



← Go to [NeurIPS 2026 Workshop AgenticLS homepage](#)

Do Repeated LLM Decisions Improve CRISPR Hit Discovery?



*Odysseus Kaloidas*¹

¹ University of California, Berkeley (berkeley.edu)

27 Sept 2026 (modified: 27 Sept 2026)
 NeurIPS 2026 Workshop AgenticLS Submission AgenticLS, Reviewers, Authors
 Revisions CC BY 4.0

Keywords: biomedical AI agents, large language models, CRISPR screening, experimental design, feedback, active learning, agent evaluation

TL;DR: Experimental feedback can improve LLM-guided CRISPR discovery without establishing an advantage over an informed numerical controller.

Abstract:
Does an LLM that benefits from experimental feedback also choose better experiments than an informed numerical policy? We study retrospective CRISPR hit discovery, where a nearest-neighbour procedure expands five proposed genes into a batch. Our numerical rival samples centres from observed hits while retaining the same initial batch, features, expansion procedure and test budget. Across six models and six readouts from four sources, correct outcomes improve LLM yield over withheld outcomes by 18.18 hits per 640 tests, yet the rival finds 9.98 more hits. Complete-campaign checks with expanded response budgets and a native JSON interface retain this pooled distinction. The native-interface model requires no output recovery, but wins on one source and loses on another. Initialization also changes which controller performs better. Finally, LLM proposals contain some useful candidates that a hypothetical selector with access to future outcomes could exploit, but the tested selectors do not establish a gain over equally sized numerical candidate pools. These conditional findings separate feedback use from the added yield of repeated LLM control. Evaluating that added yield requires an alternative that retains both outcome information and the numerical procedure used to turn proposals into experiments.

Email Sharing: We authorize the sharing of all author emails with Program Chairs.
Data Release: We authorize the release of our submission and author names to the public in the event of acceptance.
Submission Number: 128

Filter by reply type...

Filter by author...

Search keywords...

Sort: Newest First

Everyone

Program Chairs

Submission128 Authors

1 / 1 reply shown

Paper Decision



Decision by Program Chairs



30 Sept 2026, 01:05 (modified: 30 Sept 2026, 07:34)



Program Chairs, Authors



Revisions

Decision: Accept (Spotlight)

Comment:

The study provides unusually careful evidence about when feedback helps and when repeated language-model decisions fail to add value over a matched non-language-model controller. Its factorial decomposition, source sensitivities, initializer reversals and interface controls turn a negative result into a strong methodological contribution. The conclusions remain specific to retrospective released-hit replay and the evaluated controller-executor combinations, with unresolved candidate and incomplete-campaign mechanisms transparently bounded.

Add:

Official Comment

About OpenReview

FAQ

Sponsors

All Venues

Contact

Hosting a Venue

Terms of Use / Privacy Policy

Donate

OpenReview is a long-term project to advance science through improved peer review with legal nonprofit status. We gratefully acknowledge the support of the OpenReview Sponsors. © 2026 OpenReview