CASE STUDY

JUNE 2025

TWIST BIOSCIENCE

A CRE.AI.TIVE application of AI:

Engineering a more resilient global food supply

Phytoform's CRE.AI.TIVE platform leverages AI and CRISPR to engineer climate-resilient crops, but validating AI-designed variants at scale requires high-fidelity DNA synthesis for massively parallel reporter assays (MPRAs). In a study on drought-resistant tomatoes, CRE.AI.TIVE predicted the effect of millions of sequence variants, but synthesis challenges threatened the team's ability to validate these predictions. Twist Bioscience provided precise oligonucleotide synthesis, enabling Phytoform to construct a highly accurate MPRA library for 2,000 sequences. Experimental validation confirmed AI predictions, with 32 variants exhibiting a four-fold increase in gene expression. This study demonstrates how AI-driven design and accurate DNA synthesis can streamline trait engineering, accelerating crop improvement while optimizing resources.

Introduction

To meet the needs of a growing world population, it's expected that, by the year 2050, the global food system will need to increase its output by 50%.[1] However, scaling the food supply is far from simple, particularly as global warming threatens crops with flooding, drought, and other extreme weather. One step towards addressing this challenge is to develop crops that can better tolerate the effects of climate change and Phytoform is on a mission to do exactly that.

"Our goal is to make agriculture more sustainable by precisely engineering native plant genomes, rendering plants more resilient to the effects of climate change," explains Colleen Drapek, Trait Discovery Team Lead at Phytoform.

Situated within the historic Rothamsted Research Center—the world's longest-running agricultural research institution—Phytoform is developing a cutting-edge approach to trait development in crops, combining artificial intelligence (AI) with CRISPR to engineer more resilient plants. Current approaches to control gene expression in plants rely on a mix of unguided mutagenesis screens or transgenic technology. Phytoform leverages its CRE.AI.TIVE platform to discover and link cis regulatory elements (CRE) to gene expression functions *in silico*, Phytoform identifies key DNA changes in plant genomes that are linked to new gene expression patterns and will result in a desired trait (such as longer shelf-life, drought resistance, and other such qualities). This approach removes the need to empirically test the millions of potential variants that may exist for a given sequence. CRE.AI.TIVE guides the team to focus its engineering efforts where they're most likely to be productive, making the process far more efficient and scalable.

Yet, for CRE.AI.TIVE to be an effective tool, the team had to address a bottleneck that's currently slowing progress across the molecular sciences: The challenge of experimentally validating AI-driven designs on a large scale.

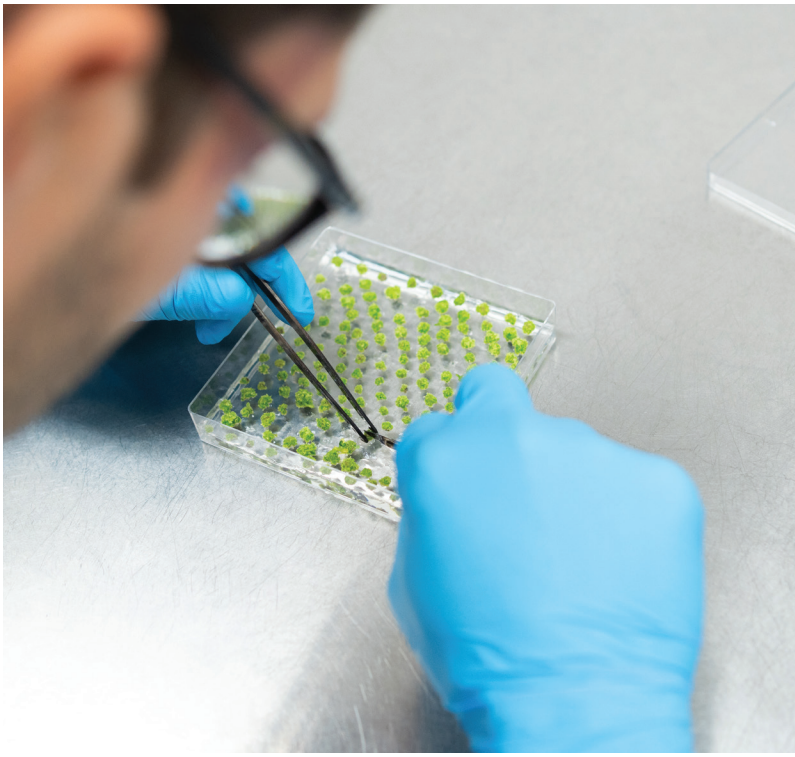
> **"Out of all the solutions we have tried, Twist's oligos are the most faithful to our designs. We absolutely need this fidelity when we're validating AI outputs."**
>
> COLLEEN DRAPEK, PHYTOFORM

ABOUT PHYTOFORM

Phytoform Labs is a crop trait development company on a mission to make agriculture more sustainable. From its headquarters in Hertfordshire, United Kingdom, Phytoform deciphers and rewrites plant DNA with its unique CRE.AI.TIVE platform to reduce inputs, lower waste, and enhance climate resilience. Through close partnerships with industry and research organizations, Phytoform delivers targeted genetic solutions that address today's most pressing agricultural challenges. Learn more about Phytoform's mission to unlock the benefits of crop genetics at phytoformlabs.com.

ABOUT TWIST BIOSCIENCE

Twist Bioscience is a leading and rapidly growing synthetic biology and genomics company that has developed a disruptive DNA synthesis platform to industrialize the engineering of biology. The core of the platform is a proprietary technology that pioneers a new method of manufacturing synthetic DNA by "writing" DNA on a silicon chip. Twist is leveraging its unique technology to manufacture a broad range of synthetic DNA-based products, including synthetic genes, tools for next-generation sequencing (NGS) preparation, and antibody libraries for drug discovery and development. Twist is also pursuing longer-term opportunities in digital data storage in DNA and biologics drug discovery. Twist makes products for use across many industries including healthcare, industrial chemicals, agriculture, and academic research. Learn more at twistbioscience.com.



The challenge: AI validation in the wet lab

CRE.AI.TIVE has the potential to accelerate trait engineering, but its computational predictions must be validated through large-scale massively parallel reporter assays (MPRAs) (see Figure 1). Doing so requires precise synthesis of AI-designed DNA sequences—an area where synthesis errors and oligo length constraints can limit assay scale and efficiency.

For example, the maximum oligo length determines the genomic window in which variation can be tested, while synthesis fidelity dictates how well the model is trained. Synthesis errors not only reduce the representation of specific variants, but it dilutes the testing library with sequences that lack experimental value.

"The main problem with this is the waste of resources at each step," explained Drapek. "Usually we build a library so that each promoter variant has ~100 barcodes, but the variant must be perfect every time to score it. If promoter synthesis is error prone, you have to compensate by making more of each variant, creating a larger library without actually adding additional information. You will then have to collect more RNA material to ensure your chance of catching a signal, and then you still lose valuable NGS resources to thousands of reads that will be thrown away in the bioinformatics QC pipeline. It ends up spilling into all stages of the building, testing, and learning cycle."

Put simply, poor-quality DNA synthesis can be a bottleneck in the validation process, limiting the number of AI-designed variants that the lab can reasonably test. Not only does this slow the screening process, but it can potentially lead to the exclusion of valuable sequences from analysis.

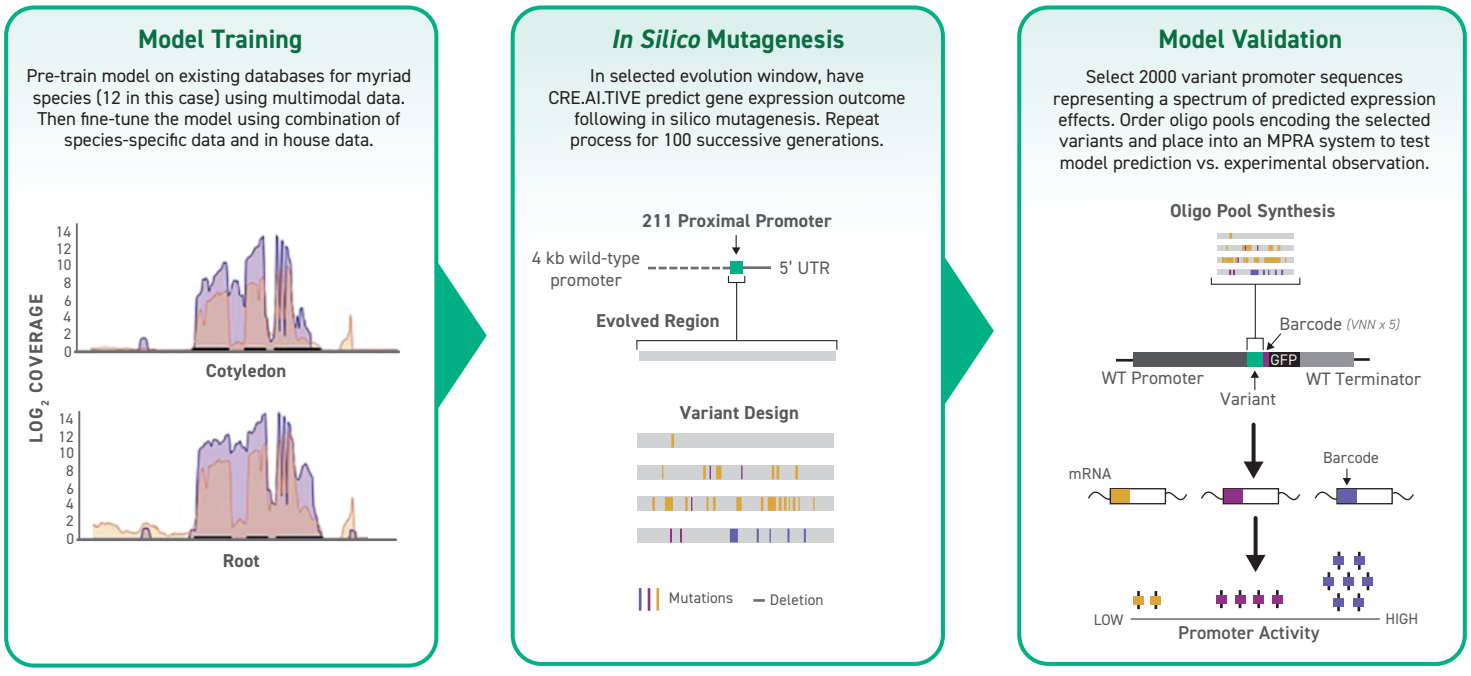
The solution: precision DNA synthesis

In a recent project,[3] the Phytoform team sought to develop a more drought resistant tomato. To do this, they set out to evolve a proximal promoter of a relevant gene in the tomato genome, SlbHLH96, previously linked to drought tolerance through overexpression. The AI model predicted gene activity for more than 10 million designed DNA sequences, and the team chose 2,000 to proceed with synthesis and MPRA validation. However, synthesizing the novel sequences proved technically challenging as many of them contained AT-rich content and multiple stretches of homopolymers.

Drapek explains that "this is not unusual or exceptional in the plant kingdom—many plant promoters have complex features. But this made it particularly difficult for companies to synthesize our oligo pools. Some flat out refused to do so."

Initially, Phytoform ordered a small pool of oligos from a popular DNA provider, but the challenging sequences proved too difficult. As Drapek described, "It was awful, we had to chuck out almost 60% of the pool." Knowing they would need better, the Phytoform team enlisted Twist Bioscience's help to produce the needed 2,000 AI-designed variants for MPRA validation.

"When training and testing the model, we expect oligo synthesis to be exactly as we designed it," Drapek explained. "Out of all the solutions we have tried, Twist's oligos are the most faithful to our designs. We absolutely need this fidelity when we're validating AI outputs."



> **"Combining AI with high quality oligo pools allowed us to approach sequence evolution in a guided way. Without these tools, we would have to burn resources building out much larger libraries without knowing whether we're searching in the right variant space or not."**
>
> COLLEEN DRAPEK, PHYTOFORM

Results: AI-guided trait engineering

To validate CRE.AI.TIVE's predictions in a real-world setting, the team set about evolving a proximal promoter upstream of the tomato SlbHLH96 gene. In this context, evolution involves the iterative mutagenesis of the promoter sequence, with each new variant paired with a predicted gene expression outcome. The SlbHLH96 proximal promoter was subject to 100 rounds of *in silico* mutagenesis with iterative AI feedback, producing millions of unique digital sequence variants with a wide range of predicted effects on SlbHLH96 expression. From these millions, 2,000 sequences representing a range of predicted gene expression effects were selected for MPRA screening.

An oligo pool coding for the 2,000 AI-designed sequences was synthesized by Twist and delivered to the Phytoform team. "Our expected loss is 25%," Drapek emphasizes. "With the Twist pool, we saw a loss of 15%, and we can attribute most of that to errors in amplification, not synthesis."

The variants were then transfected into tomato cotyledon protoplasts for MPRA screening. Results showed a positive correlation between predicted and measured gene expression changes. Of the 1,700 variants predicted to increase gene expression, 163 were found to have at least a two-fold increase in gene expression, with 32 demonstrating a four-fold increase.

Impact and future directions

Phytoform's study is a strong proof of principle for the plant engineering community. "Combining AI with high quality oligo pools allowed us to approach sequence evolution in a guided way," Drapek explained. "Without these tools, we would have to burn resources building out much larger libraries without knowing whether we're searching in the right variant space or not."

In bringing these tools together, Drapek's team is helping to push the field to new heights.

"Increasing gene expression four-fold is an important achievement in the field," she adds. "It's been much easier to 'break' promoters with random insertion/deletions via CRISPR, almost always resulting in decreased gene expression." Options for increasing gene expression, she adds, remain limited. "There are some cases of natural variation or EMS-induced SNPs in upstream regions that result in gene overexpression, but these are usually on the order of ~2 fold increase."[2]

Beyond demonstrating the ability to engineer gene activity increase, this study is also significant because it is the first of its kind to demonstrate the predictive potential of AI-algorithms in plant cells—a model system with a strong track record for producing results that translate to the farmer's field. "Instead of just going straight to plant studies with a large selection of variants, you can efficiently home in on a small group of promising candidates," says Drapek.

In short, CRE.AI.TIVE allowed the team to rapidly explore target sequence space for higher gene activity variants, predicting the effects of millions of edits and guiding the selection of 2,000 candidates for wet-lab validation. Twist Bioscience's accurate synthesis platform enabled the team to faithfully transfer these sequences to the wet lab for protoplast MPRA screening. Ultimately, this workflow can help conserve resources and ensure that only the most promising candidates move to the more costly stage of *in vivo* testing.

Future directions

Drapek and the team at Phytoform will continue to refine CRE.AI.TIVE and explore its predictive potential across species in silico, in vitro, and in vivo. Of interest to the team is the use of Twist's new Multiplexed Gene Fragments, which enable the large-scale synthesis of double-stranded DNA fragments between 300 and 500 bp in length.

"Evolving longer stretches of DNA is where the field is going," Drapek states. "If you give the model more space to explore, it increases your chances of finding the best solution."

Ultimately, CRE.AI.TIVE holds significant potential for engineering climate-resilient crops, making trait development faster and more flexible across multiple species. "It's just so much faster now for researchers to engineer traits—not just in one well-studied species, but across many different species," Drapek concludes. "I'm excited to see how CRE.AI.TIVE plays out in so many applications for precision breeding. It's harnessing the power of DNA and prediction for agricultural impact!"

REFERENCES

1. IPCC. "Chapter 5 - Special Report on Climate Change and Land." Ipcc.ch, Special Report on Climate Change and Land, 2019, www.ipcc.ch/srccl/chapter/chapter-5/

2. Zhang, Junli, et al. "Sequencing 4.3 Million Mutations in Wheat Promoters to Understand and Modify Gene Expression." Proceedings of the National Academy of Sciences of the United States of America, vol. 120, no. 38, 13 Sept. 2023, https://doi.org/10.1073/pnas.2306494120.

3. Jevtic, Sania, et al. A Scalable Method for Modulating Plant Gene Expression Using a Multispecies Genomic Model and Protoplast-Based Massively Parallel Reporter Assay. 9 Dec. 2024, https://doi.org/10.1101/2024.12.05.626999.

Want to use Twist's DNA synthesis Services for your own AI-guided studies? Visit **twistbioscience.com.**

If you want to see how Phytoform and their CRE.AI.TIVE platform can support your work, visit **phytoformlabs.com.**

WHAT CAN TWIST DO FOR YOU?

sales@twistbioscience.com
twistbioscience.com

#WeMakeDNA