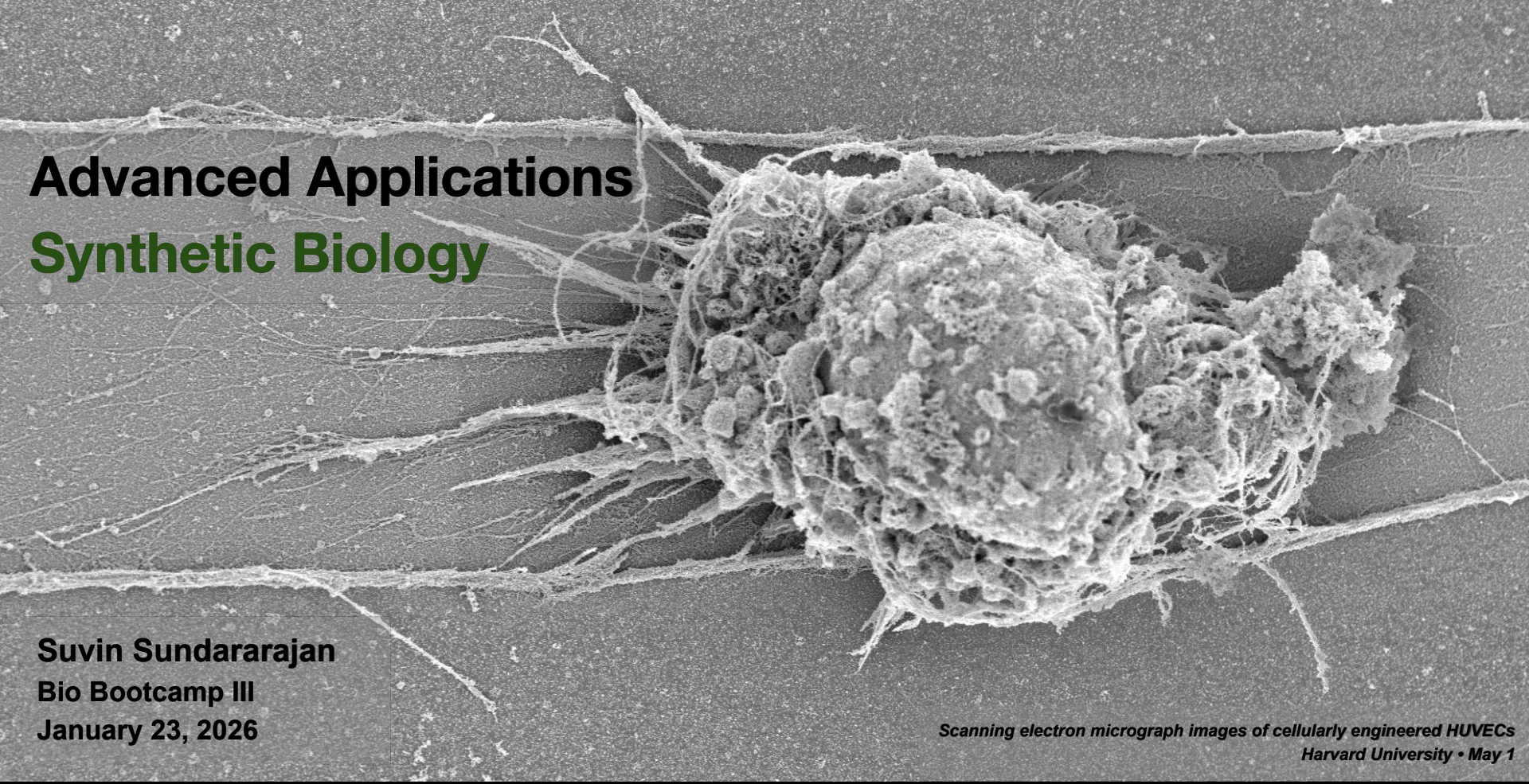
Advanced Applications

Synthetic Biology

Suvin Sundararajan

Bio Bootcamp III

January 23, 2026



Scanning electron micrograph images of cellularly engineered HUVECs
Harvard University • May 1

Agenda

Agenda

Part 1: Developments in the Field (10 minutes)

Agenda

Part 1: Developments in the Field (10 minutes)

Part 2: Role of Automation and AI (10 minutes)

Agenda

Part 1: Developments in the Field (10 minutes)

Part 2: Role of Automation and AI (10 minutes)

Part 3: Live Demonstrations (15 min)

Agenda

Part 1: Developments in the Field (10 minutes)

Part 2: Role of Automation and AI (10 minutes)

Part 3: Live Demonstrations (15 min)

- *RF Diffusion, AlphaFold*

Agenda

Part 1: Developments in the Field (10 minutes)

Part 2: Role of Automation and AI (10 minutes)

Part 3: Live Demonstrations (15 min)

- *RF Diffusion, AlphaFold*
- *Lab Automation* *(Time Pending)*

Agenda

Part 1: Developments in the Field (10 minutes)

Part 2: Role of Automation and AI (10 minutes)

Part 3: Live Demonstrations (15 min)

- *RF Diffusion, AlphaFold*
- *Lab Automation* *(Time Pending)*

Part 4: Case Study (15 min)

Part 1: Developments in the Field

How can we improve experimental discovery to advance society and improve wellbeing?

1. Protein Prediction and Analysis Milestones

1. Protein Prediction and Analysis Milestones

- **2021: Structure prediction becomes usable (AlphaFold2)**
  - Accurate protein structures at scale shift biology to model-informed decisions.

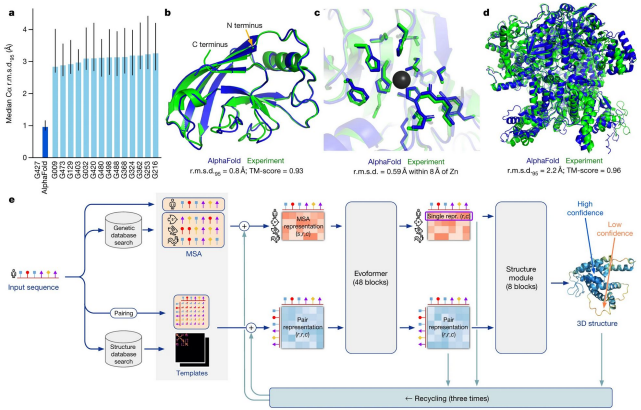
Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

1.96m Accesses | **20k** Citations | **4010** Altmetric | Metrics



1. Protein Prediction and Analysis Milestones

- **2021: Structure prediction becomes usable (AlphaFold2)**
  - Accurate protein structures at scale shift biology to model-informed decisions.
- **2022: Protein language models scale the "sequence → insight" loop (ESMFold / ESM-2)**
  - Faster, cheaper inference and massive sequence-space coverage enable large metagenomic structure atlases and numerous hypotheses.

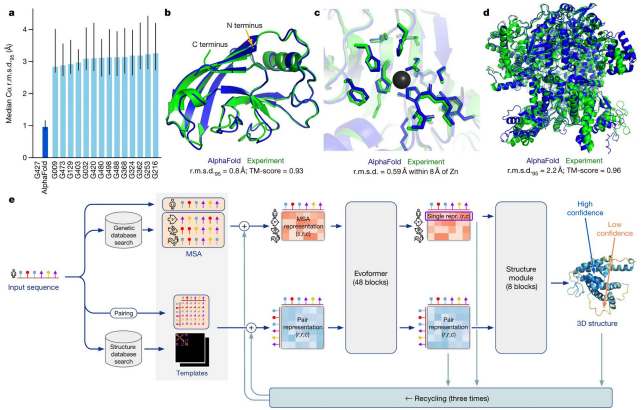
Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

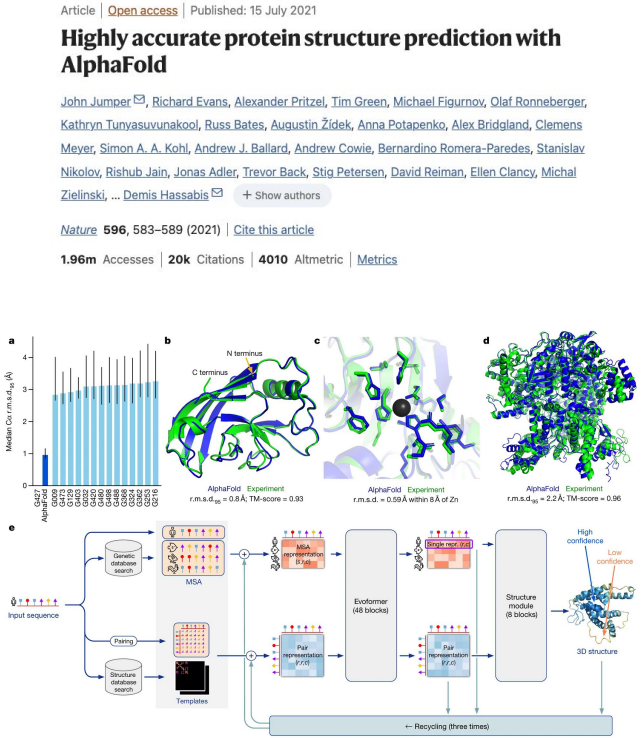
Nature **596**, 583–589 (2021) | Cite this article

1.96m Accesses | **20k** Citations | **4010** Altmetric | Metrics



1. Protein Prediction and Analysis Milestones

- **2021: Structure prediction becomes usable (AlphaFold2)**
  - Accurate protein structures at scale shift biology to model-informed decisions.
- **2022: Protein language models scale the "sequence → insight" loop (ESMFold / ESM-2)**
  - Faster, cheaper inference and massive sequence-space coverage enable large metagenomic structure atlases and numerous hypotheses.
- **2023: Generative design arrives (RFdiffusion)**
  - Models begin proposing new sequences/molecules, not just predicting nature.



Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

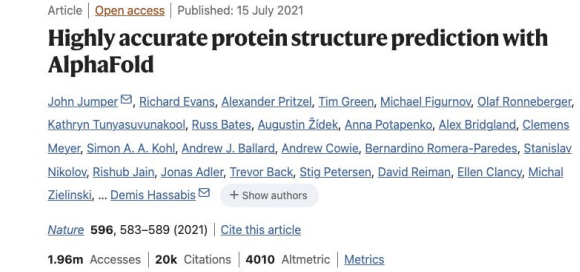
John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

1.96m Accesses | **20k** Citations | **4010** Altmetric | Metrics

1. Protein Prediction and Analysis Milestones

- **2021: Structure prediction becomes usable (AlphaFold2)**
  - Accurate protein structures at scale shift biology to model-informed decisions.
- **2022: Protein language models scale the "sequence → insight" loop (ESMFold / ESM-2)**
  - Faster, cheaper inference and massive sequence-space coverage enable large metagenomic structure atlases and numerous hypotheses.
- **2023: Generative design arrives (RFdiffusion)**
  - Models begin proposing new sequences/molecules, not just predicting nature.
- **2024: Interaction-centric modeling (AlphaFold 3)**
  - Focus shifts from "folded shape" to binding interactions (complexes, ligands, RNA/DNA).



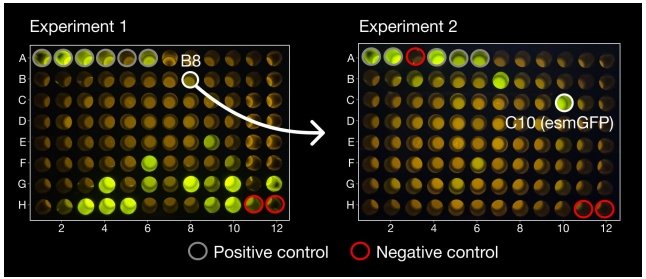
Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

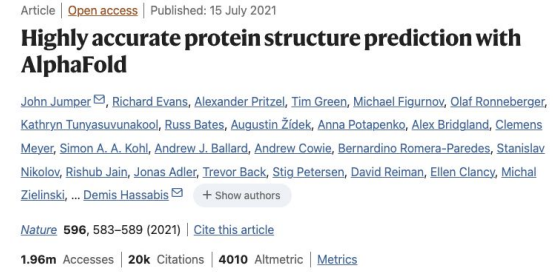
Nature **596**, 583–589 (2021) | Cite this article

1.96m Accesses | **20k** Citations | **4010** Altmetric | Metrics



1. Protein Prediction and Analysis Milestones

- **2021: Structure prediction becomes usable (AlphaFold2)**
  - Accurate protein structures at scale shift biology to model-informed decisions.
- **2022: Protein language models scale the "sequence → insight" loop (ESMFold / ESM-2)**
  - Faster, cheaper inference and massive sequence-space coverage enable large metagenomic structure atlases and numerous hypotheses.
- **2023: Generative design arrives (RFdiffusion)**
  - Models begin proposing new sequences/molecules, not just predicting nature.
- **2024: Interaction-centric modeling (AlphaFold 3)**
  - Focus shifts from "folded shape" to binding interactions (complexes, ligands, RNA/DNA).
- **2025: Truly novel functional proteins and target-conditioned design emerge (ESM3 → esmGFP)**
  - Foundation/diffusion models generate novel, functional proteins, making target-conditioned design practical and moving protein engineering toward an iterative workflow.



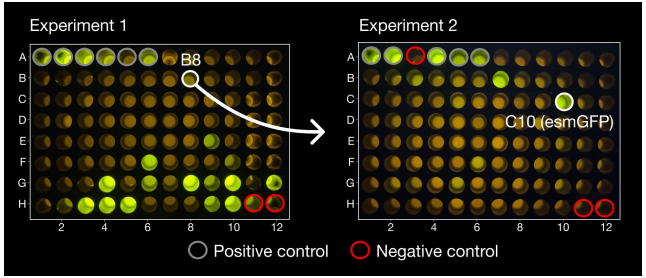
Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis + Show authors

Nature **596**, 583–589 (2021) | Cite this article

1.96m Accesses | **20k** Citations | **4010** Altmetric | Metrics



2. GLP-1 (Ozempic) Launch

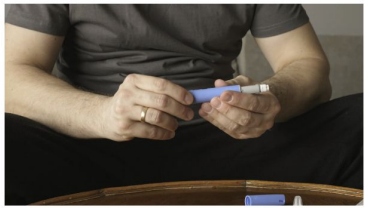
2. GLP-1 (Ozempic) Launch

2. GLP-1 (Ozempic) Launch

What Is Ozempic and Why Is It Getting So Much Attention?

More people are turning to a diabetes medication to induce weight loss — but experts say it's not a miracle drug.

Share full article 775



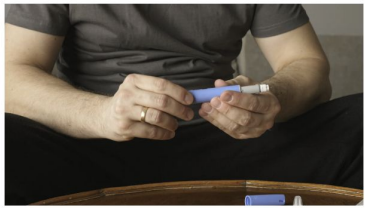
Getty Images

2. GLP-1 (Ozempic) Launch

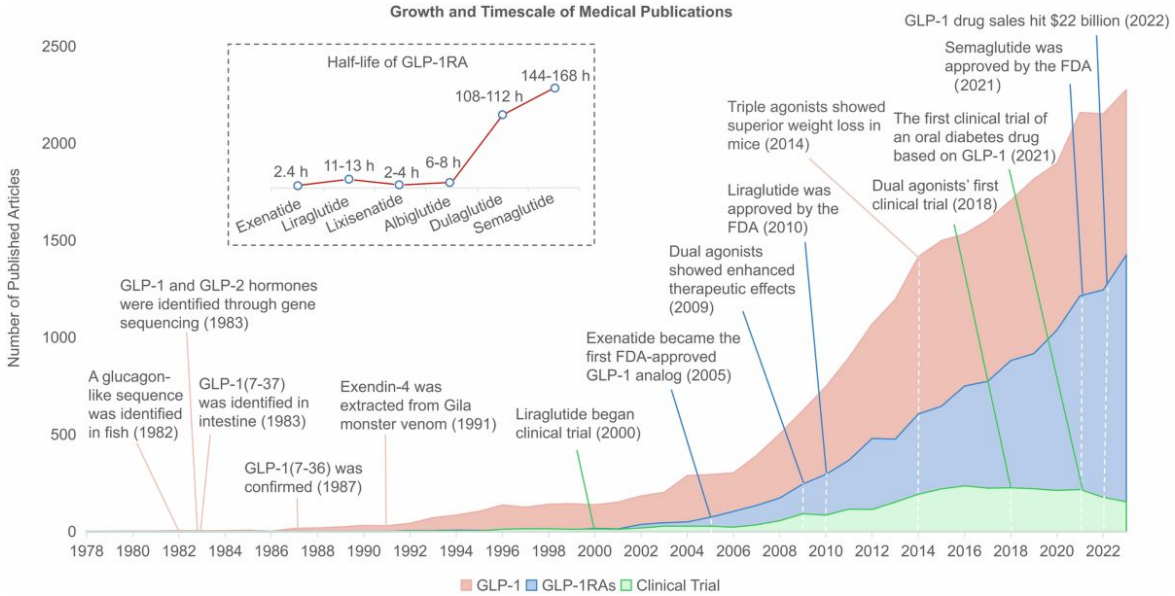
What Is Ozempic and Why Is It Getting So Much Attention?

More people are turning to a diabetes medication to induce weight loss — but experts say it's not a miracle drug.

Share full article 775



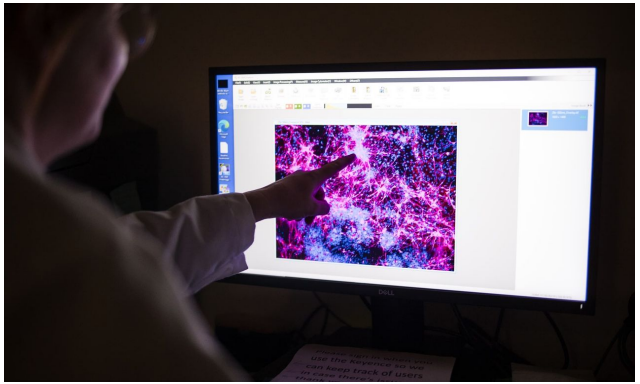
Getty Images



3. Personalized CRISPR Therapy

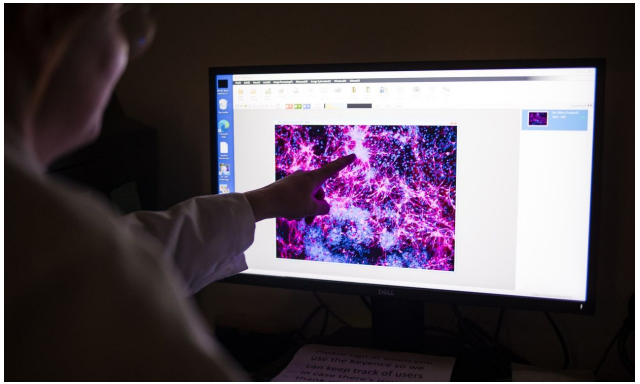
3. Personalized CRISPR Therapy

- **Traditional Model**
  - One editor applied to one disease program over many years, treating many patients.



3. Personalized CRISPR Therapy

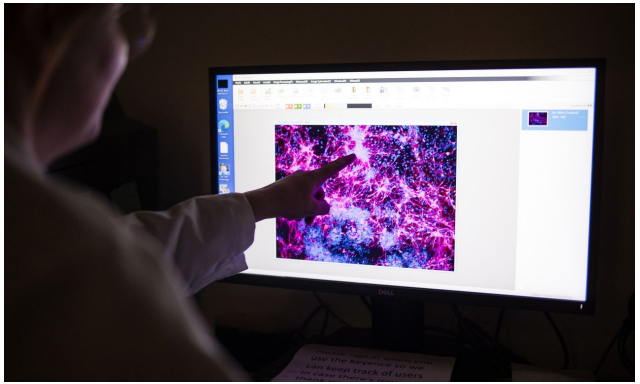
- **Traditional Model**
  - One editor applied to one disease program over many years, treating many patients.
- **New Frontier (N-of-1)**
  - Patient-specific editing for ultra-rare mutations (e.g., bespoke therapy for a child by CHOP/Penn).



https://www.chop.edu/news/worlds-first-patient-treated-personalized-crispr-gene-editing-therapy-childrens-hospital

3. Personalized CRISPR Therapy

- **Traditional Model**
  - One editor applied to one disease program over many years, treating many patients.
- **New Frontier (N-of-1)**
  - Patient-specific editing for ultra-rare mutations (e.g., bespoke therapy for a child by CHOP/Penn).
- **Shift in Value**
  - The "product" is now the capacity to rapidly generate safe, specific edits. Value is placed on speed and correctness, not just novelty.

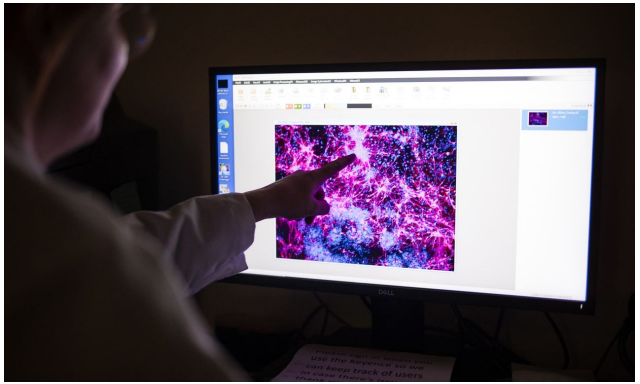


https://www.chop.edu/news/worlds-first-patient-treated-personalized-crispr-gene-editing-therapy-childrens-hospital

3. Personalized CRISPR Therapy

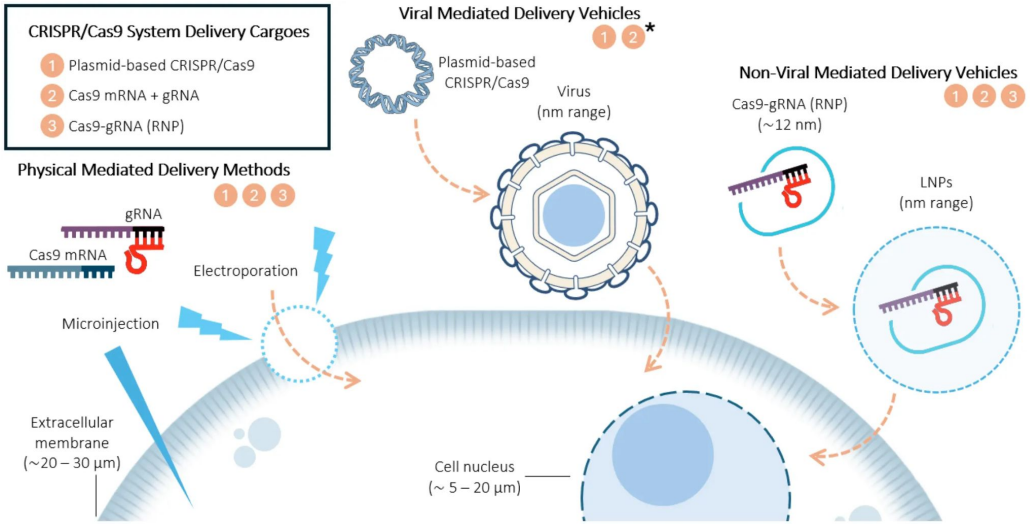
- **Traditional Model**
  - One editor applied to one disease program over many years, treating many patients.
- **New Frontier (N-of-1)**
  - Patient-specific editing for ultra-rare mutations (e.g., bespoke therapy for a child by CHOP/Penn).
- **Shift in Value**
  - The "product" is now the capacity to rapidly generate safe, specific edits. Value is placed on speed and correctness, not just novelty.

Gene editing is becoming a more repeatable process, reframing medicine as an engineering discipline where biology is increasingly programmable.



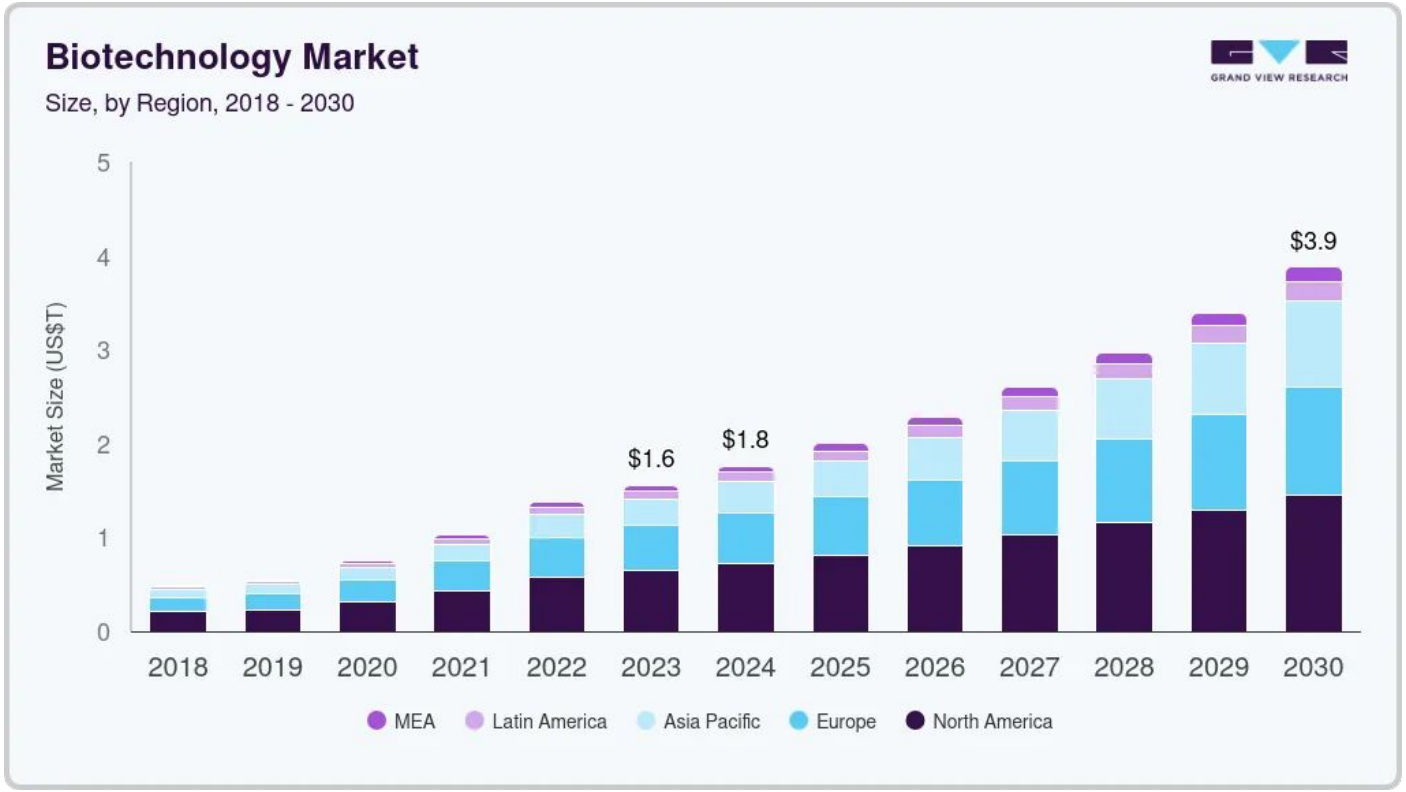
https://www.chop.edu/news/worlds-first-patient-treated-personalized-crispr-gene-editing-therapy-childrens-hospital

3. Personalized CRISPR Therapy - Challenges



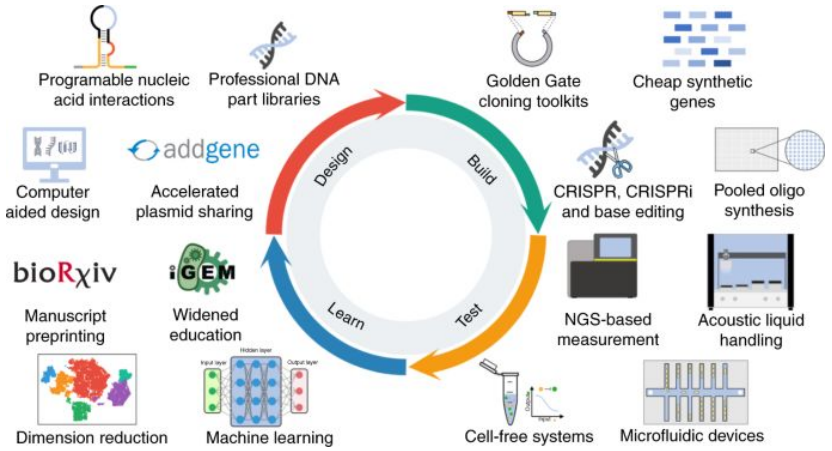
https://www.synthego.com/blog/delivery-crispr-cas9/

4. Market Growth



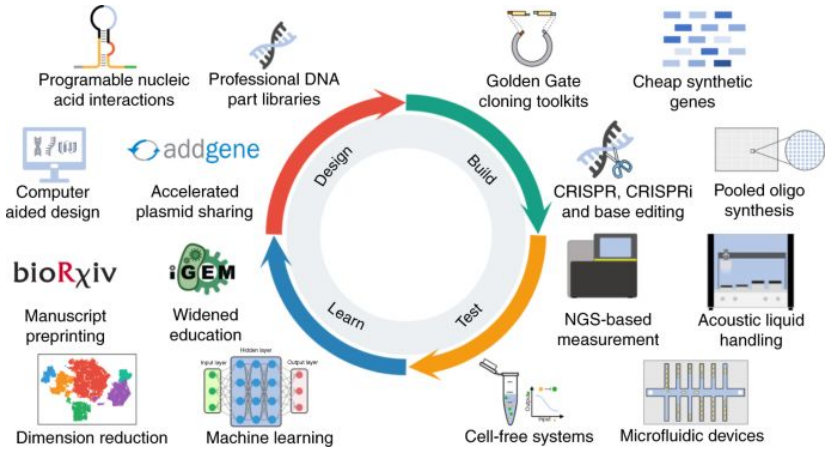
Future Developments

- **Biology is shifting from running single experiments to coordinated campaigns of experiments, Bayesian optimization, and active learning.**



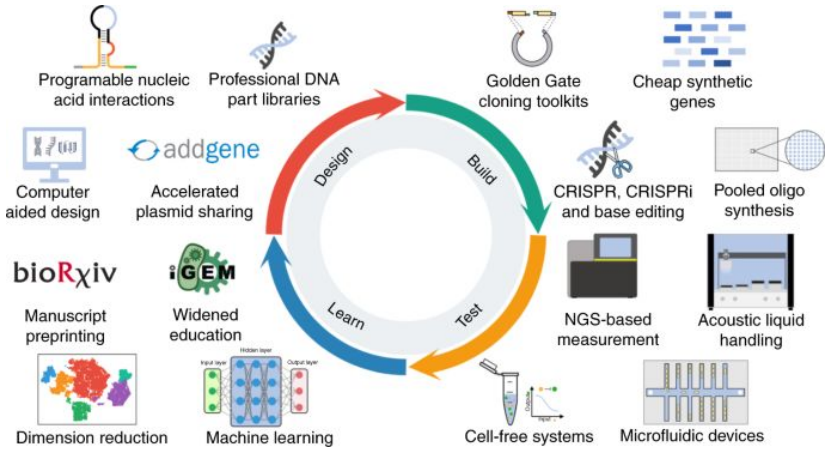
Future Developments

- Biology is shifting from running single experiments to coordinated campaigns of experiments, Bayesian optimization, and active learning.
- **This shift is happening now because models are useful, lab robots are accessible, and assays are high-throughput, so iteration speed has become the limiting factor.**



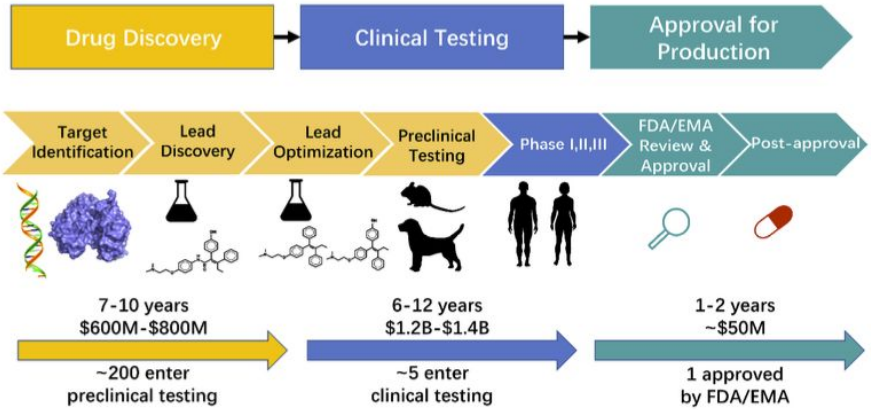
Future Developments

- Biology is shifting from running single experiments to coordinated campaigns of experiments, Bayesian optimization, and active learning.
- This shift is happening now because models are useful, lab robots are accessible, and assays are high-throughput, so iteration speed has become the limiting factor.
- **The core workflow is a repeatable loop: design the next set of conditions, execute them, measure outcomes, learn from the data, and choose the next run—on a daily or weekly cadence rather than quarterly.**



Future Developments

- Biology is shifting from running single experiments to coordinated campaigns of experiments, Bayesian optimization, and active learning.
- This shift is happening now because models are useful, lab robots are accessible, and assays are high-throughput, so iteration speed has become the limiting factor.
- **The core workflow is a repeatable loop: design the next set of conditions, execute them, measure outcomes, learn from the data, and choose the next run—on a daily or weekly cadence rather than quarterly.**



Part 2: Role of Automation and AI

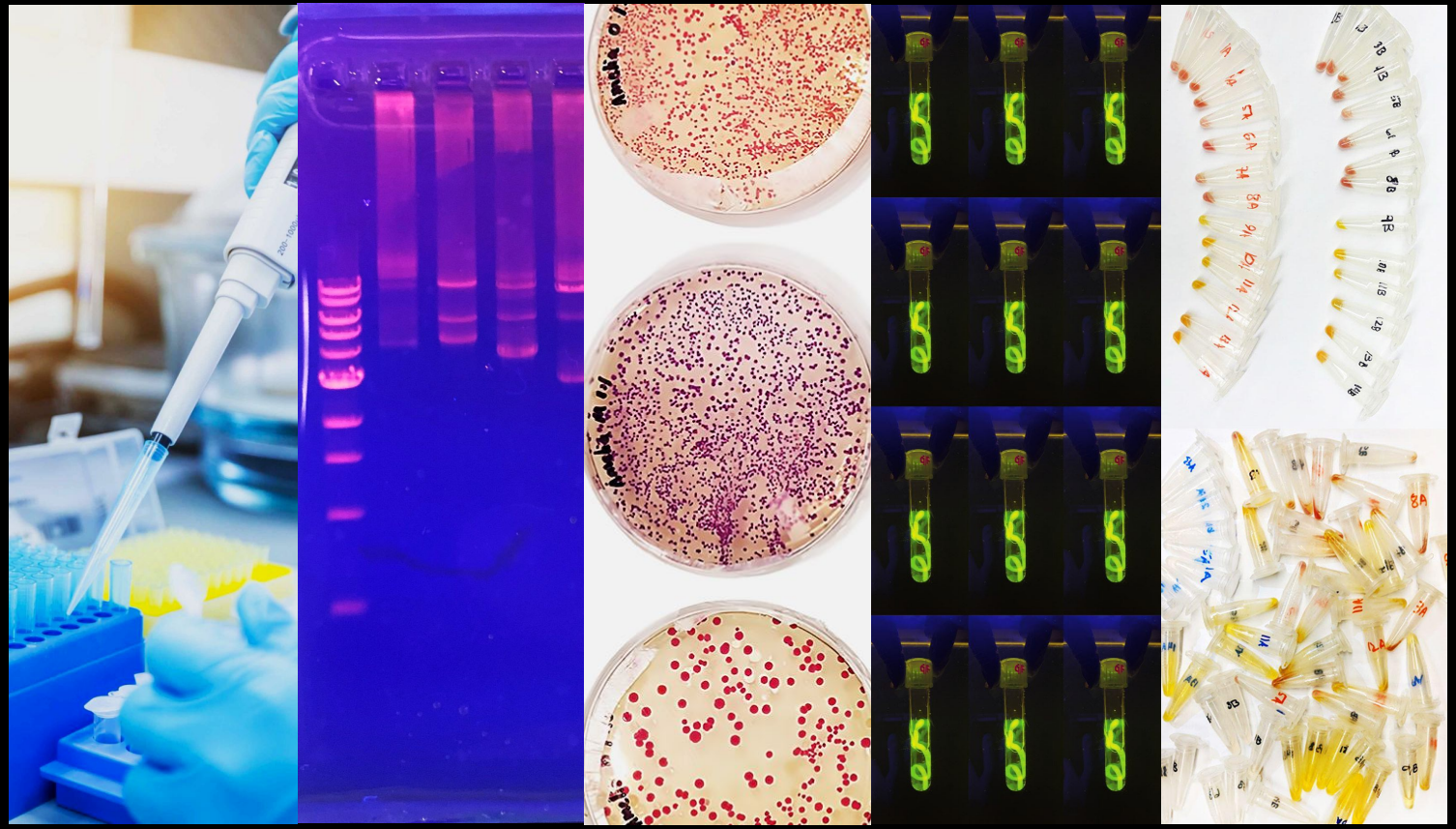
What is the most common tool for synthetic biologists?

What is the most common tool for synthetic biologists?

How can we improve this?

=

Sneak Peak



Sneak Peak

How to Grow (Almost) Anything 2025
Lab Protocol: Bioproduction of Beta-Carotene and Lycopene

Protocol | Part 1: Overnight Cultures

Time Estimate: 30 Minutes, 24 Hour Incubation

Media, Equipment and Consumables

- LB with pre-added antibiotic (chloramphenicol)
- Pipette set, serological pipettes and pipette tips
- Incubation room
- Culture tubes

You'll be testing two genes, two media combinations, one input, and two temperature conditions shown above. You'll need to set up **16 unique overnight cultures** in given culture tubes, alongside duplicates for each, and two cultures with just media (so a total of 34 cultures in total)

Condition #	Plasmid	Culture temp	Growth Medium
1 and 2	pAC-LYC	30°C, 37°C	Luria Broth (LB)
3 and 4	pAC-LYC	30°C, 37°C	LB + fructose
5 and 6	pAC-LYC	30°C, 37°C	2YT
7 and 8	pAC-LYC	30°C, 37°C	2YT + fructose
9 and 10	pAC-BETA	30°C, 37°C	LB
11 and 12	pAC-BETA	30°C, 37°C	LB + fructose
13 and 14	pAC-BETA	30°C, 37°C	2YT
15 and 16	pAC-BETA	30°C, 37°C	2YT + fructose

1. For each conditions above, set up overnight cultures that contain the following
    a. 5 mL of the specified media (already supplemented with antibiotic)
    b. 10 uL of *E. coli* from the starter culture with the specified plasmid
2. Grow cultures for 24 hours in the circular roller drum in the appropriate warm room

How to Grow (Almost) Anything 2025
Lab Protocol: The Chromophore Color Cloning Quest

Edited December 29, 2024

Protocol | Part 1: PCR

Time Estimate: 30 Minutes, 24 Hour Incubation

Media, Equipment and Consumables

- Ice bucket
- Phusion High-Fidelity PCR Master Mix
- Primers (working stocks, at 10 uM)
- UltraPure Water
- PCR tubes
- Thermocycler
- P20 pipette and 10uL tips
- P200 pipette and 200uL tips

Setup the following PCR reactions:

Universal DNA Fragment (Univ Fwd and Univ Primers)				
Reaction	**Reagent**	**Stock Conc.**	**Desired Conc.**	**Volume**
All	Mini Prepped mUAV	Check Stock	Typically 10 ng	X uL
All	Univ Forward Primer	10 uM	0.5 uM	1.25 uL
All	Univ Reverse Primer	10 uM	0.5 uM	1.25 uL
All	Phusion HF PCR Mix	2X	1x	12.5 uL
All	Nuclease-free water	n/a	n/a	10 - X uL
All	Total Volume			25 µL

Sneak Peak

How to Grow (Almost) Anything 2025
Lab Protocol: Bioproduction of Beta-Carotene and Lycopene

Protocol | Part 1: Overnight Cultures

Time Estimate: 30 Minutes, 24 Hour Incubation

Media, Equipment and Consumables

- LB with pre-added antibiotic (chloramphenicol)
- Pipette set, serological pipettes and pipette tips
- Incubation room
- Culture tubes

You'll be testing two genes, two media combinations, one input, and two temperature conditions shown above. You'll need to set up **16 unique overnight cultures** in given culture tubes, alongside duplicates for each, and two cultures with just media (so a total of 34 cultures in total)

Condition #	Plasmid	Culture temp	Growth Medium
1 and 2	pAC-LYC	30°C, 37°C	Luria Broth (LB)
3 and 4	pAC-LYC	30°C, 37°C	LB + fructose
5 and 6	pAC-LYC	30°C, 37°C	2YT
7 and 8	pAC-LYC	30°C, 37°C	2YT + fructose
9 and 10	pAC-BETA	30°C, 37°C	LB
11 and 12	pAC-BETA	30°C, 37°C	LB + fructose
13 and 14	pAC-BETA	30°C, 37°C	2YT
15 and 16	pAC-BETA	30°C, 37°C	2YT + fructose

1. For each conditions above, set up overnight cultures that contain the following
   a. 5 mL of the specified media (already supplemented with antibiotic)
   b. 10 uL of *E. coli* from the starter culture with the specified plasmid
2. Grow cultures for 24 hours in the circular roller drum in the appropriate warm room

How to Grow (Almost) Anything 2025
Lab Protocol: The Chromophore Color Cloning Quest

Edited December 29, 2024

Protocol | Part 1: PCR

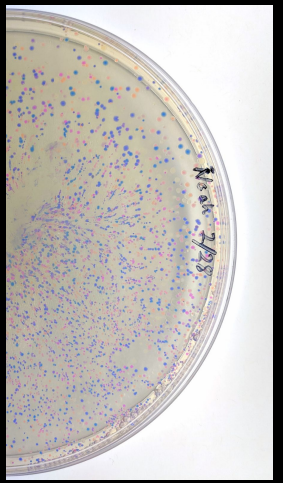
Time Estimate: 30 Minutes, 24 Hour Incubation

Media, Equipment and Consumables

- Ice bucket
- Phusion High-Fidelity PCR Master Mix
- Primers (working stocks, at 10 uM)
- UltraPure Water
- PCR tubes
- Thermocycler
- P20 pipette and 10uL tips
- P200 pipette and 200uL tips

Setup the following PCR reactions:

Universal DNA Fragment (Univ Fwd and Univ Primers)				
Reaction	**Reagent**	**Stock Conc.**	**Desired Conc.**	**Volume**
All	Mini Prepped mUAV	Check Stock	Typically 10 ng	X uL
All	Univ Forward Primer	10 uM	0.5 uM	1.25 uL
All	Univ Reverse Primer	10 uM	0.5 uM	1.25 uL
All	Phusion HF PCR Mix	2X	1x	12.5 uL
All	Nuclease-free water	n/a	n/a	10 - X uL
All	Total Volume			25 µL



Sneak Peak

How to Grow (Almost) Anything 2025
Lab Protocol: Bioproduction of Beta-Carotene and Lycopene

Protocol | Part 1: Overnight Cultures

Time Estimate: 30 Minutes, 24 Hour Incubation

Media, Equipment and Consumables

- LB with pre-added antibiotic (chloramphenicol)
- Pipette set, serological pipettes and pipette tips
- Incubation room
- Culture tubes

You'll be testing two genes, two media combinations, one input, and two temperature conditions shown above. You'll need to set up **16 unique overnight cultures** in given culture tubes, alongside duplicates for each, and two cultures with just media (so a total of 34 cultures in total)

Condition #	Plasmid	Culture temp	Growth Medium
1 and 2	pAC-LYC	30°C, 37°C	Luria Broth (LB)
3 and 4	pAC-LYC	30°C, 37°C	LB + fructose
5 and 6	pAC-LYC	30°C, 37°C	2YT
7 and 8	pAC-LYC	30°C, 37°C	2YT + fructose
9 and 10	pAC-BETA	30°C, 37°C	LB
11 and 12	pAC-BETA	30°C, 37°C	LB + fructose
13 and 14	pAC-BETA	30°C, 37°C	2YT
15 and 16	pAC-BETA	30°C, 37°C	2YT + fructose

1. For each conditions above, set up overnight cultures that contain the following
   a. 5 mL of the specified media (already supplemented with antibiotic)
   b. 10 uL of *E. coli* from the starter culture with the specified plasmid
2. Grow cultures for 24 hours in the circular roller drum in the appropriate warm room

How to Grow (Almost) Anything 2025
Lab Protocol: The Chromophore Color Cloning Quest

Edited December 29, 2024

Protocol | Part 1: PCR

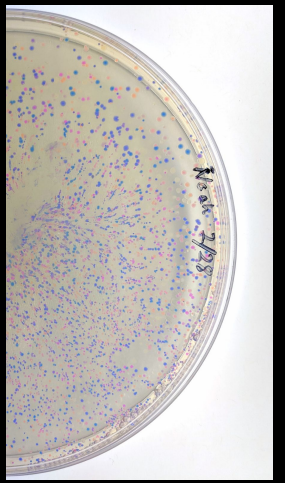
Time Estimate: 30 Minutes, 24 Hour Incubation

Media, Equipment and Consumables

- Ice bucket
- Phusion High-Fidelity PCR Master Mix
- Primers (working stocks, at 10 uM)
- UltraPure Water
- PCR tubes
- Thermocycler
- P20 pipette and 10uL tips
- P200 pipette and 200uL tips

Setup the following PCR reactions:

Universal DNA Fragment (Univ Fwd and Univ Primers)				
Reaction	**Reagent**	**Stock Conc.**	**Desired Conc.**	**Volume**
All	Mini Prepped mUAV	Check Stock	Typically 10 ng	X uL
All	Univ Forward Primer	10 uM	0.5 uM	1.25 uL
All	Univ Reverse Primer	10 uM	0.5 uM	1.25 uL
All	Phusion HF PCR Mix	2X	1x	12.5 uL
All	Nuclease-free water	n/a	n/a	10 - X uL
All	Total Volume			25 µL



Lab Work is Constrained by Physical Capacity

68-089 (HTGAA Lab Space)

Total Bench Width: ~101.5 ft

Bench Depth: ~2.5 ft

~15.5 ft lab bench

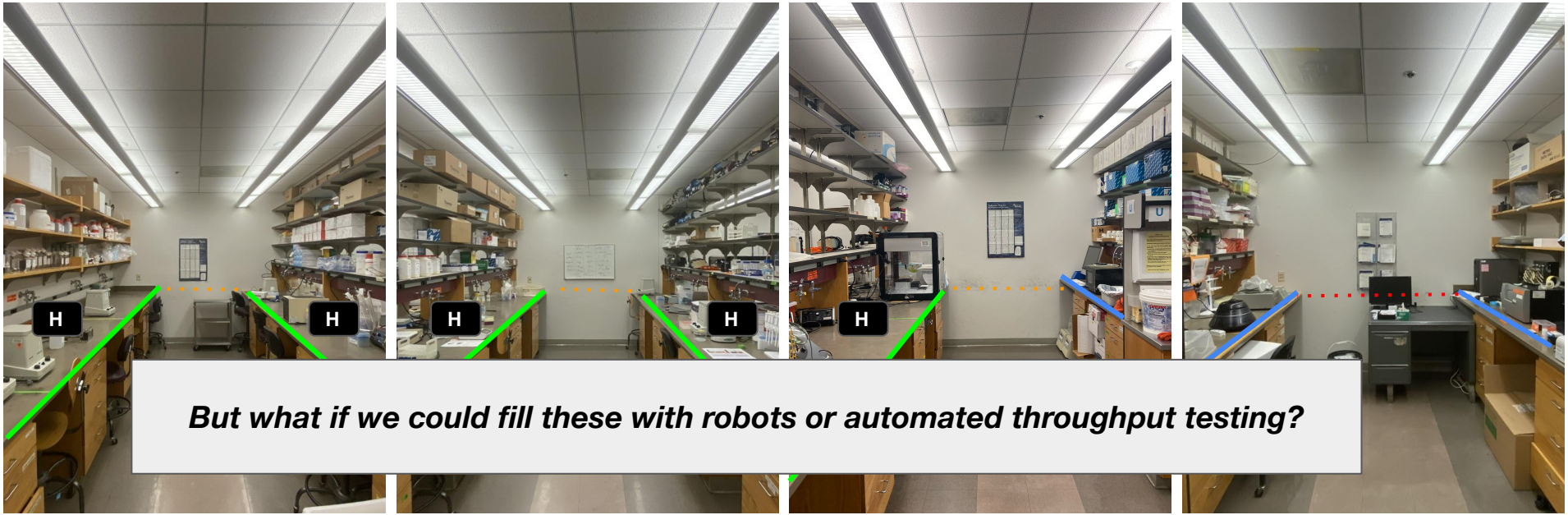
~8 ft lab bench

~5 ft space

~5.5 ft space

NOTE: HTGAA Currently uses the first 5 lab benches in this room, denoted by **H**

Lab Work is Constrained by Physical Capacity



68-089 (HTGAA Lab Space)

Total Bench Width: ~101.5 ft

Bench Depth: ~2.5 ft

~15.5 ft lab bench

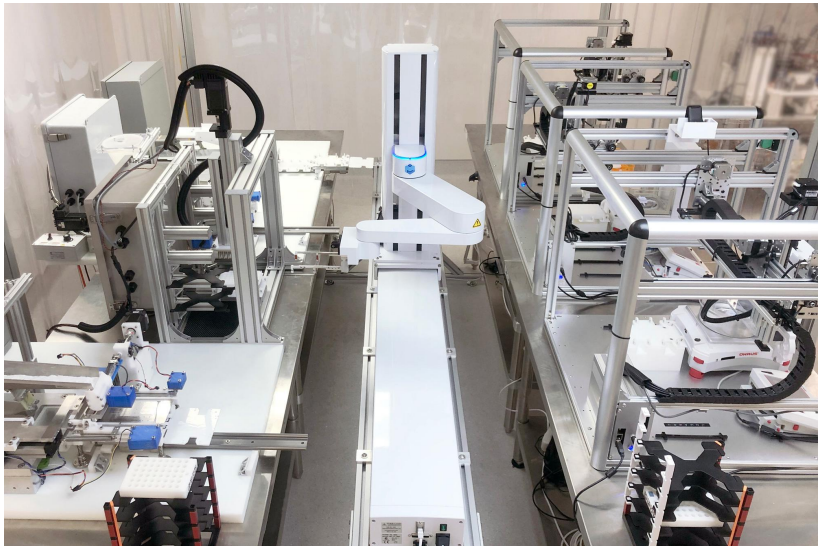
~8 ft lab bench

~5 ft space

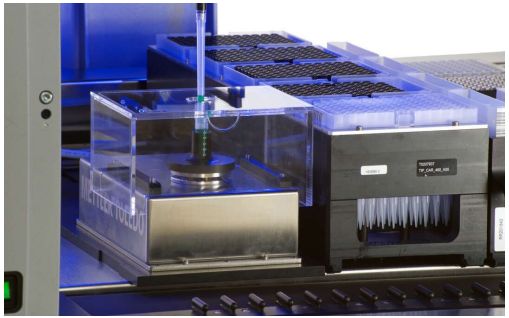
~5.5 ft space

NOTE: HTGAA Currently uses the first 5 lab benches in this room, denoted by H

Lab Modules



Lab Modules



Opentrons Thermocycler module

Opentrons heater shaker

On-deck
plate reader

Lab Devices

>$300k

Hamilton STAR

<$10k

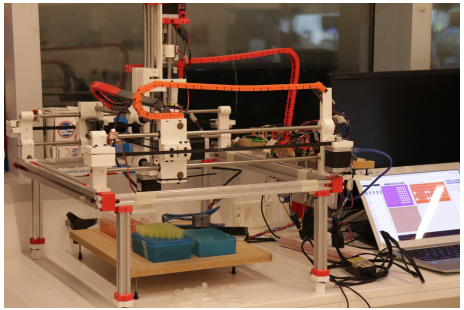
Opentrons OT-2*

>$300k

Tecan Freedom EVO

>$300k

Hamilton Vantage



Open source robot

Many more…

AI Scientist and Discovery

nature chemical engineering

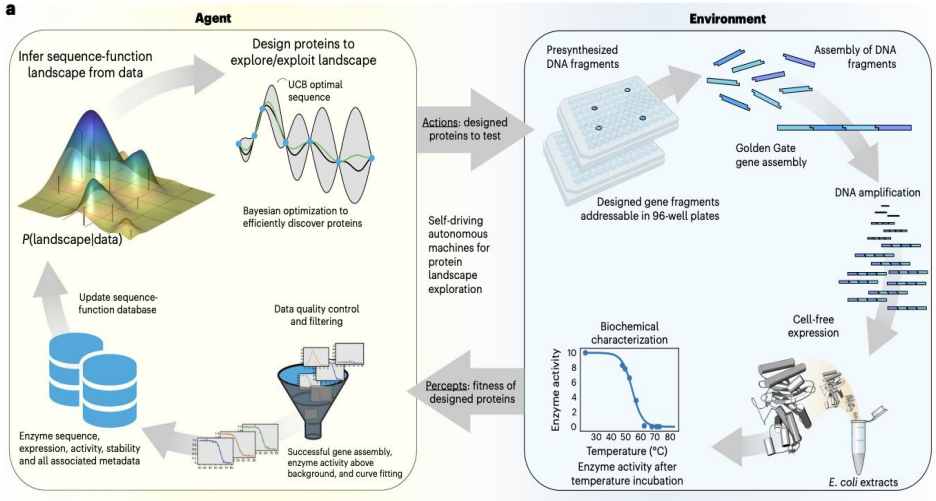
Article

https://doi.org/10.1038/s44286-023-00002-4

Self-driving laboratories to autonomously navigate the protein fitness landscape

Received: 20 July 2023

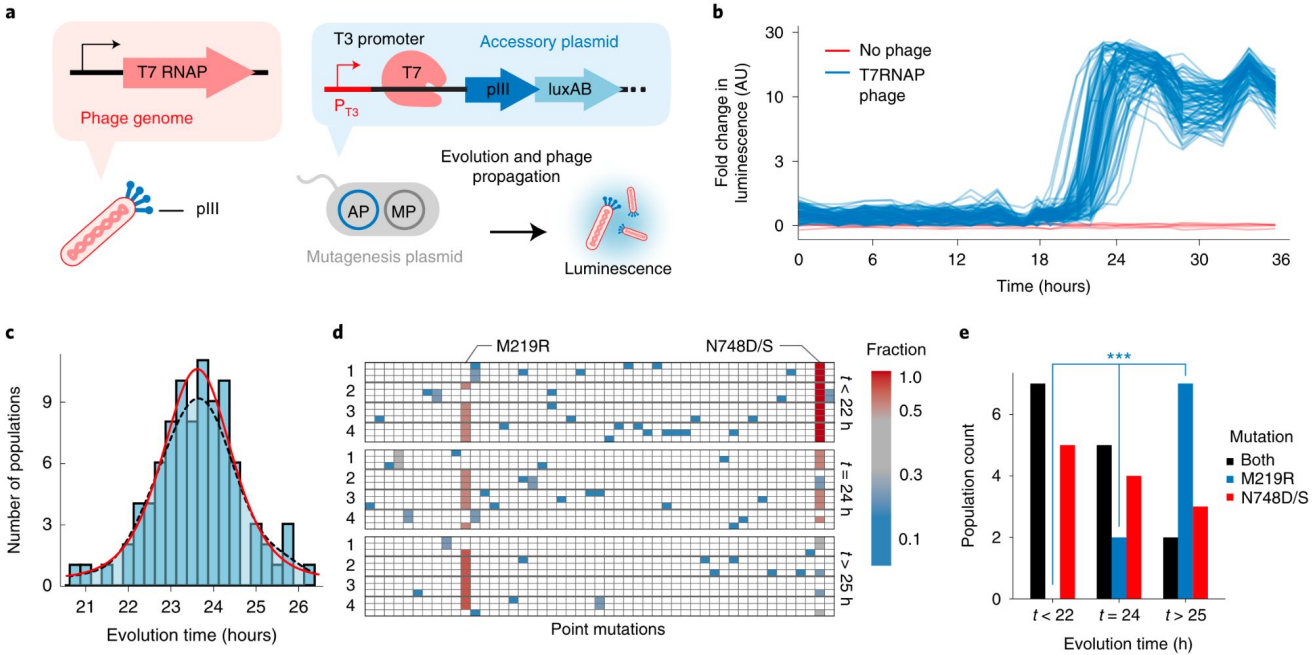
Jacob T. Rapp[1], Bennett J. Bremer[1] & Philip A. Romero[1,2]



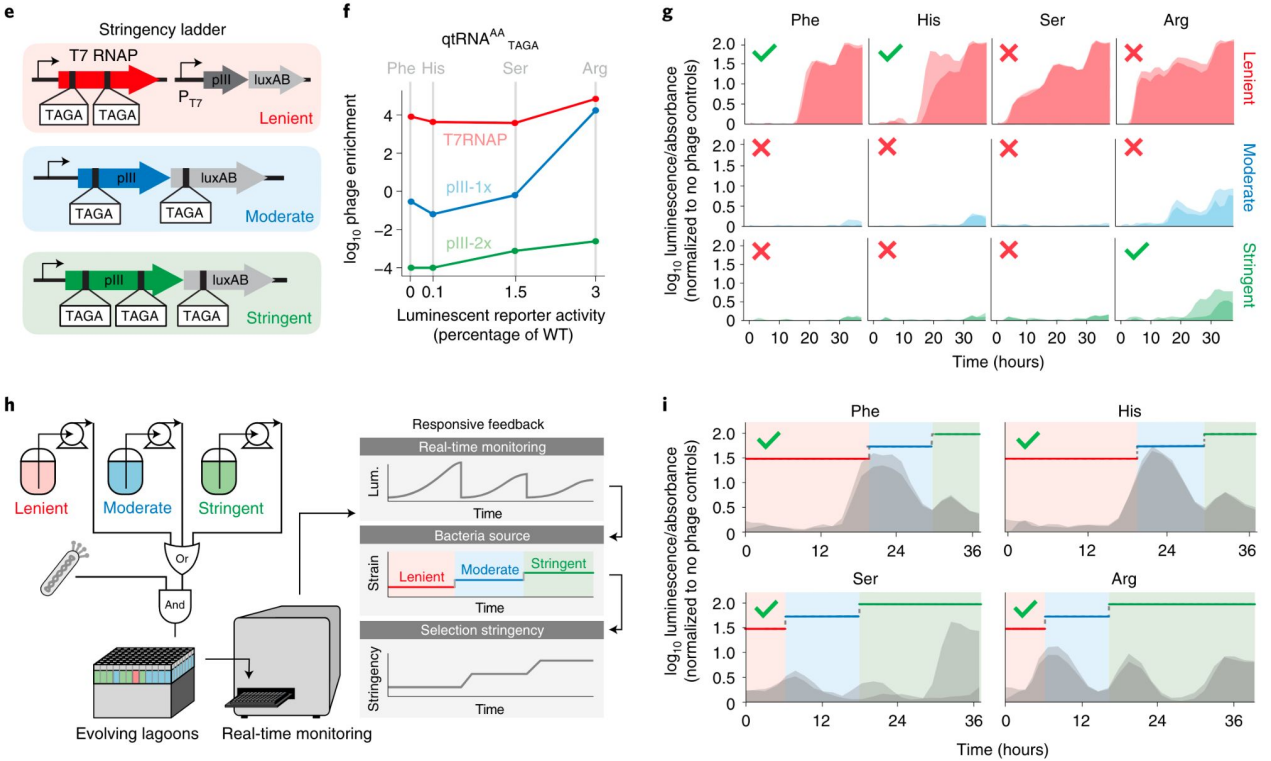
Automation-enabled real-time monitoring and feedback

Example: Continuous phage development

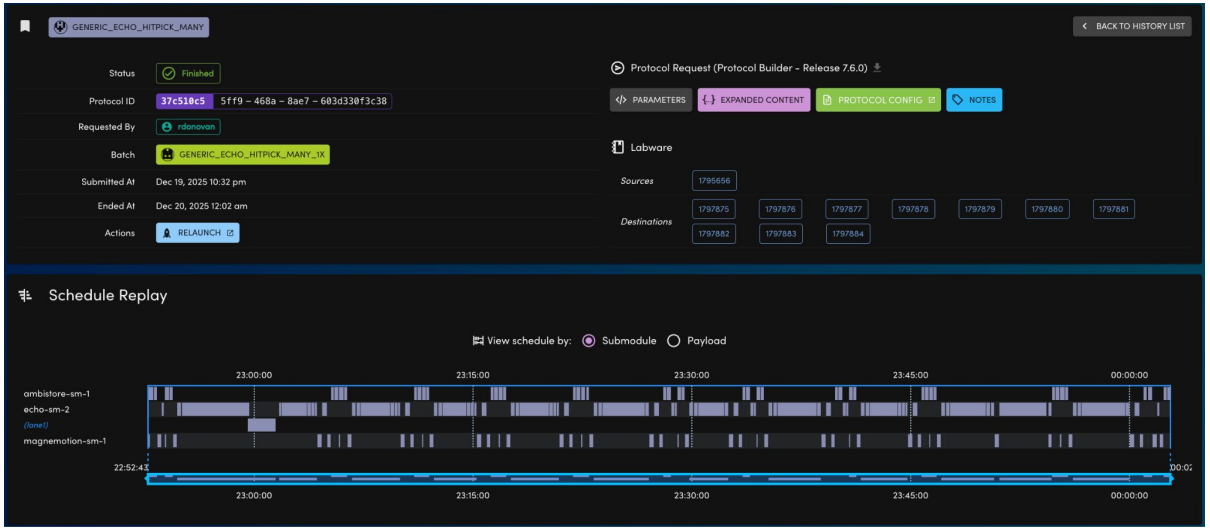
Can be used to automate conditions and DNA constructs for expression



Automation-enabled real-time monitoring and feedback



Automation-enabled real-time monitoring and feedback



Esther, Suvin, Ronan conducting fluorescent artwork preparation with a laboratory scheduler

Part 3: Demonstrations

Demo 1: RF Diffusion, AlphaFold

Demo 2: Lab Automation

Part 4: Case Study