Lineage Tracking for Synthetic Biology: CellRepo® version control platform for Microbial Research

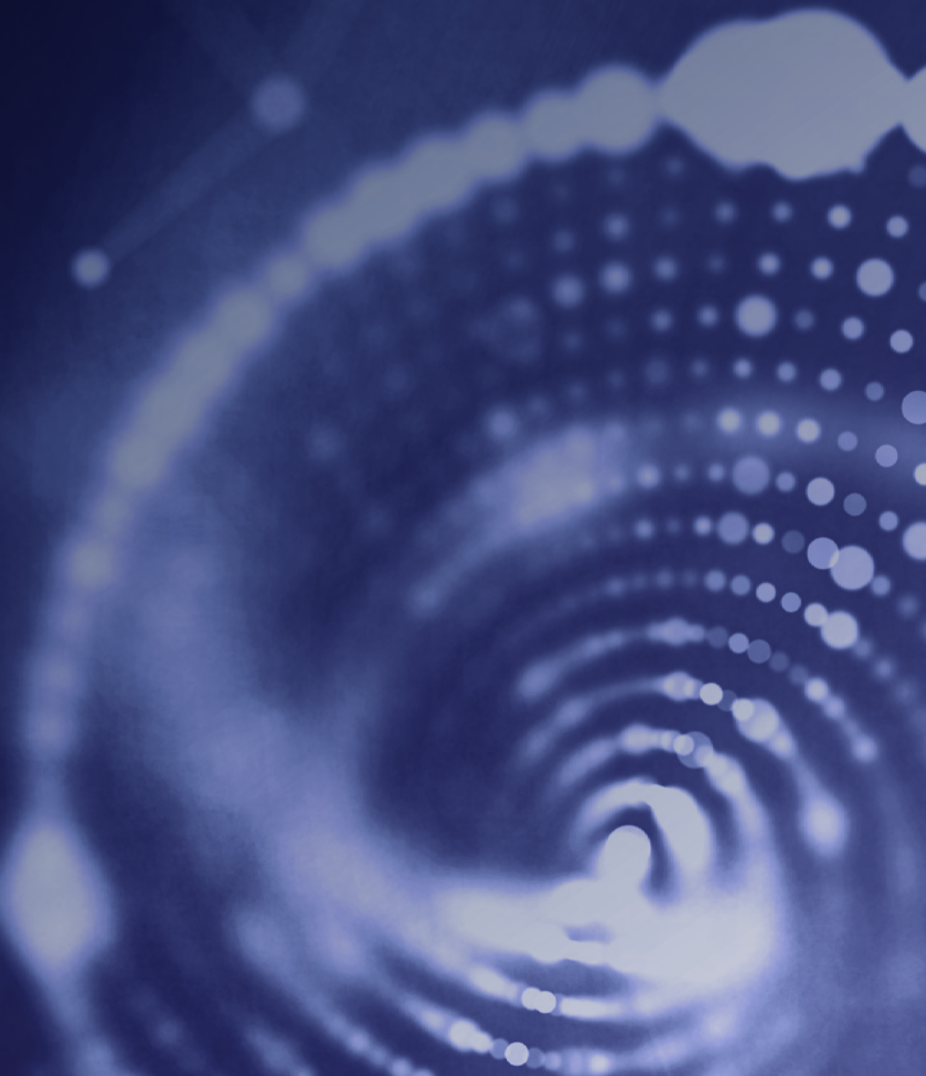
---

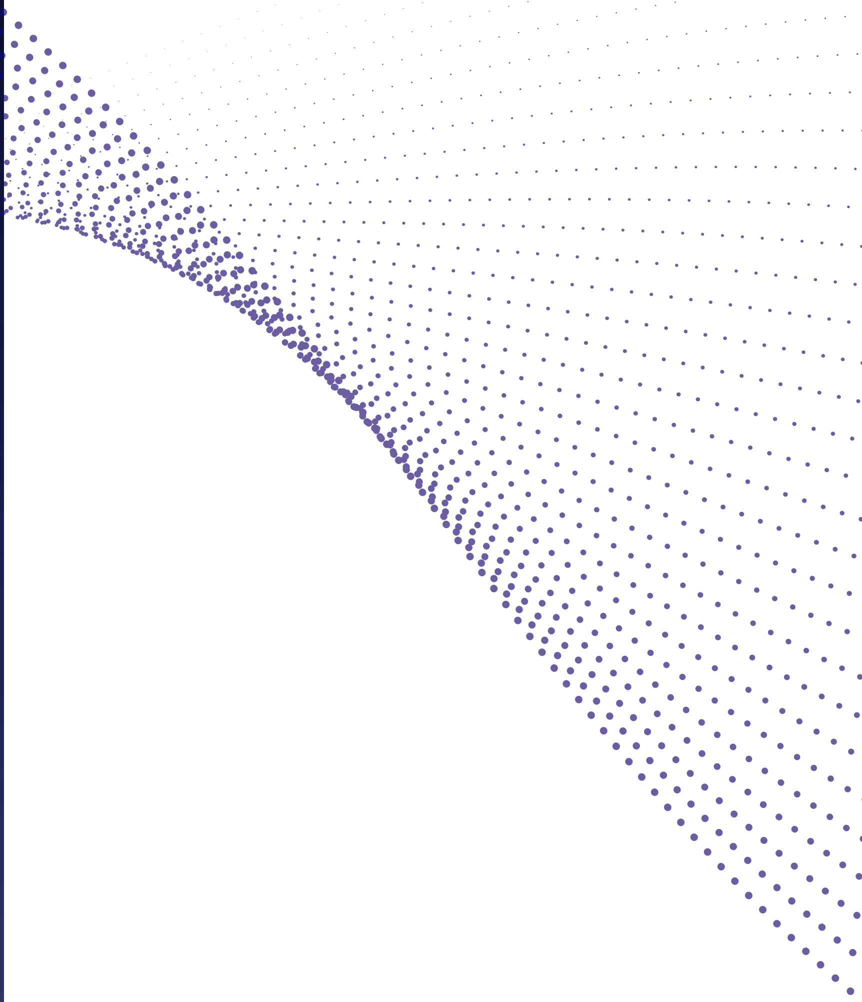
GitLife
Biotech

INTRODUCTION

Microbiology and synthetic biology are evolving rapidly. Labs are generating more data, building more constructs, and iterating faster than ever. But many researchers still rely on fragmented tools to manage the complexity, this leaves them vulnerable to errors, wasted time, and lost work.

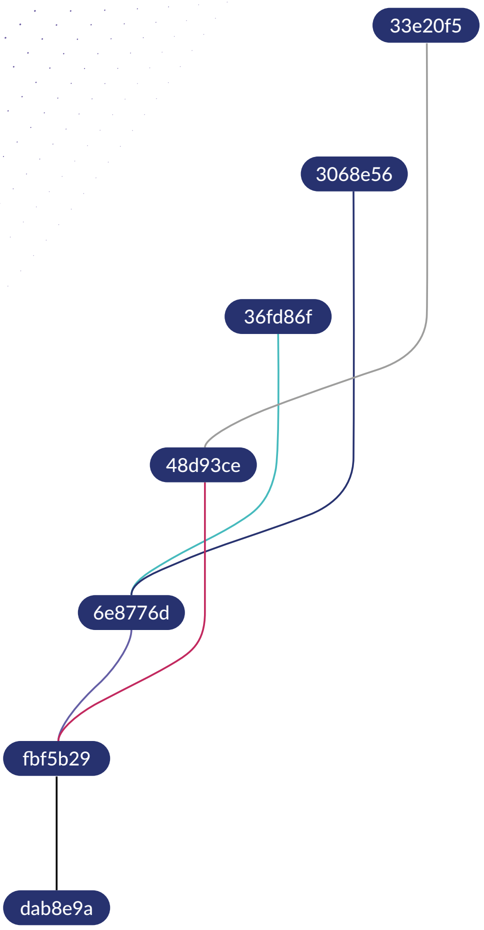
In microbial research, where constructed DNA and strains are constantly modified, reused, and shared across teams, tracking what changed, when, and why can be surprisingly difficult. Mislabelling a plasmid, updating a protocol without version control, or losing context on strain history can derail experiments and compromise reproducibility.





CELLREPO ADDRESSES THIS BY BRINGING BIOLOGICAL VERSION CONTROL INTO THE LAB.

Every change to a strain, plasmid, or protocol is captured as a structured, timestamped, and attributed commit hereby creating a complete, auditable record of each step. With built-in data templates, branching support, and traceability across users and time, CellRepo turns biological recordkeeping into a reliable, collaborative system.



33e20f5
Mutant Δabc2_pRegulate
Nov 14 2025, 16:15 GMT
Author: Gilly

3068e56
Mutant Δabc1_pRegulate
Nov 14 2025, 16:11 GMT
Author: Gilly

36fd86f
Mutant Δabc1Δabc2
Nov 14 2025, 16:00 GMT
Author: Gilly

48d93ce
Mutant Δabc2
Nov 14 2025, 15:57 GMT
Author: Gilly

6e8776d
Mutant Δabc1
Nov 14 2025, 15:55 GMT
Author: Gilly

fbf5b29
Wildtype strain
Nov 14 2025, 15:11 GMT
Author: Gilly

dab8e9a
Initial Commit
Nov 14 2025, 15:03 GMT

How CellRepo Works in Practice

CellRepo is built on top of Git, the world-standard version control system powering the software industry. In CellRepo, each biological record, whether a strain, plasmid, or protocol, is stored in a repository. When users make changes, such as updating the genotype or attaching a sequencing file, they do so by creating a new commit. This commit:

- Preserves all previous data
- Logs the time, author, and change description
- Supports comparison with prior versions

Mapping Git Concepts to CellRepo

To help teams manage complexity and collaborate effectively, CellRepo adopts core concepts from Git but tailored for synthetic biology:

- Repositories: Each record or information of a construct lives in its own repository, capturing the full version history of their relevant data. Just like a Git tracks a software project.
- Commits: Every change is saved as an immutable snapshot with full traceability: what changed, when, by whom and for what purpose. Capturing information that cannot be gleaned from simply genome (re)sequencing.

- Commit Templates: Structured templates provide convenience and flexibility while ensure consistent data collection. For example, we provide templates for genotyping, plasmid maps, and antibiotic resistance markers, ensuring comparable records across users and entries.
- Branches & Merges: Different lines of experimentation or documentation can be maintained in parallel. For instance, separate teams creating a divergent construct can branch from a common base strain. If needed, users can then merge branches while preserving full version history.
- Projects and Organisations: Like modern version control systems used in software development, CellRepo provides a structured framework for managing related repositories across individuals, teams, labs, and institutions. This enables clear ownership, access control & privacy, and coordination across collaborative research environments.
- Forks: Users can fork repositories to explore changes independently of the original repository, this supports parallel work without losing traceability.

Real-life Example:

A synthetic biology team is iterating on a plasmid design for a reporter assay. One researcher modifies the promoter and uploads the new version to CellRepo, complete with sequence, plasmid maps, and growth conditions. A few days later, another team member notices that the antibiotic resistance marker in the plasmid doesn’t match the one used in the current selection plates.

Instead of manually editing the file or making a guess, they check CellRepo’s commit history. The mismatch is traced to an earlier version, and a corrected construct is created via a new commit—preserving the full change history without confusion. The team avoids repeating several failed transformation attempts, and everyone stays aligned on which version is in use.

By making every change traceable and versionable, CellRepo helps researchers avoid wasted time, miscommunication, and data loss—challenges that are common in everyday synthetic biology work.

Structured data through Commit Templates

CellRepo's commit templates enforce lab-specific structured data collection. Fields such as host organism, genotype, growth conditions, and protocol steps can be configured as required fields with controlled vocabularies. This reduces ambiguity and makes entries comparable.

Example Use Case: In a teaching lab, students submit plasmid constructs through CellRepo using a shared template. Fields like 'plasmid name', 'backbone', 'promoter', and 'reporter gene' are filled in consistently. Instructors can then compare submissions or track changes made during troubleshooting.

Traceability and Collaboration

In multi-user environments, traceability is essential. CellRepo logs every change with full author attribution and timestamps.

Example Use Case: A consortium project across three labs is tracking CRISPR trial strains. Each lab maintains its own branch in CellRepo. When data is consolidated, differences in construct data (e.g. sgRNA design or resistance cassette) can be compared across branches without loss of context. Any branch can be merged or preserved separately.

Genosignatures®: Embedding data in DNA

CellRepo supports Genosignatures, DNA barcodes containing repository specific data that can be embedded into a strain. This enables recovery of the construct's digital history from physical samples without the need to carry out a full genome sequencing and annotation.

Example Use Case: A strain retrieved from a biobank is found to contain a Genosignature. Sequencing the Genosignature alone, reveals the CellRepo commit ID, which links to the full digital footprint behind this strain: who created it, what plasmids it contains, and its passage history, growing conditions, use licenses, etc. This is especially useful for tracing provenance after sample transfer.

CONCLUSION

Reproducibility is not just a scientific ideal, it is a foundation for scalable, investable, and trustworthy biology. CellRepo doesn’t merely improve documentation; it introduces a new paradigm: structured, immutable, and collaborative version control for living systems that integrates seamlessly with your laboratory workflows.

By linking digital provenance to physical constructs by embedding Genosignatures directly into engineered strains and plasmids, CellRepo transforms the practice of experimental biology from opaque to auditable, from isolated to collaborative.

In a world of accelerating biotech innovation, CellRepo ensures that your data, and the value it carries, can keep up.

For more information please go to
www.cellrepo.com

GitLife
Biotech

gitlifebiotech.com