TelesisBio

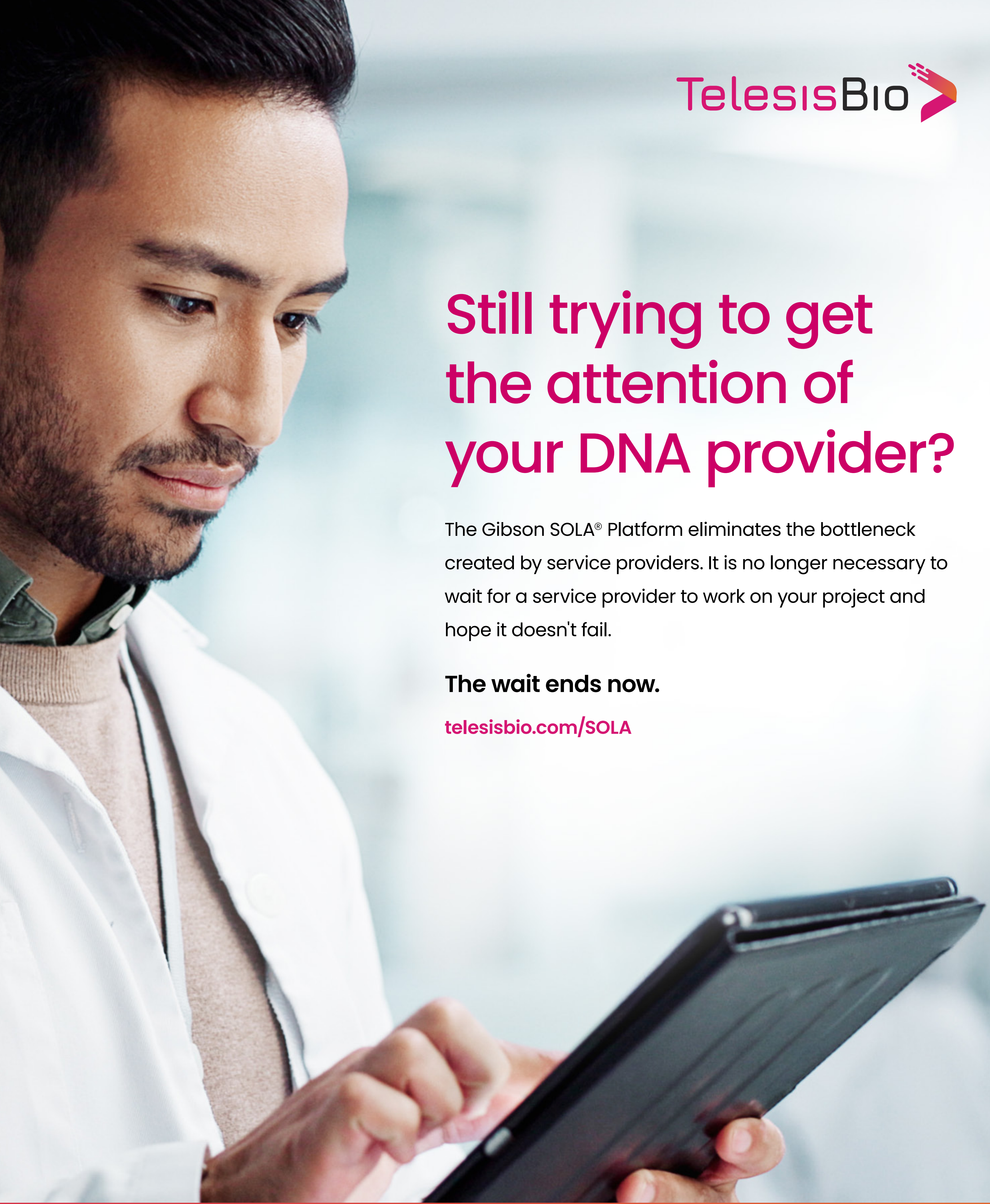
Breaking the Bottleneck

On-Demand Biology in the Age of AI

A SPONSORED PUBLICATION FROM

GEN Genetic Engineering & Biotechnology News

TelesisBio

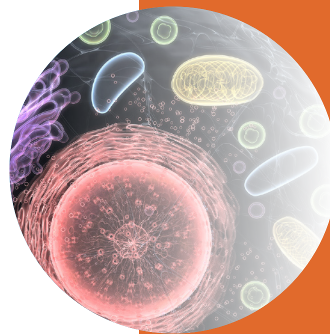


Still trying to get the attention of your DNA provider?

The Gibson SOLA® Platform eliminates the bottleneck created by service providers. It is no longer necessary to wait for a service provider to work on your project and hope it doesn't fail.

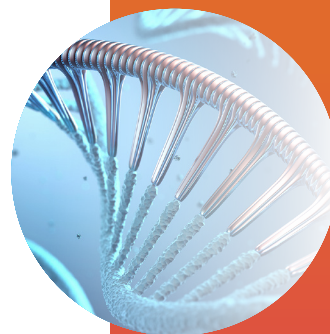
The wait ends now.

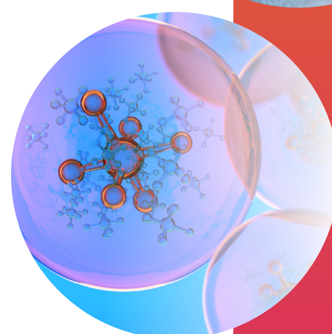
telesisbio.com/SOLA

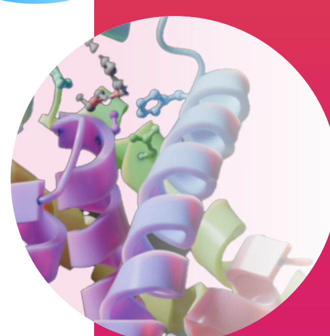


Breaking the Bottleneck

On-Demand Biology in the Age of AI







Cover: Metamorworks/Getty Images

A Faster Path from Idea to Impact in Modern Biology Workflows

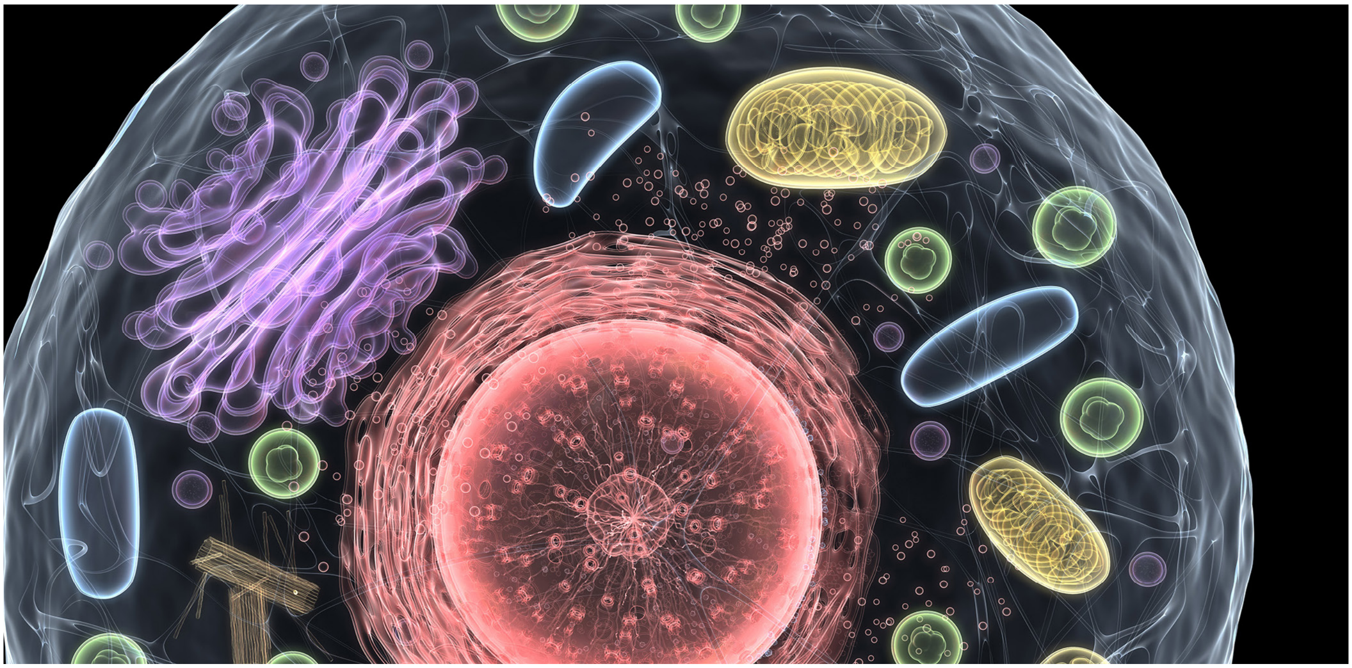
The worlds of synthetic biology and artificial intelligence are becoming increasingly interconnected, reshaping how scientific breakthroughs are achieved. What previously required months of careful benchwork and slow iteration can now be completed in days. With the help of machine learning models and on-demand synthesis tools, researchers can design, build, and test biological systems at a pace that simply wasn't feasible before.

Across drug discovery, crop engineering, and protein design, researchers are turning to AI to generate hypotheses and select the most promising candidates with greater confidence. At the same time, tools like Telesis Bio's Gibson SOLA Platform allow labs to synthesize DNA and mRNA on-demand, bypassing the wait times and constraints of historically outsourced services. The net effect is speed coupled with flexibility, repeatability, and better decision-making early in the pipeline.

In therapeutic development, deep learning models drive advances in antibody and mRNA design. In the field of protein engineering, integrated platforms identify candidate molecules and feed results back into the design loop, creating an iterative cycle of improvements. And in agriculture, insights from pangenomic datasets are helping scientists understand the role of paralogs in traits that can be edited with high precision. These capabilities are turning data into action much more efficiently without sacrificing quality or creativity.

What connects all these innovations is a shift in how research teams approach the process itself. They're no longer limited by long synthesis timelines or slow feedback loops. Instead, they're moving ideas into experiments more quickly, testing more options in less time, and refining their approach with each cycle. Whether you're working in a startup, academic lab, or biopharma R&D team, the chapters ahead offer a window into what's working now and what's possible next.

Synthetic Biology Meets AI to Craft Life with Computational Precision



enot-poloskun / Getty Images

Synthetic biology and machine learning might sound like buzzwords from science fiction, but together they're reshaping how scientists create and modify organisms. Imagine being able to program bacteria to produce medicine, predict exactly how gene editing will turn out, or design entirely new biological entities. Thanks to artificial intelligence (AI), these futuristic scenarios are becoming today's reality.

Metabolic Pathway Optimization

One key area significantly impacted by artificial intelligence (AI) in synthetic biology is metabolic pathway optimization. Metabolic pathways are the series of chemical reactions that occur inside cells, converting nutrients into energy, building blocks, or specific products like biofuels, pharmaceuticals, or industrial chemicals. Optimizing these pathways involves adjusting cellular metabolism to boost the production of desired compounds. Previously, this process required extensive experimentation using traditional methods of trial and error. It was necessary to systematically test different genetic modifications in cells such as bacteria or yeast, observing which alterations led to higher yields. Unfortunately, this conventional approach was slow, expensive, and inefficient.

Today, machine learning algorithms have transformed this scenario. These AI models are trained on massive datasets containing genomic data, experimental outcomes, and biochemical knowledge. With this information, they can precisely predict the impact of genetic changes on a cell's metabolic functions. Instead of time-intensive trial and error, it is now possible to receive targeted guidance on which genes to modify for optimal results. This computational approach drastically reduces the experimental workload, accelerates development timelines, and cuts costs. Consequently, AI-driven metabolic pathway optimization empowers synthetic biologists to quickly and reliably engineer microbes that efficiently produce valuable substances, greatly enhancing both research capabilities and industrial applications.

Mapping Metabolic Pathways

Flux balance analysis (FBA) is a computational approach to analyze how substances flow within a cell's metabolic networks. Think of it as constructing a comprehensive map illustrating the precise routes that molecules, like nutrients or metabolites, take inside cells such as bacteria or yeast. Each cell contains numerous interconnected pathways, each performing specific functions like generating energy, discarding waste, or synthesizing valuable products like pharmaceuticals (figure 1). Historically, enhancing these pathways to boost production involved extensive trial-and-error experimentation. It was necessary to manually modify genes without certainty of success, resulting in a lengthy, costly, and uncertain process.

Recently, scientists have begun integrating FBA with machine learning algorithms to significantly streamline this process. Machine learning systems can analyze vast amounts of biological data, including genetic sequences and outcomes from prior experiments, enabling precise predictions about how specific genetic changes will influence metabolic outputs. This advanced approach eliminates much of the uncertainty involved, providing scientists with accurate, evidence-based guidance on the best genetic modifications to implement. Consequently, this powerful combination of AI and FBA has dramatically accelerated scientific discovery, enabling more efficient microbial engineering for producing medications, renewable energy sources, and various industrially relevant

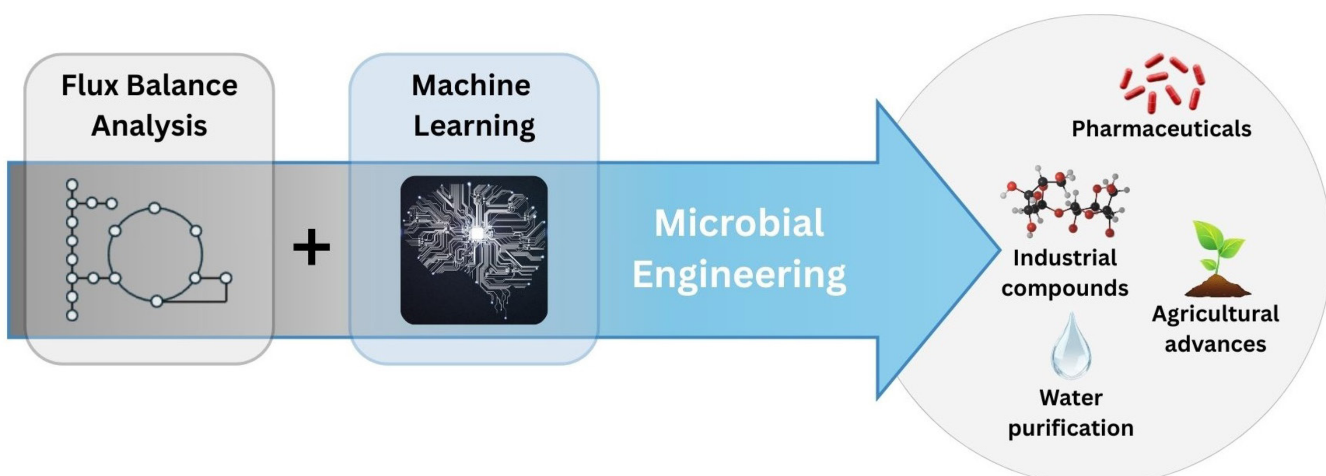


Figure 1. Computational Flux Balance Analysis. AI models accelerate pathway determination for microbial engineering in applications ranging from drug development to industrial compound engineering.

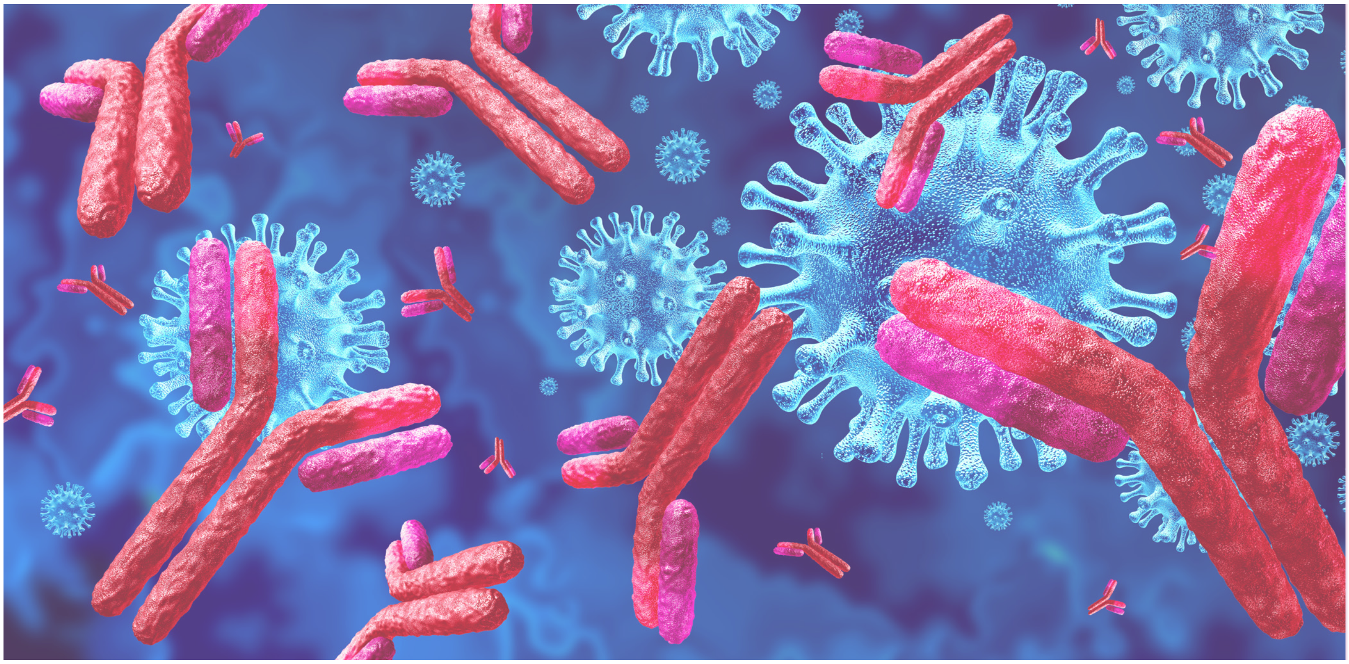


Figure 2. AI Models Predict Higher Probability Antibody Targets. AI models can identify patterns and predict high-affinity, specific binders and also support virtual screening and affinity maturation.

compounds, while simultaneously lowering costs and improving accessibility.

Streamlining Antibody Discovery

Similar to how scientists focused on Metabolomics are using AI to shift from serendipitous discovery toward rational, data-driven discovery, AI and ML are revolutionizing traditional antibody discovery workflows. By dramatically accelerating the design, screening, and optimization of therapeutic antibodies scientists can drive towards more efficient use of resources and improve overall productivity. Traditionally, antibody discovery relied on time-consuming processes such as hybridoma screening or phage display, which required labor-intensive wet-lab work and iterative trial-and-error optimization. AI now enables in silico prediction of antigen–antibody binding, immunogenicity, and developability, allowing researchers to narrow down candidate pools with far greater efficiency and precision before even entering the lab.

These models trained on massive datasets such as antibody sequence–structure–function relationships that can identify patterns and predict high-affinity, specific binders and also support virtual screening and affinity maturation, dramatically reducing the time and cost needed to advance a lead candidate (figure 2). These tools are ultimately streamlining development pipelines and enabling more personalized and potent therapies while getting new therapies to market faster and with reduced developmental cost.

Gene Editing

In addition to streamlining traditional workflows, AI is significantly improving the safety and accuracy of genome editing, particularly when combined with CRISPR technology. While CRISPR revolutionized biological research by enabling precise genetic modifications, it still faces significant challenges—particularly unintended edits or off-target effects. These accidental alterations can lead to unwanted side effects that can be potentially harmful in therapeutic settings or problematic in genetically modified crops.

AI, specifically machine learning, is addressing this challenge head-on. Machine learning models analyze extensive datasets containing genetic sequences and past editing outcomes to identify patterns that human researchers might miss. By detecting subtle relationships between gene sequences and editing accuracy, these algorithms predict where unintended edits are likely to occur. With these insights, it's now possible to carefully select CRISPR targets and design experiments that dramatically reduce the chances of errors.

The practical implications of this AI-driven precision are enormous. In medicine, it means gene therapies can become safer, paving the way for treatments of genetic disorders without the fear of harmful side effects. In agriculture, genetically modified crops can be developed with greater confidence, enhancing their nutritional value or resilience to climate conditions. Ultimately, by making genome editing safer and more reliable, AI amplifies CRISPR's potential to beneficially transform human health and agriculture.

Removing Barriers and Delays

One exciting innovation breaking barriers in synthetic biology is overnight DNA synthesis using the Gibson SOLA Platform by Telesis Bio. Traditionally, synthesizing precise DNA sequences required substantial time, delaying research and limiting experimentation flexibility. However, the SOLA Platform dramatically changes this by enabling rapid, automated DNA synthesis directly within researchers' labs overnight. This technology

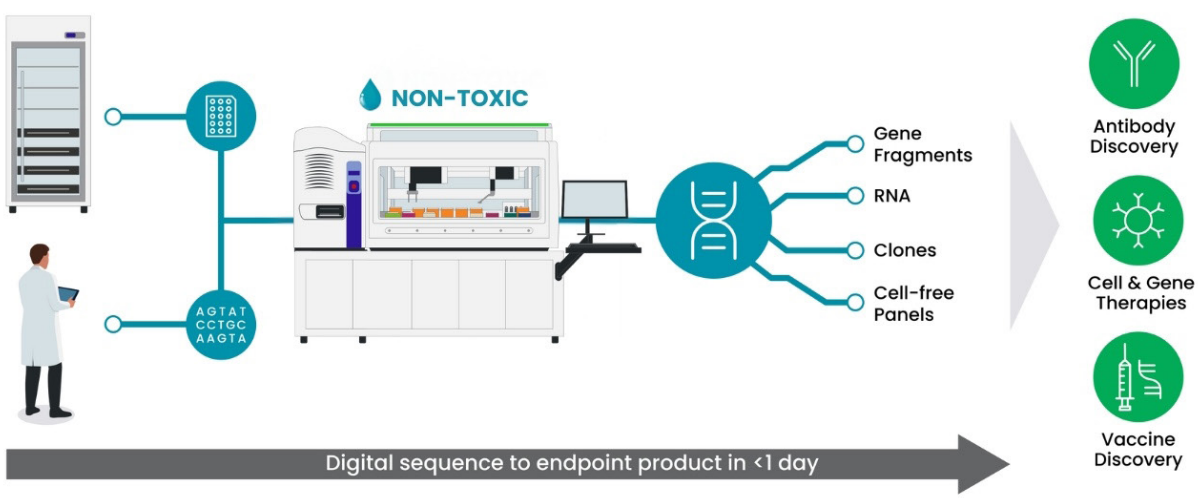


Figure 3. Digital Sequence to Biological Endpoint in 1 Day. The Gibson SOLA Platform is a 'digital to biological converter' that generates DNA or mRNA in less than one day.

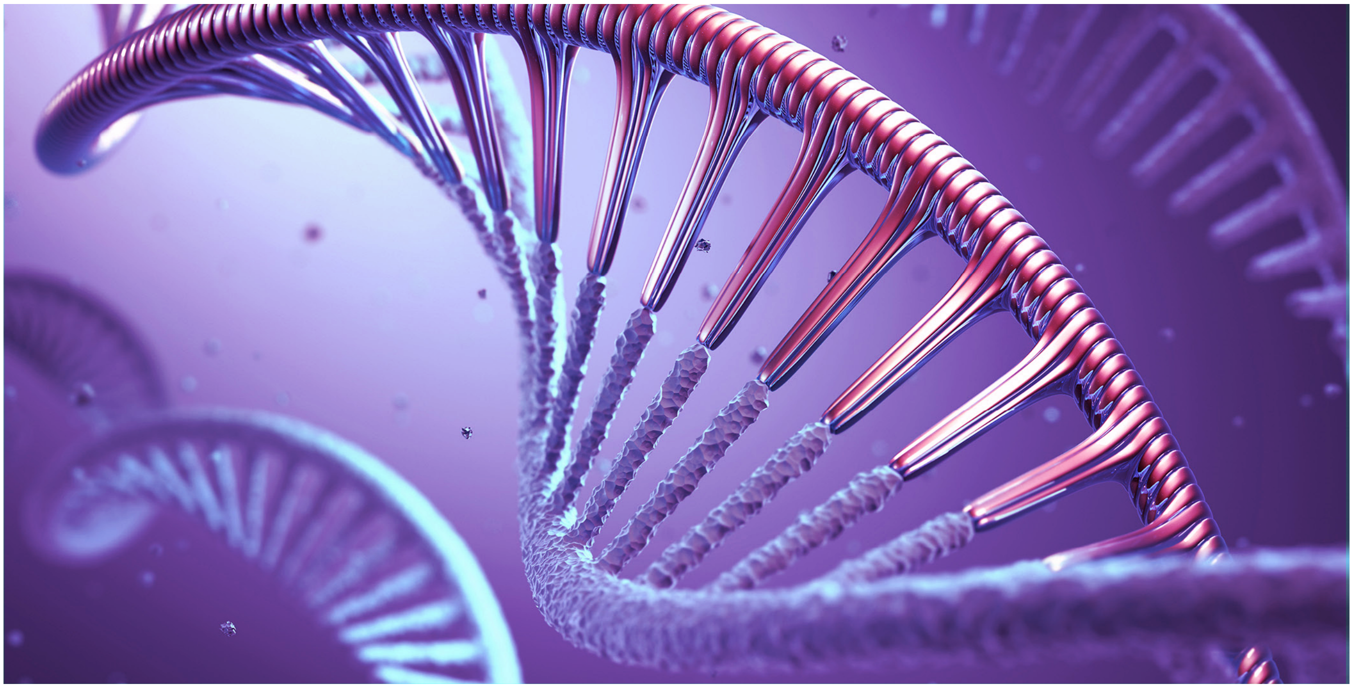
combines state-of-the-art enzymatic methods with advanced robotics, removing the dependency on external service providers and shipping delays (figure 3).

The Gibson SOLA Platform streamlines workflows significantly, making it possible to quickly design, test, and iterate genetic constructs within just one day. This acceleration allows researchers to rapidly advance their therapeutic development and protein engineering projects, speeding up the overall pace of scientific discovery. Furthermore, it reduces logistical hurdles like supply chain disruptions or delays caused by external providers. When paired with automation tools like the Echo 525 Acoustic Liquid Handler and the Biomek i7 Automated Workstation from Beckman Coulter Life Sciences, labs can consistently output more than 50kb of DNA per week.

By making DNA synthesis fast, reliable, and immediately accessible, Telesis Bio's Gibson SOLA Platform not only enhances research productivity but also lowers barriers to innovation, allowing labs to participate actively in groundbreaking synthetic biology research.

For more information on how the Gibson SOLA Platform is transforming DNA synthesis, visit telesisbio.com/sola.

Accelerating Drug Discovery with Generative AI



ktsimage / Getty Images

Generative AI is changing how scientists design new medicines. Traditional drug discovery can be slow and expensive, especially when it comes to finding the right molecule to fight a specific disease. With the help of powerful computer models, scientists can now create and test molecular designs much faster than ever before. Two of the biggest breakthroughs in this area are generative adversarial networks (GANs) and diffusion models. These systems can design completely new molecules—sometimes so new they wouldn't have occurred to any human researcher.

GANs use a "generator" that attempts to create something such as a new molecule while a "discriminator" judges if that molecule matches desired criteria. These two networks learn from each other, so the generator keeps improving its designs until the discriminator can't tell them apart from a known set of good examples (figure 1). GAN-based tools such as the DrugGEN system, which has been used to identify AKT1 inhibitors for the treatment of certain types of cancer, and MedGAN which can predict effective molecular structures to aid in repurposing existing drug candidates, are redefining how traditional drug discovery can be accomplished.

Meanwhile, diffusion models start with random noise and repeatedly remove that noise in tiny steps. Eventually, they form a clear image of a molecule that could meet the specified criteria. The ability of these models to incorporate structure-based conditioning, multi-modal inputs, and gradient-free sampling makes them extremely powerful for next-generation drug discovery

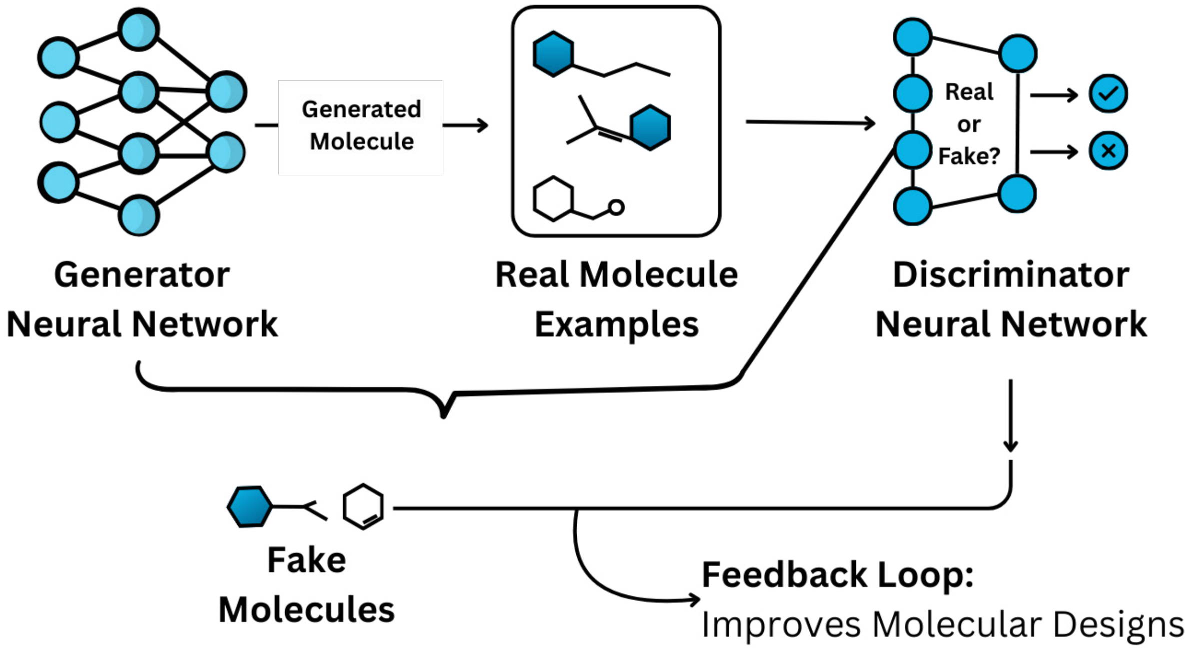


Figure 1. GAN Networks. Generative Adversarial Networks (GANs) use a generator to create novel molecular candidates. These are evaluated by a discriminator against known molecules, providing feedback that improves future outputs. Gibson SOLA Platform is a 'digital to biological converter' that generates DNA or mRNA in less than one day.

pipelines. With both methods, what used to take years of trial and error in a lab can now happen digitally in weeks or even days.

Antibodies are among the most exciting targets for these AI systems. Antibodies are proteins the body uses to fight infections, and they're also the basis of many modern therapies. To train an AI model to invent an antibody from scratch, scientists feed it huge amounts of data on existing antibodies, along with details on how they bind to specific targets. The AI models can suggest unique designs that are predicted to bind more efficiently than known antibodies and can even improve the generation of engineered multivalent antibodies to further improve avidity and efficacy. Of course, once the AI model generates these promising ideas, researchers must still test them in cells or animals. But researchers can skip a lot of the guesswork, because they already have reason to believe these designs will do well.

Another fascinating area is the use of AI for mRNA-based therapies. Over the past few years, mRNA has gained attention for its utility in vaccines and other treatments. Designing mRNA that's stable and effective isn't simple. The RNA sequence can affect how well it's translated into proteins and how the body's immune system responds to it. With AI models, researchers can explore thousands of possible mRNA sequences and quickly narrow down which ones are most likely to succeed. They can then bring those sequences into the lab to evaluate performance and efficiency.

A critical complement to AI models is the ability to generate DNA and mRNA on-demand, even overnight for fastest results. Historically, sequence

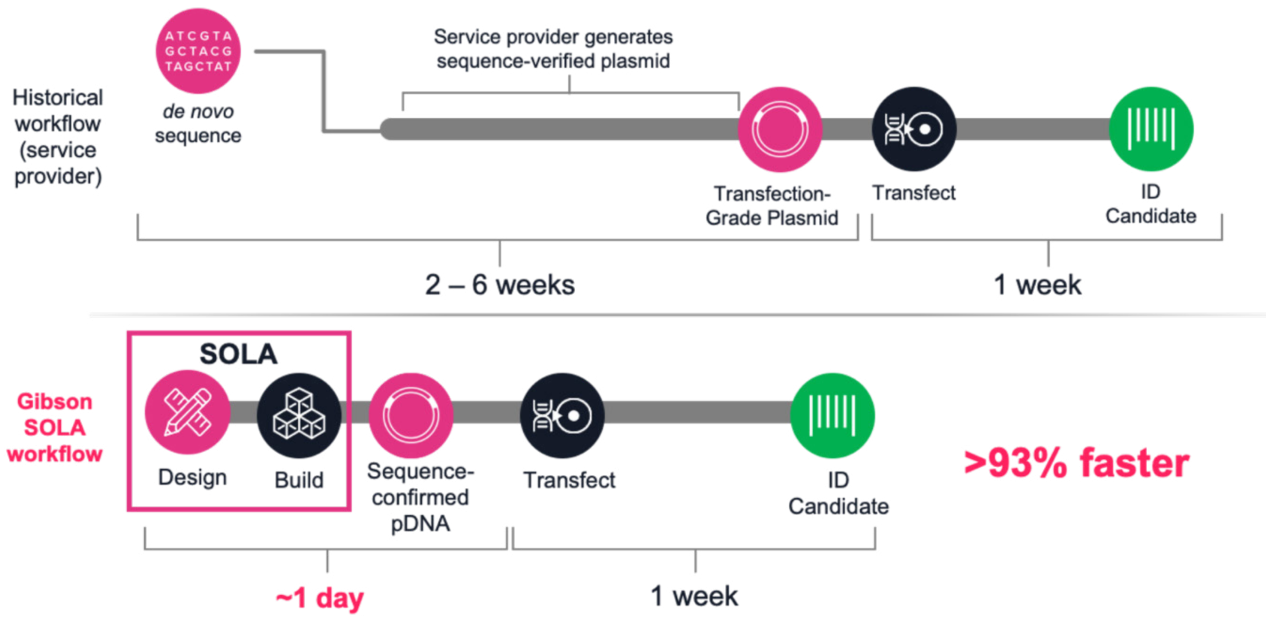


Figure 2. Complete projects more quickly. Eliminate bottlenecks and reduce dependence on external providers by building biology overnight. Historically, service providers can require up to 6 weeks to prepare transfection-grade plasmids (top). The Gibson SOLA Platform generates pDNA in 1 day (bottom).

designs were sent to an external service provider who would need a month or longer to generate the DNA or mRNA. If the sequences fail upon arrival, the process begins again with a new timeline, extending the time to answers. Companies that have their own synthesis capabilities can produce the DNA or mRNA in-house quickly. They proceed from digital sequences to biological data in a much shorter time frame which allows for significantly faster iterative cycles. These companies also retain confidentiality of the designs by preventing them from being added to databases maintained by service providers.

Telesis Bio's novel Gibson SOLA Platform empowers researchers to make the DNA or mRNA they need, when they need it. The platform assembles DNA from universal building blocks on-demand, without the hassle that usually comes with building large or complex sequences. It can also be adjusted for mRNA-related projects, which is a critical advancement for researchers working on vaccines or gene therapies. Instead of waiting for outsourced DNA or mRNA, researchers can analyze lab results within days. The Gibson SOLA Platform is a game changer that removes the barriers and delays that were previously considered unavoidable (figure 2).

Consider a realistic scenario in which a proprietary AI model designs a new antibody that targets a cancer-related protein more effectively than anything else known. Using the Gibson SOLA Platform, that team assembles the template DNA for that antibody on-demand. The DNA is transfected into cells, producing a small batch of the antibody, which is tested for reactivity with cancer cells. If the results look good, the team can move on to more complex studies. If the results are disappointing, the designs are iteratively improved until a candidate is selected. Teams that generate DNA or

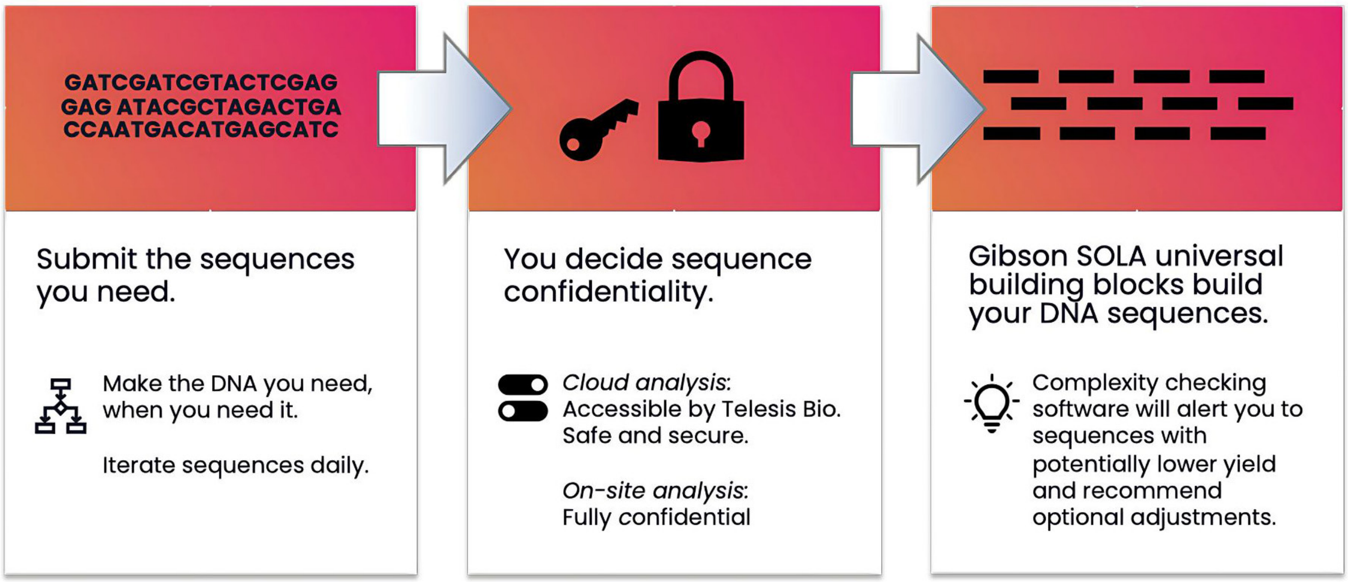


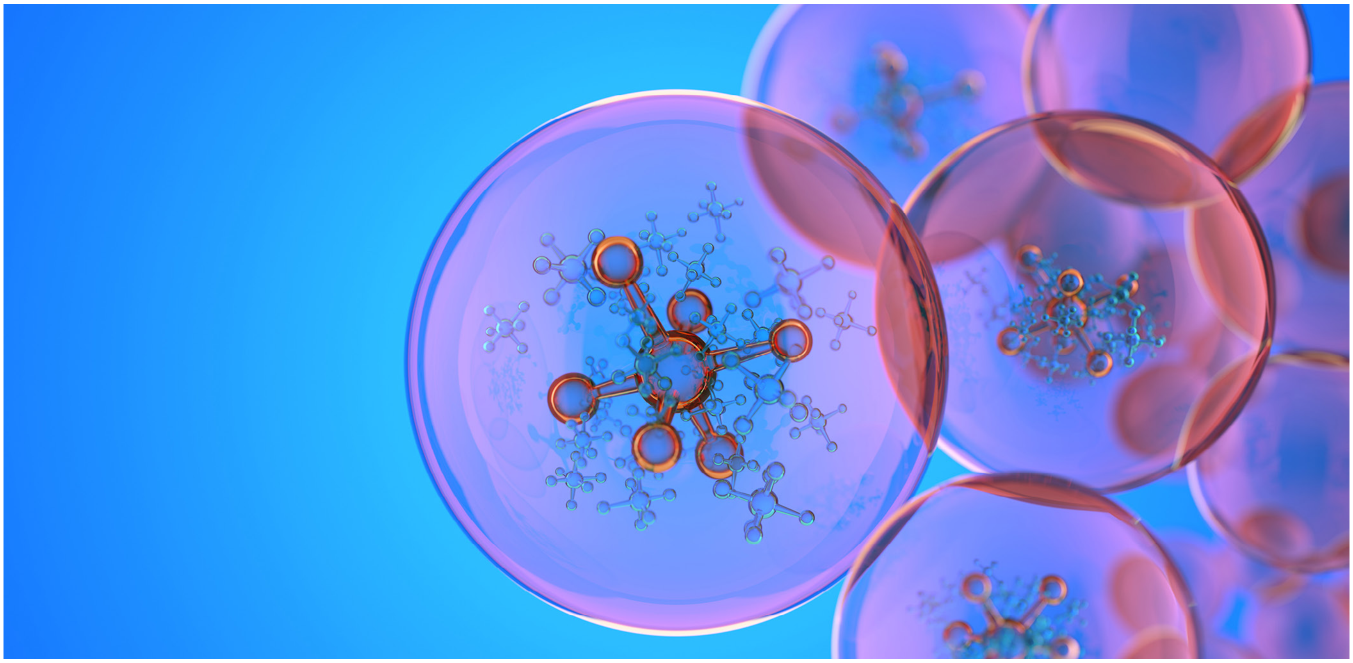
Figure 3. Automated DNA synthesis platform in your lab. The Gibson SOLA Platform is a 'digital to biological converter' that generates DNA and RNA on-demand in your lab. Maintain your sequences as trade secret by generating DNA in-house rather than sending them out to an external service provider for inclusion in their databases.

mRNA in-house using the Gibson SOLA Platform achieve research milestones much faster than teams who outsource DNA or mRNA generation to external service providers (figure 3).

In the future, these techniques could lead to personalized medicine, where AI algorithms generate treatments tailored to an individual's genetic makeup or their specific version of a disease. And because in-house synthesis and testing processes are now more accessible, labs of all sizes could potentially create breakthroughs in medicine, food security, and protein engineering.

Gibson SOLA and the Future of Pharma

MF3d / Getty Images



Drug development is notoriously complex, expensive, and high-risk. One of the biggest determinants of success is the quality of decisions made early in the R&D pipeline—particularly in selecting a promising starting point, or lead candidate.

A growing number of pharmaceutical teams are recognizing that to improve those odds, you need two things: a deep understanding of how biological sequences are constructed, and the ability to combine that knowledge with modern AI tools. The convergence of these two capabilities is reshaping how—and how fast—new therapeutics get developed.

The Importance of Sequence Construction

At its core, every biologic drug originates from a DNA or RNA sequence. These sequences encode proteins, regulate gene expression, or serve as the drug itself (as in mRNA therapeutics). Being able to design and synthesize these sequences accurately is a foundational part of the development process.

Understanding not just *what* the sequence is, but *how it's built*, provides crucial context for predicting biological function, selecting targets, optimizing efficacy, and avoiding failure-prone candidates.

Legacy Limitations: The Bottlenecks of Traditional DNA Synthesis

Historically, DNA and RNA synthesis has relied heavily on external vendors. Researchers submit sequence designs, then wait—often for days or weeks—for a shipment. The process is slow, subject to delays, and introduces risk at the most critical phase: early validation.

This bottleneck reduces the pace of experimentation and lengthens the feedback loop

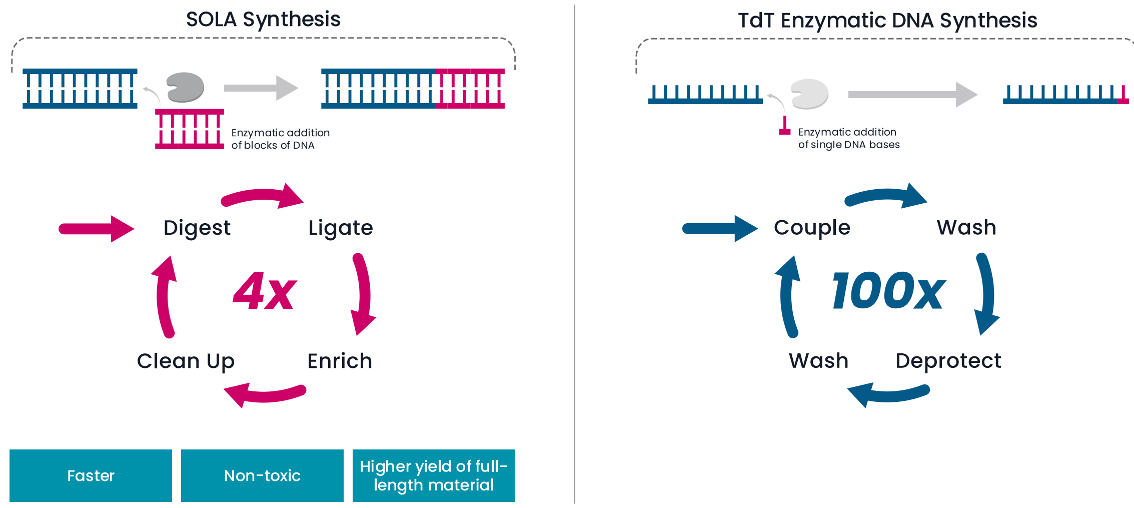


Figure 1 *(left)*. By starting with pre-qualified, short building blocks, the Gibson SOLA Platform enables the on-demand synthesis of 100bp DNA constructs with significantly fewer steps than traditional DNA synthesis methods. This results in faster completion times compared to outsourcing to service providers while leveraging non-toxic reagents to achieve a higher yield of the desired full-length product.

Figure 1 *(right)*. Historical approaches to DNA synthesis build DNA strands one base at a time, introducing complexity and errors with each base. The accumulation of mis-bases leads to a reduction in usable yield and restricts buildable sequences.

for decision-making. Worse, if the sequence has synthesis errors, it might generate false results or require rework.

Accelerating R&D with Gibson SOLA

Telesis Bio's Gibson SOLA Platform addresses this head-on. It enables researchers to synthesize high-fidelity DNA or mRNA *in-house*, overnight.

Rather than constructing long DNA sequences one nucleotide at a time using traditional phosphoramidite chemistry, Gibson SOLA uses enzymatic methods to assemble short, pre-designed oligonucleotides from a universal stock library. This approach significantly lowers error rates and speeds up the process (see figure 1).

The result: researchers gain control over the build process, reduce dependency on third parties, and significantly compress development timelines.

The AI Factor: Data-Driven Discovery

AI models in drug discovery span a range of machine learning approaches, from traditional algorithms like random forests and support vector machines to advanced deep learning frameworks such as convolutional neural networks (CNNs) and transformer-based architectures. Each model type is suited to different data modalities—CNNs might analyze 3D protein structures, while transformers excel at interpreting long DNA or RNA sequences. These models are trained on vast datasets that include genomic annotations, high-throughput screening data, molecular docking simulations, and clinical trial outcomes. Once trained, they can predict which compounds are likely to bind a biological target, how a mutation might affect protein folding, or whether a molecule is likely to pass safety thresholds. Importantly, they also generate probabilistic confidence scores, helping

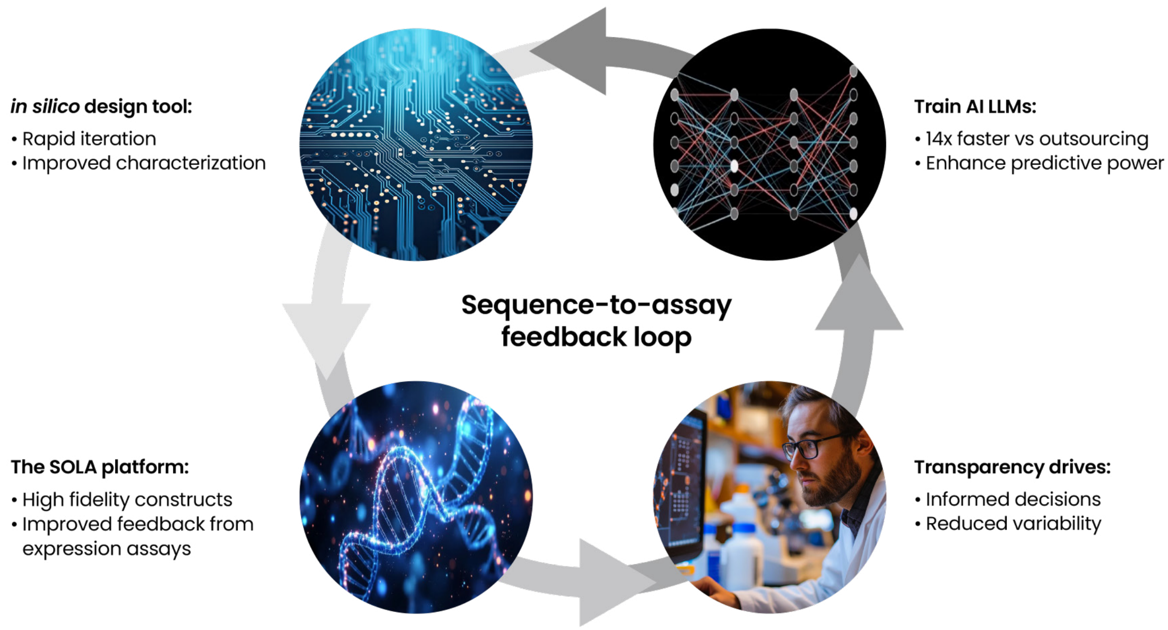


Figure 2. Gibson SOLA enhances productivity in therapeutic development. High fidelity DNA and mRNA generated on the Gibson SOLA Platform reinforces accuracy of AI models, leading to improved predictions for drug targets.

scientists prioritize candidates more objectively. This computational layer not only speeds up decision-making but can uncover previously hidden biological insights that push discovery into entirely new territory.

However, the quality of an AI's output is tightly coupled to the quality of its input. Inaccurate or low-fidelity sequence data undermines predictive performance. Worse, gaps in data caused by external service providers who are unwilling to build the desired sequences result in missed connections and insights. That's why pairing high-quality, real-time synthesis tools like Gibson SOLA with AI workflows is such a powerful combination (figure 3).

When AI Meets Fast, Accurate Synthesis

AI models can identify promising lead candidates based on predictive models. As more data is fed into the models, they improve in accuracy of predicted target candidates and mechanisms. The Gibson SOLA Platform allows those candidates to be synthesized on-demand for accelerated insights. Together, they create a tight feedback loop (figure 3). This cycle, repeated rapidly, allows teams to test more leads, refine ideas faster, and avoid dead-ends early in the process.

Empowered R&D Teams

Bringing synthesis in-house and combining it with AI models accelerates timelines and also changes how teams think and operate.

- **Agility:** Scientists can test hypotheses the same week they're formed.
- **Scalability:** The system supports everything from pilot studies to production.

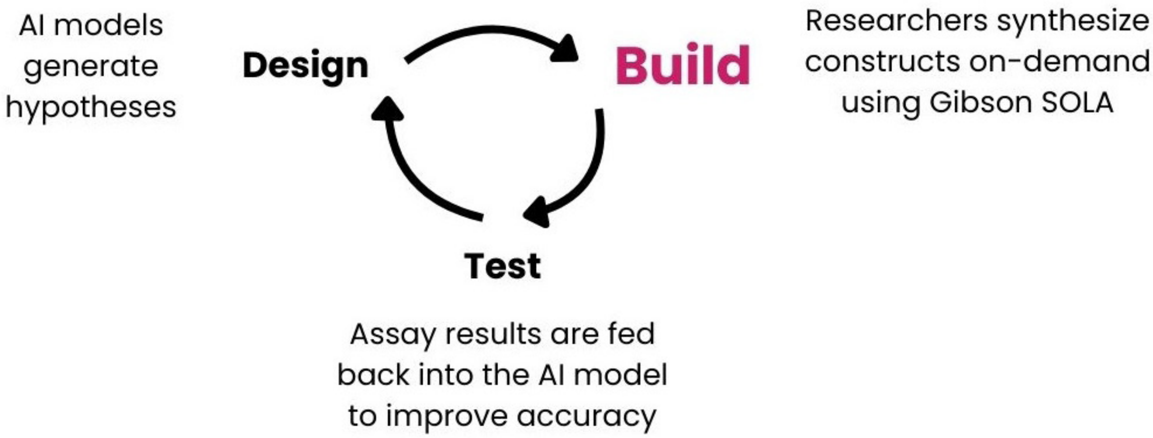


Figure 3. In-House DNA/mRNA Synthesis Accelerates Innovation. The bottleneck in research often shifts from target identification ("design") to molecule synthesis ("build") when DNA or mRNA production is outsourced. By generating DNA and mRNA on-demand within their own facilities, scientists eliminate this delay, transforming a traditional constraint into a catalyst for speed and discovery. In-house synthesis empowers teams to iterate faster, reduce dependency on external timelines, and achieve high-impact breakthroughs with greater efficiency.

- **Innovation:** With fewer logistical delays and more control, teams can take more creative risks.

The result is a pipeline that moves faster, costs less, and is more likely to produce viable therapeutics.

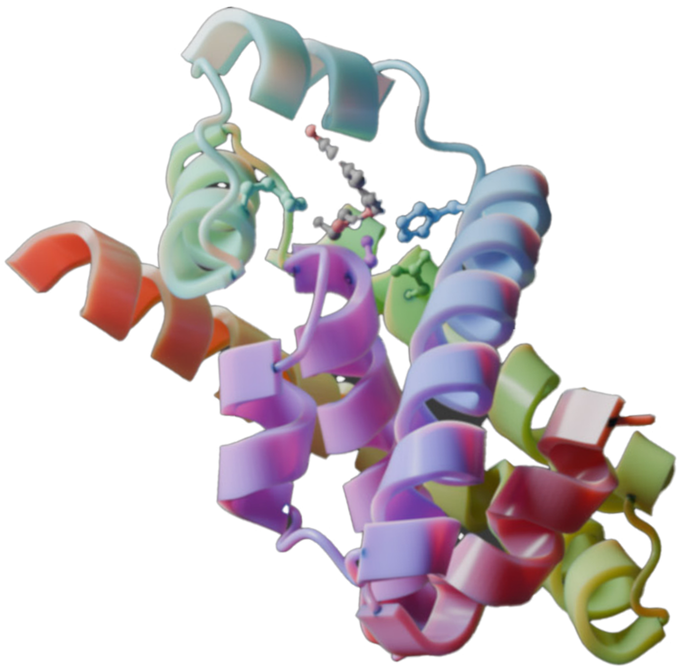
Conclusion

Pharma R&D teams need to make smarter bets earlier in the pipeline. That means starting with better information, better tools, and better feedback loops.

Platforms like Gibson SOLA give researchers control over one of the most foundational aspects of drug development: sequence construction. The cyclic process of build, test, iterate happens more quickly with potentially more valuable insights obtained in each cycle.

In an industry where every delay can cost millions, and every failed candidate sets back progress, that kind of speed and precision isn't just convenient, it's transformative.

AI-Driven Protein Design Produces Enzyme that Mimics Natural Hydrolase Activity



Ian C. Haydon, UW Institute for Protein Design

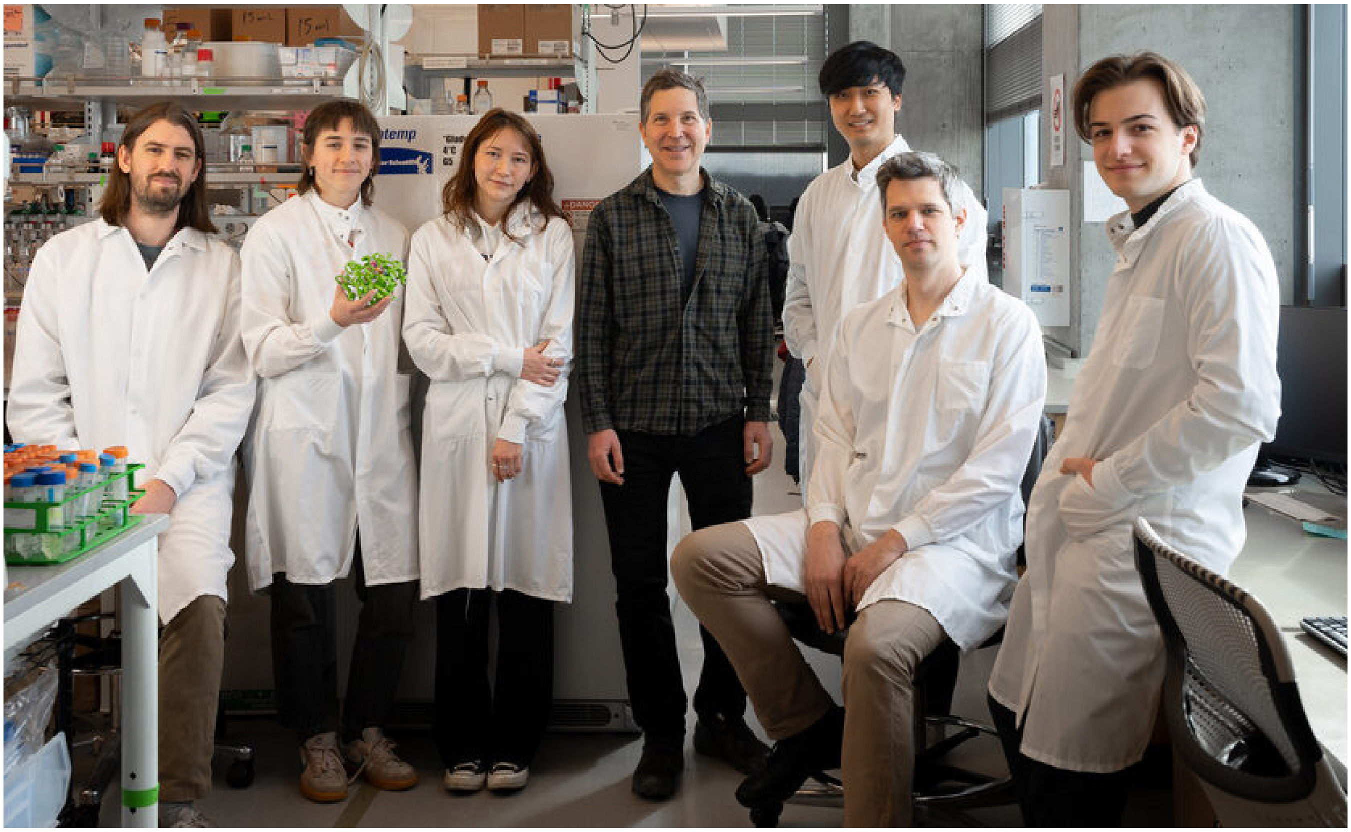
The AI-designed structure of a novel serine hydrolase enzyme. [Samuel J. Pellock, UW Institute for Protein Design]

A new method that successfully designs serine hydrolase enzymes capable of catalyzing ester hydrolysis with high efficiency, demonstrates a computational approach for creating de novo enzymes that catalyze complex, multistep reactions.

The research was published in a paper titled, "Computational design of serine hydrolases," in *Science*. The findings provide a framework for engineering enzymes with intricate active sites, expanding the possibilities for synthetic biocatalysts.

Designing functional enzymes from scratch remains a significant challenge in protein engineering. Traditional approaches often rely on inserting active sites into existing protein scaffolds, which can limit catalytic efficiency due to structural constraints. Until recently, computationally designed enzymes have had reduced efficiencies compared to natural enzymes, however, advances in machine learning and AI open new opportunities for more complex protein design options.

"We thought that designing a serine hydrolase, a well-characterized but fairly complex enzyme, would be an ideal model system to test out the latest AI tools, hopefully helping us uncover general methods that can be applied to the design of other enzymes," co-lead author Anna Lauko, PhD, a post-doctoral researcher in the Baker lab at the University of Washington, shared with GEN.



AI engineered enzyme designers. From left to right: Sam Pellock, PhD, Anna Lauko, PhD, Kiera Sumida, David Baker, PhD, Donghyo Kim, PhD, Indrek Kalvet, PhD, and Seth Woodbury.

Lauko and colleagues developed a novel machine learning model called PLACER (Protein-Ligand Atomistic Conformational Ensemble Reproduction), which predicts the active site conformations of designed enzymes.

"We can configure PLACER to predict the conformations of the active site in each step of the chemical reaction, allowing us to check computationally that the catalytic sidechains actually adopt the active conformation throughout the reaction mechanism," Lauko explained to GEN.

The team began their work using the established generative AI framework, RFdiffusion, to design proteins with complex catalytic sites. They then evaluated the active site conformation with PLACER, which makes predictions by analyzing the protein backbone, amino acid identities, and the chemical structures of bound molecules, allowing for a high degree of precision in enzyme design. With this method, the team successfully designed serine hydrolases that efficiently catalyze ester hydrolysis.

"Enzymes are very challenging to design because the conformation of the active site has to be atomically accurate for catalysis to occur at all," co-lead author, Sam Pellock, PhD, also at the University of Washington, pointed out to *GEN*. "This requires precise design and prediction of not just the overall protein fold but also the conformations of individual sidechains, which is very challenging, especially for enzymes that contain multiple catalytic residues or use multi-step catalytic mechanisms."

The approach eventually resulted in the team generating enzymes with minimal active site specification while maintaining high catalytic efficiency. Initially, the "designed enzymes could only perform the first half of the reaction mechanism before becoming inactivated," Lauko reported to *GEN*.

The team redesigned simplified versions containing three of the five catalytic groups in the natural enzymes first, then pivoted to creating a more complex design containing all five catalytic groups found in the native serine hydrolases.

"We thought it might be easier to accurately design a simpler active site," Lauko explained. "We tried making much more complex designs…and we were overjoyed when we saw that some of these designs could catalyze the whole reaction."

Through experimental characterization of the created serine hydrolases, the team found that the novel enzymes retained folds unique from natural serine hydrolases with high catalytic efficiencies. These novel enzymes contained structures that closely matched the computational design models, as confirmed by crystal structure analysis. "We were really excited when we saw how well it matched our predicted structure," Lauko shared.

By screening designed enzymes with active site preorganization, the team was able to preselect novel enzyme designs that had a higher likelihood of success in reality. The study's findings highlight the potential of computational tools in enzyme engineering, particularly for creating biocatalysts with industrial and pharmaceutical applications. By integrating AI tools into their experimental methods, the researchers established an adaptable strategy for designing enzymes with tailored functions, with high utility in synthetic biology applications.

"We hope that the concepts and methods we used in this paper will be applicable to designing new enzymes in the future that act on important substrates or perform new chemistry," Pellock concluded. ■

Solanum Pan-Genome Unveils Paralogs' Role in Genome Engineered Crops

Andrii Yalanskyi / 1267333699 / iStock / Getty Images Plus

Eggplant plantations grow in the field on a sunny day.

Advancements in genomics, next-generation sequencing, and genome editing are driving forward a new era of crop breeding. About 75% of the world's food comes from 12 plants. However, scientists estimate up to 30,000 species are edible. One opportunity in broadening our food supply lies in exchanging genotype-to-phenotype knowledge between globally and locally cultivated crops. However, many genetic variants are species-specific. And methods of selecting for advantageous traits can produce different results in related species.

"There's a lot of wonderful food crops out there," said Zachary Lippman, PhD, Cold Spring Harbor Laboratory (CSHL) professor & HHMI investigator. "How many of them have not received the attention they would benefit from, compared to 'major' crops?"

Now, CSHL researchers and colleagues around the globe have established a pan-genome of the crop-rich genus Solanum. The team sequenced dozens of complete genomes for the plant genus that includes tomatoes, potatoes, and eggplants. The new, high-quality pan-genome was then used to map the genes behind specific traits of agricultural significance across the genus, and target those genes to create desirable mutations.

Yuuji / Getty Images

This work is published in *Nature* in the paper, "Solanum pan-genetics reveals paralogues as contingencies in crop engineering."

The team's research reveals the importance of understanding the evolution of paralog genes in predicting genome editing outcomes. How paralogs relate to physical changes across species has not been deeply studied—until now. And, in this study, the biggest breakthroughs came from the African eggplant: a tomato relative indigenous to the sub-Saharan region, African eggplant varies highly in fruit shape, color, and size.

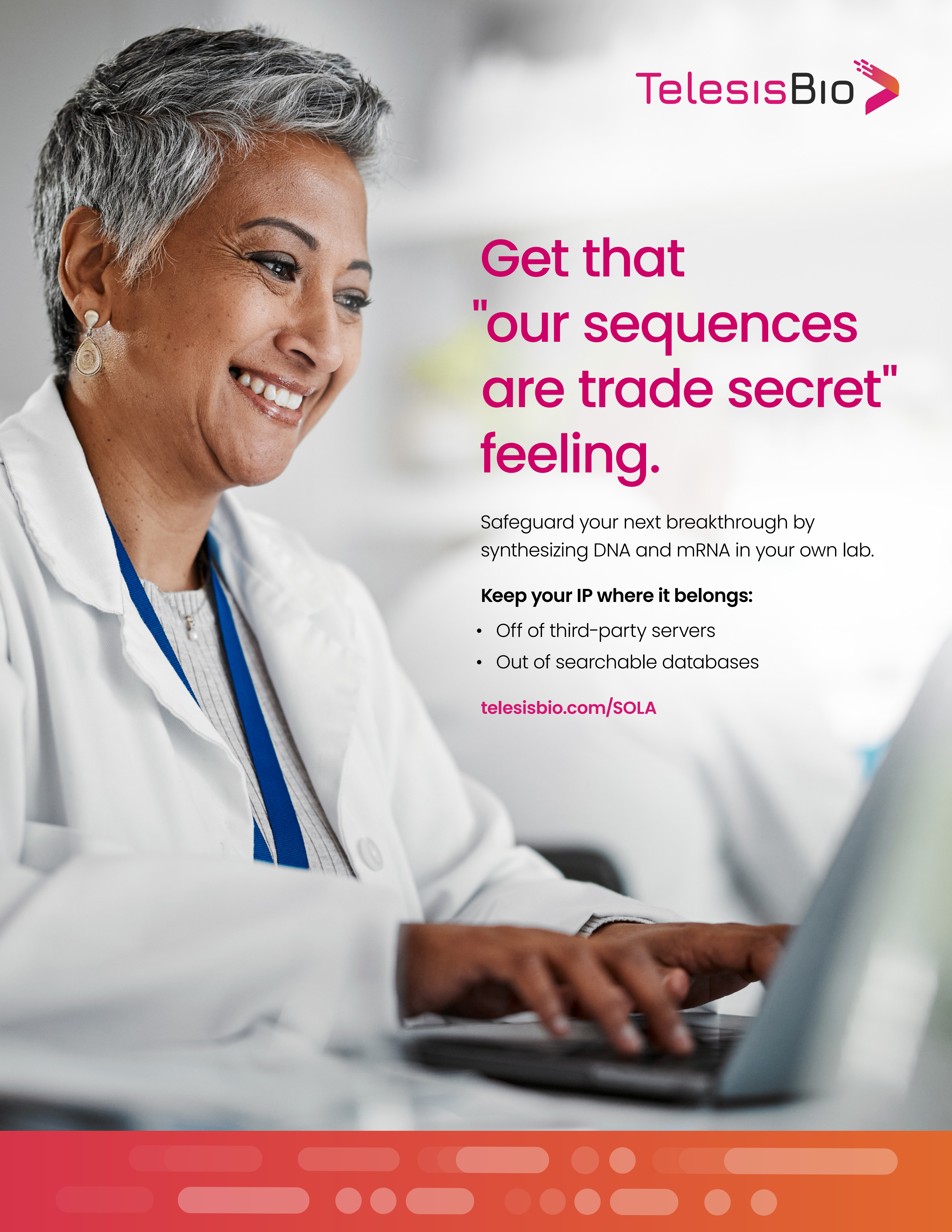
The authors wrote, "Despite broad conservation of gene macrosynteny among chromosome-scale references for 22 species, including 13 indigenous crops, thousands of gene duplications, particularly within key domestication gene families, exhibited dynamic trajectories in sequence, expression, and function. By augmenting our pan-genome with African eggplant cultivars and applying quantitative genetics and genome editing, we dissected an intricate history of paralogue evolution affecting fruit size."

Lippman and longtime collaborator Michael Schatz, PhD, professor of computational biology and oncology at Johns Hopkins University, turned to a breeder in Uganda to exchange ideas and expertise. Mapping tens of thousands of paralogs, the team identified a previously unknown gene in African eggplant that affects fruit size. The paralog has the same function in tomatoes. The researchers discovered they could influence tomato size by editing it.

"Reciprocal exchange between indigenous and major crops creates new, predictable paths for better breeding," said Benoit. "This is key to boost the diversity and resilience of the food system."

The findings, the authors suggest, demonstrate that "paralogue diversifications over short timescales are underexplored contingencies in trait evolvability. Exposing and navigating these contingencies is crucial for translating genotype-to-phenotype relationships across species."

"Crop diversity benefits nutrition, choice, and health," Lippman added. "Determining how related paralogs function across species could help improve crop yields, flowering times, and food selection. In other words, it's a win-win-win for scientists, farmers, and consumers everywhere." ■


TelesisBio
Get that "our sequences are trade secret" feeling.
Safeguard your next breakthrough by synthesizing DNA and mRNA in your own lab.
Keep your IP where it belongs:
• Off of third-party servers
• Out of searchable databases
telesisbio.com/SOLA