustainable biofuel production processes by integrating green chemistry principles into the pretreatment of lignocellulosic feedstocks.
- It serves as a model for developing energy-efficient, resource conserving, and environmentally friendly technologies.

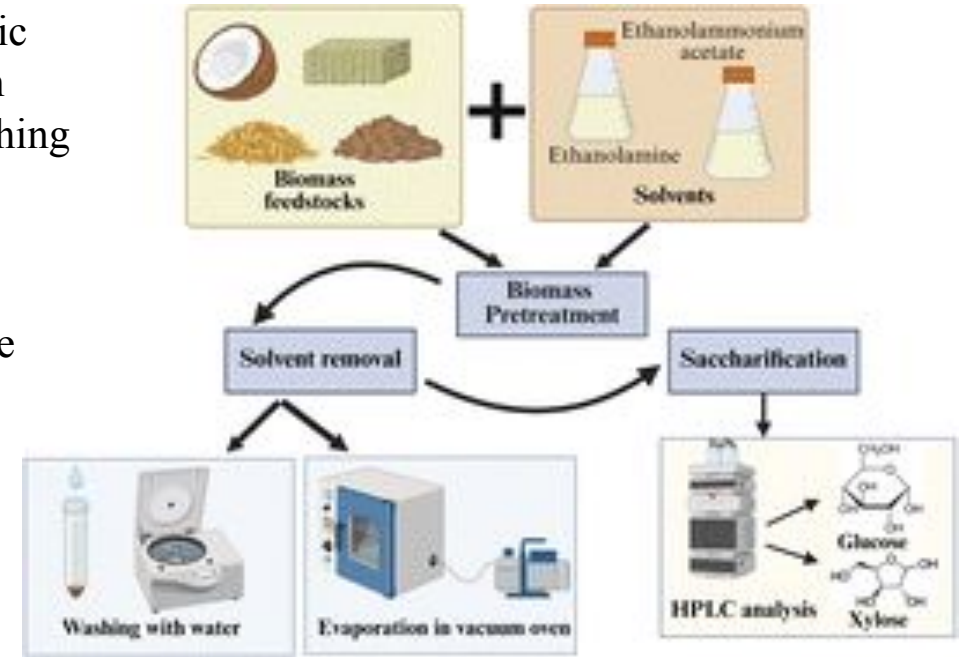


Figure caption: : Flowchart depicting the process of pretreatment of various lignocellulosic feedstocks with alkanolamines, followed by the two solvent recovery approaches – water washing and solvent evaporation via vacuum distillation.

# Can protein expression be ‘solved’?

## Background/Objective

- Heterologous protein expression is a fundamental technique very frequently used in modern day biology.
- Protein expression experiments primarily remain an experience-guided trial and error situation.

## Approach

- Here, we discuss the state of the field and identify the lack of large, high-fidelity datasets as the primary bottleneck to progress.
- We outline a proposed path toward an extensible experimental platform for collecting soluble overexpression data across organisms

## Results

- We suggest that the resulting data should be used to train predictive models of protein expression toward answering the question: can protein expression be solved?

## Significance/Impacts

- A predictive model of protein expression would have a profound commercial impact and could replace countless hours of experimentation with a higher-probability directed approach.

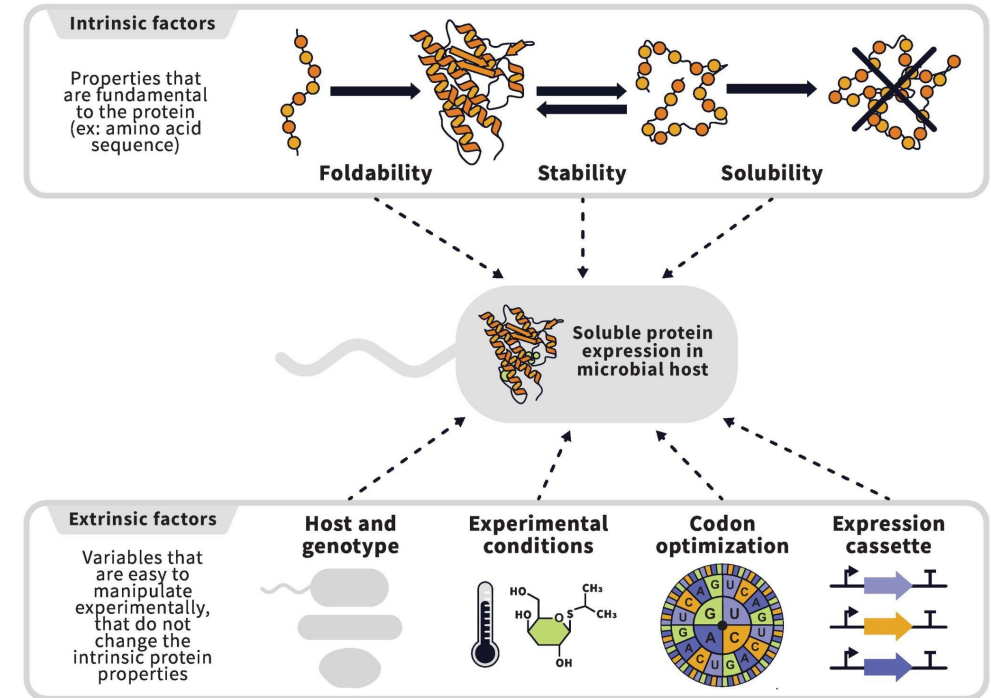


Figure caption: Intrinsic factors, those ‘hard-coded’ into the protein via its sequence, and extrinsic factors, those that can be changed experimentally, are both important to explore in a protein expression dataset.

# Regulatory orchestration of FK506 biosynthesis in *Streptomyces tsukubaensis* NRRL 18488 revealed through systematic analysis

## Background/Objective

- *Streptomyces tsukubaensis* NRRL 18488, the primary producer of the immunosuppressant FK506, was analyzed to elucidate regulatory features of secondary metabolism, which utilizes building blocks, such as CoA and amino acids, derived from primary metabolism.

## Approach

- A systematic approach to integrating genomic, primary transcriptomic, transcriptomic, and translational data from *S. tsukubaensis* to understand the multi-level regulatory mechanisms orchestrating secondary metabolism

## Results

- High-quality genome completion enabled precise FK506 BGC re-annotation
- Metabolic shift from primary to FK506-related pathways was observed
- Post-transcriptional regulation of allylmalonyl-CoA synthesis was discovered
- Ribosome pausing at the TTA codon within the FK506 BGC was elucidated

## Significance/Impacts

- Provides a roadmap for a fundamental synthetic biology approach to efficient BGC expression for the natural product biosynthesis in *Streptomyces*
- Demonstrates the power of integrating multi-omic data to reveal the complex, systemic nature of cellular mechanisms

*Streptomyces tsukubaensis* NRRL 18488 → FK506 (Tacrolimus)

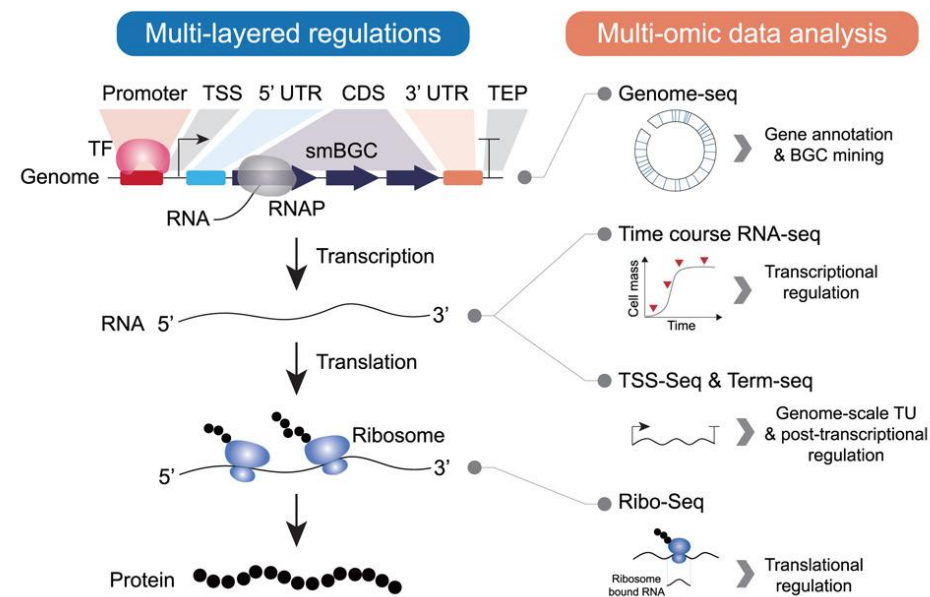
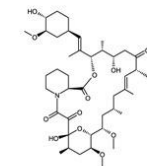
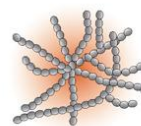


Figure caption: Multi-layered regulations from transcription to translation of the genes related to FK506 biosynthesis in *S. tsukubaensis* NRRL 18488 were elucidated by the multi-omic data analysis.

# Genomic and Transcriptomic Characterization of Carbohydrate-Active Enzymes in the Anaerobic Fungus *Neocallimastix cameroonii* var. *constans*

## Background/Objective

- Characterize novel anaerobic gut fungal isolate from stable fungal-methanogen natural consortium
- Evaluate how the isolate pairs in synthetic co-culture with methanogens

## Approach

- Genomically and transcriptomically sequence the fungal isolate
- Measure methane output from co-cultures of fungi with methanogens

## Results

- *N. cameroonii* var. *constans* genome encodes more glycoside hydrolases than most other anaerobic gut fungi
- Co-cultures of this fungus and methanogens produce methane from grass

## Significance/Impacts

- Few anaerobic gut fungi have been characterized this extensively
- This isolate can act as a versatile chassis organism to evaluate anaerobic gut fungal mechanisms and interactions in co-culture

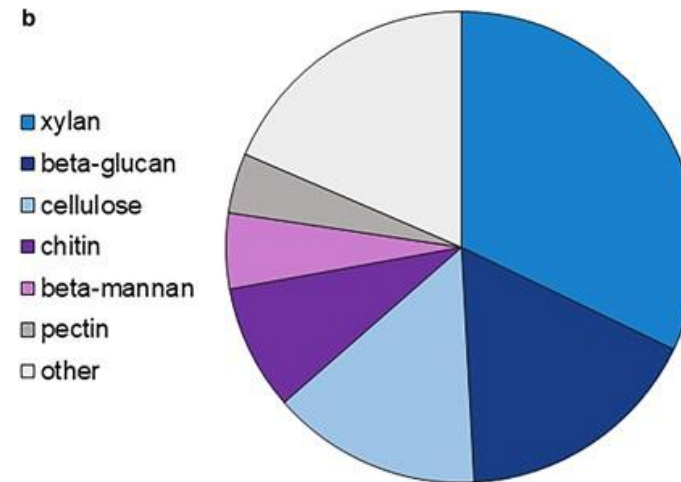
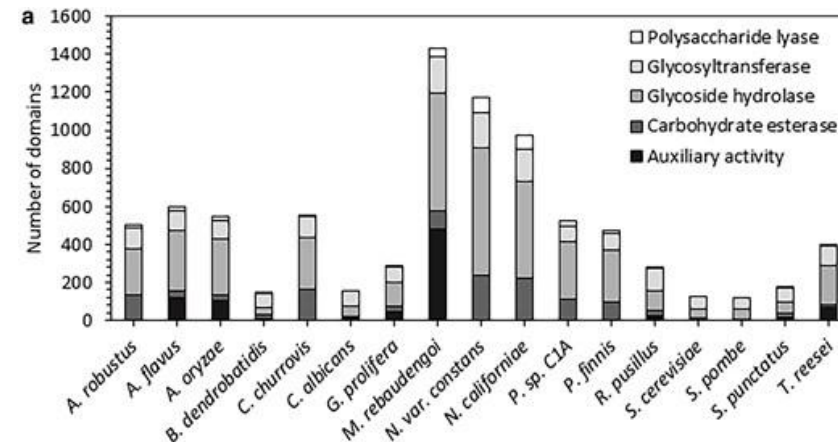


Figure caption: (A) *N. cameroonii* var. *constans* has among the highest CAZyme domain content compared to other sequenced fungi. (B) Substrate prediction for CAZymes identified in *N. cameroonii* var. *constans* using dbCAN-sub.

# The importance of ester cleavage in the butylamine pretreatment of hybrid poplar

## Background/Objective

- Butylamine was found to be very effective for pretreatment for a variety of lignocellulosic feedstocks
- Goal was to understand the mechanisms of the solvents action on the biomass

## Approach

- Butylamine pretreatment of hybrid poplar feedstock was studied under a variety of conditions
- Multi-modal spectroscopy and microscopy was used to study the ways butylamine changed the structure of lignocellulose

## Results

- Butylamine appears to selectively act on ester linkages within the lignin and hemicellulose components of the biomass

## Significance/Impacts

- Knowledge of mechanisms lets us select a family of promising volatile amine reagents for pretreatment
- Awareness of spectroscopic signatures of pretreatment enables in-situ characterization which is valuable for scale-up

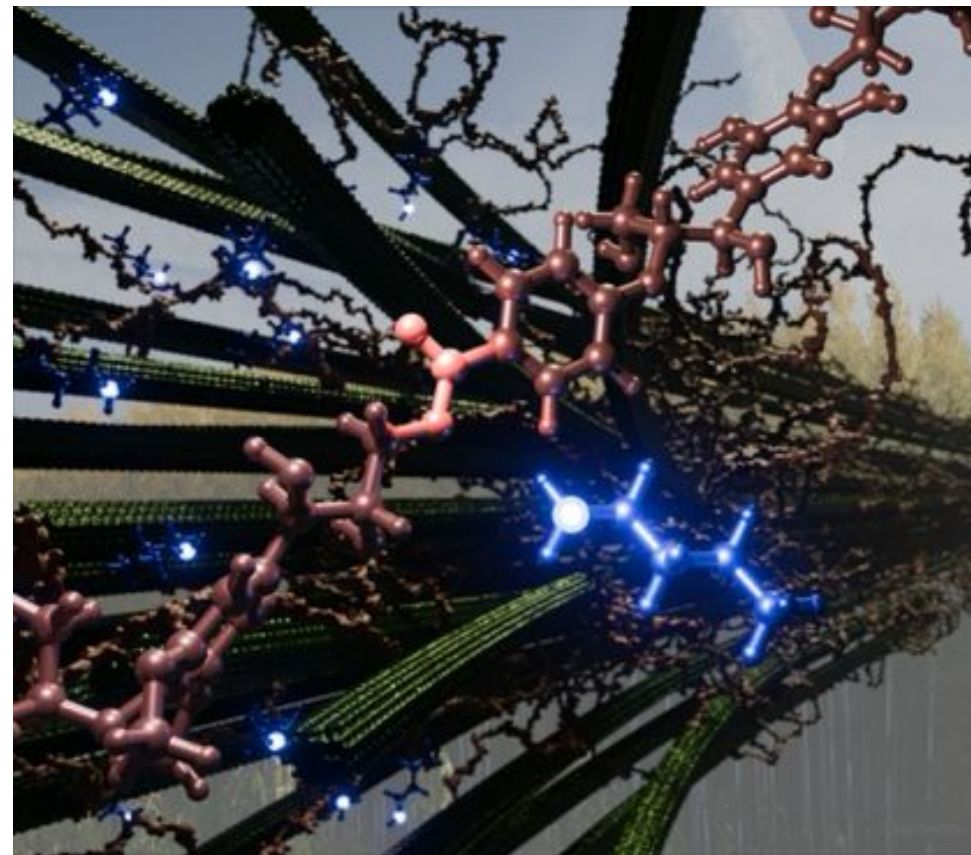


Figure caption: Cover art for the manuscript published in Green Chemistry (produced in-house at JBEI).

## Background/Objective

- Achieving a sustainable economy will require biorefineries to replace the current production of energy and chemicals from fossil carbon sources
- This work aims to inform the research community about the opportunities, challenges, and perspectives in developing truly sustainable lignocellulose-based biorefineries

## Approach

- This review discusses the current state of lignocellulosic biomass processing using ionic liquids (ILs) and deep eutectic solvents (DESs), focusing on pretreatment chemistry, process flows, and the valorization of each biomass component, along with sustainability assessments and key technological challenges

## Significance/Impacts

- ILs and DESs enhance the efficiency and sustainability of biomass processing but require careful evaluation of cost, toxicity, and recovery to ensure industrial feasibility and environmental compatibility
- Successful pilot deployments demonstrate the scalability and potential of IL- and DES-based technologies to enable sustainable, circular biorefineries and advance the biofuels and bioproducts industries

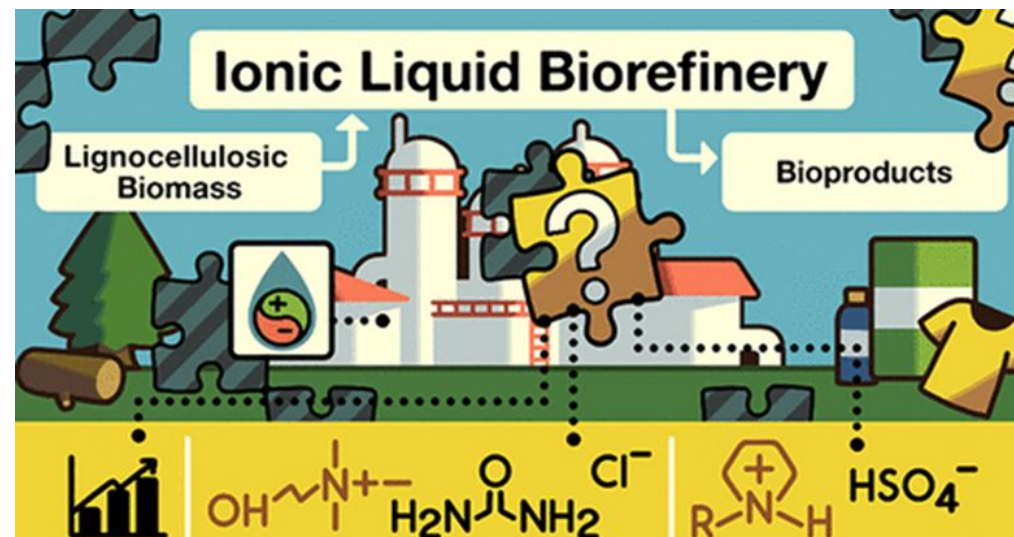


Figure caption: Puzzling out ionic liquid- and deep eutectic solvent-based lignocellulosic biorefineries for chemical and material production.

# Land-based resources for engineered carbon dioxide removal in the United States exceed the expected needs

## Background/Objective

- Perennial grasses can play an important role in converting atmospheric CO<sub>2</sub> into usable or storable form. This article explores the potential scale of biomass w/ C removal and storage (BiCRS)

## Approach

- Identified marginal and abandoned lands
- Modeled limited conversion of these lands at different prices

## Results

- Lands most attractive for conversion to switchgrass are in marginal climates, where dryland agriculture is possible but has an impact on yields

## Significance/Impacts

- Results highlight the importance of selecting bioenergy crops that are drought tolerant and maintain high yields and SOC accumulation under limited water

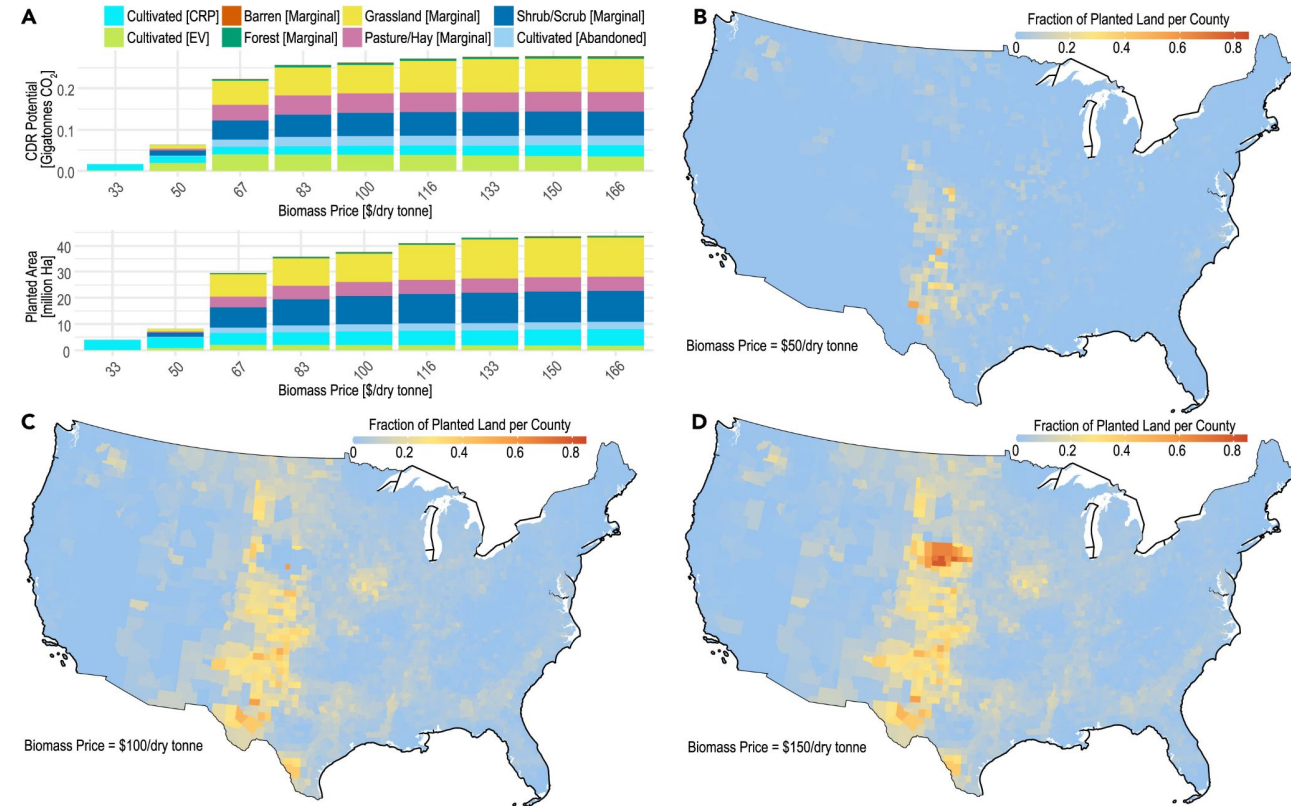


Figure 3: Availability and distribution of suitable land for carbon crop cultivation varies as biomass prices changes. (A) The availability of land for carbon crop cultivation area and the carbon dioxide removal potential varies as biomass price changes. Labels indicate land-cover class based on satellite data, and bracketed terms indicate the scenario identifying the land area for conversion (CRP, EV, marginal, and abandoned). (B–D) The spatial distribution of available land for biomass production at a farm gate biomass price of \$50, \$100, and \$150/dry tonne.

# Enabled Publications

# Dilemma of organic matter input to mitigate climate impact of rice paddies

## Background/Objective

- Application of organic matter has the potential to increase soil organic carbon; however, in rice paddy systems, it can also stimulate CH<sub>4</sub> emissions, which may offset the net benefits of carbon sequestration
- We investigated effects of organic amendments on CH<sub>4</sub> and N<sub>2</sub>O emissions from flooded rice paddies, and identify strategies through which organic amendments can enhance net soil C storage

## Approach

- In a field trial, we compared the effects of five organic matter types on net climate impact by quantifying annual gaseous fluxes and soil C stock changes with CO<sub>2</sub> equivalents

## Results

- All organic amendments increased CH<sub>4</sub> emissions compared with the no organic matter treatment (NPK)
- Biochar increased CH<sub>4</sub> emissions but also enhanced soil C stocks, leading to a net negative emission effect without compromising rice productivity

## Significance/Impacts

- Influence of organic amendments on SOC stocks in rice paddies can help optimize sustainable practices for one of the world's most important staple crops

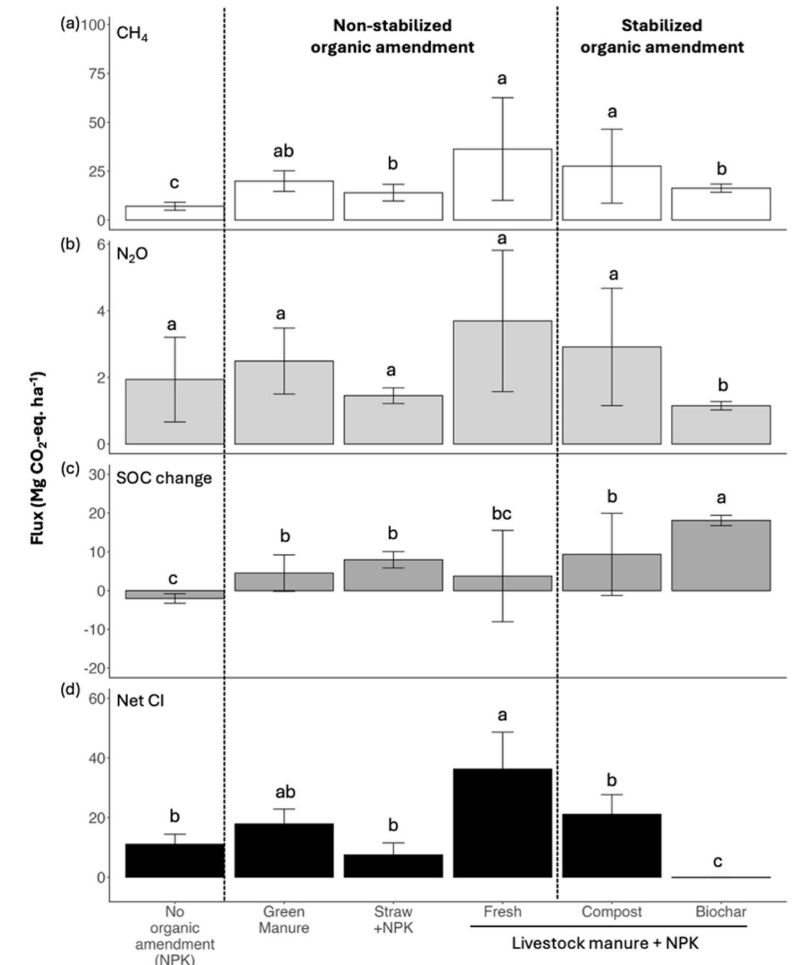


Figure caption: Rice grain yield and climate impact intensity (CII) after different organic amendment applications. Each bar shows a mean value, and the error bars represent the standard deviation.

# Consecutive low-frequency shifts in A/T content denote nucleosome positions across microeukaryotes

## Background/Objective

- Nucleosomes package ~150bp of DNA, are conserved across eukaryotes, and their positions are critical for gene regulation
- Little is known about how DNA signatures impact nucleosome organization. Here we use comparative genomic and multi-omic approaches to explore this

## Approach

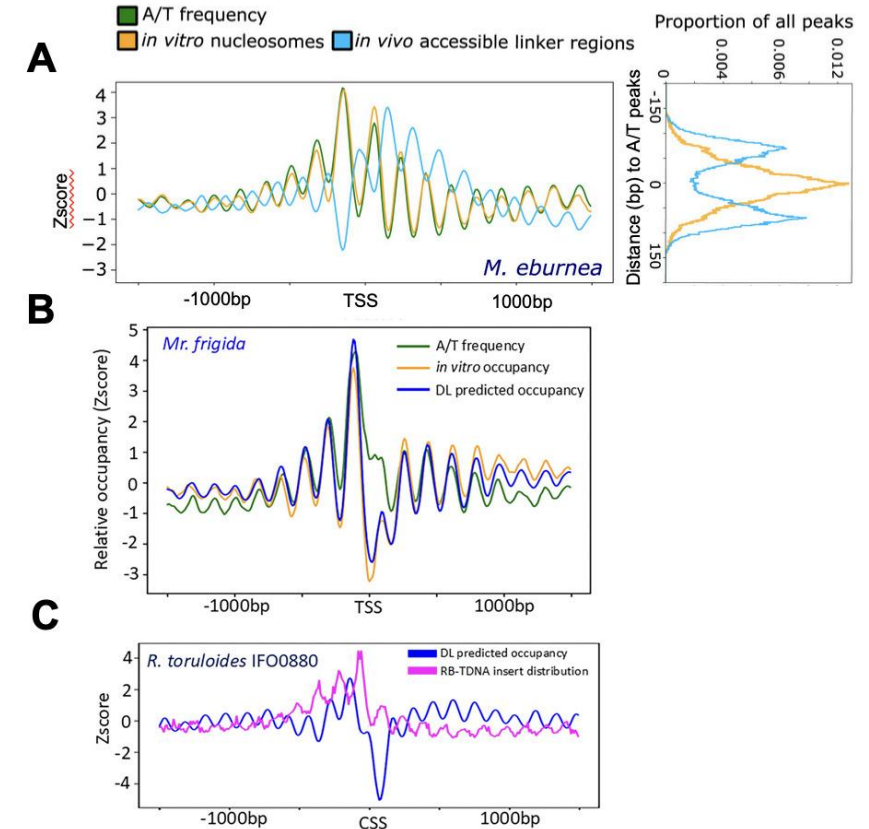
- Analyze >1000 microeukaryote genomes for ~150bp shifts in nucleotide frequencies across assemblies and gene features.
- Map nucleosomes in vitro and in vivo across fungi and algae. Compare to DNA signatures and DNA modifications.
- Train Deep Learning (DL) AI model to predict nucleosomes using DNA signatures

## Results

- ~150bp shifts in A/T frequency widespread across microeukaryotes
- A/T peaks associated with nucleosome positions in vitro and in vivo
- DL model trained using DNA features can predict fungal nucleosome positions

## Significance/Impacts

- Discovered eukaryote-wide strategy for generating cassettes of nucleosome-favorable DNAs
- Nucleosomes impact regulation and interfere with integration of foreign DNA. DL predictions can inform engineering efforts (e.g CRISPR & RB-TDNaseq)



# Engineering modular enzyme assembly: synthetic interface strategies for natural products biosynthesis applications

## Background/Objective

- Natural products remain indispensable sources of therapeutic and bioactive compounds
- Traditional discovery strategies are constrained by compound rediscovery

## Approach

- Modular biosynthetic enzymes, such as type I polyketide synthases (PKSs) and type A non-ribosomal peptide synthetases (NRPSs), offer promising platforms for combinatorial biosynthesis owing to their programmable architectures

## Results

- Synthetic interfaces including cognate docking domains, synthetic coiled-coils, SpyTag/SpyCatcher, and split inteins can function as orthogonal, standardized connectors to facilitate post-translational complex formation

## Significance/Impacts

- Synthetic interfaces can be integrated with computational tools to support a more systematic and scalable framework for modular enzyme engineering
- These approaches can accelerate the programmable assembly of biosynthetic systems and expand the accessible chemical space for natural products

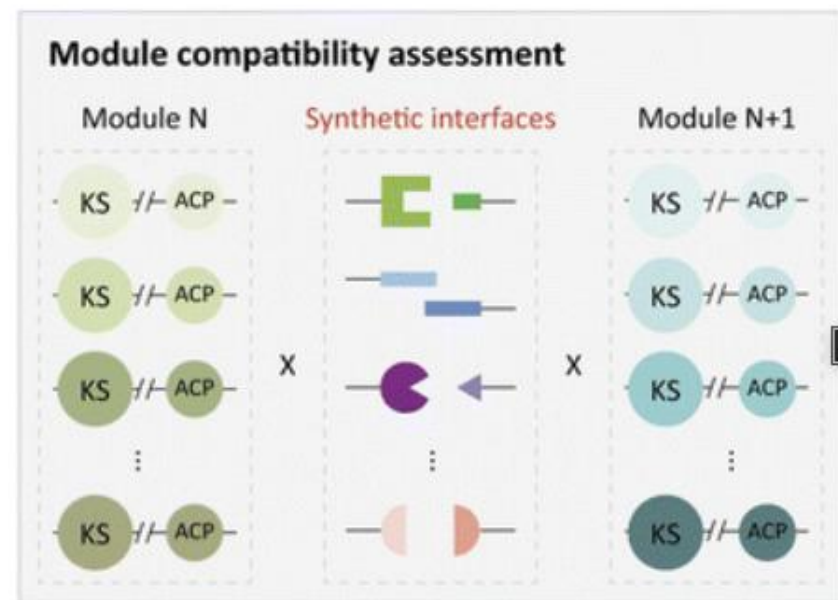
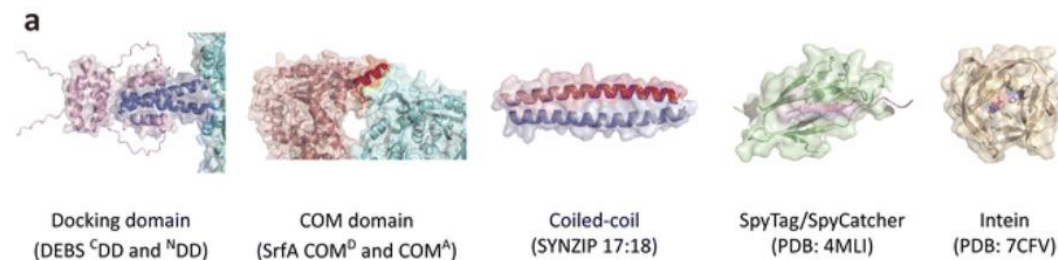


Figure caption: (Top) Representative structural models of synthetic interfaces used for modular assembly: DD (DEBS1-CDD/DEBS2-NDD, AlphaFold3), COM (srfAA-COMD/srfAB-COMA, AlphaFold3), coiled-coil (SYNZIP 17/18, AlphaFold3), SpyTag/SpyCatcher (PDB: 4MLI), and intein (PDB: 7CFV) pair. (Bottom) Functional applications of synthetic interfaces.