Received: 18 June 2021 | Revised: 31 March 2022 | Accepted: 18 May 2022

DOI: 10.1002/bmb.21638

ARTICLE

Biochemistry and Molecular Biology Education WILEY

LacOp: A free web-based lac operon simulation that enhances student learning of gene regulation concepts

Richard Charczenko[1] | **Madeline McMahon**[1] | **Kimberly Kandl**[2] | **Robert Rutherford**[3]

[1]Departments of Biology and Computer Science, Seattle University, Seattle, Washington, USA

[2]Department of Biology, St. Olaf College, Northfield, Minnesota, USA

[3]Department of Biology, Seattle University, Washington, USA

Correspondence
Rutherford Robert, Department of Biology, Seattle University, 901 12th Ave, Seattle, WA 98122, USA.
Email: robertr@seattleu.edu

Abstract

Here, we describe a free, web-based simulation of the lac operon, "LacOp," that is designed to enhance the learning of prokaryotic gene regulation and pathways in advanced high school and undergraduate genetics courses. This new electronic resource was created by a team of students in an advanced undergraduate course and is hosted online (http://flask-env.rnwhymamqf.us-west-2.elasticbeanstalk.com/lacop). LacOp has a simple web interface compatible with a range of devices, including smartphones. To determine whether the LacOp simulation enhances student learning from traditional instruction, we introduced the lac operon to undergraduate genetics students through a traditional classroom experience followed by use of the LacOp simulation. Students worked on their own using the included tutorial to create and test the effect of various genotypes on *E. coli* lactose metabolism and regulation. Upon completion of the tutorial, students showed measurable gains in conceptual understanding of the lac operon. These students also reported a generally favorable opinion of the LacOP simulation as a use of their instructional time.

KEYWORDS

bioinformatics, gene regulation, lac operon, modeling, online laboratory, pathway, Python, systems-based thinking

1 | INTRODUCTION

By their nature, biological systems are dynamic, interactive, and complex. These traits make biological systems amenable to study by modeling and simulation activities. Computer simulations can have the same benefits as other active learning strategies.[1] The current pandemic makes simulations that are compatible with physical distancing and remote instruction, especially timely.[2] Here, we describe a simulation of the lactose operon, LacOp, a free, simple to use online simulation that expands student comprehension of prokaryotic gene regulation.

Simulations have grown in importance as a tool to aid learning. Almost a decade ago, leaders in undergraduate education[3,4] called for a fundamentally new approach that emphasized the importance of developing competencies that include the process of science, the ability to use quantitative reasoning and the ability to use modeling and simulation. In parallel, genetics educators have built and tested a range of simulations, for example to model the central dogma of molecular biology and microarray technology,[5] multiplex PCR,[6] and simulated transgenics.[7] PhET, a group at the University of Colorado, has developed a number of STEM simulations (https://phet.colorado.edu/en/simulations) including one that models the lac operon.[8]

This informative simulation shows individual molecules and diffusion, but has limited quantitative outputs

and does not the support combinations of mutations often studied when students learn about the lac operon. The Cell Collective[9] has also developed an innovative tool which allows students to build and test their own model of the lac operon,[10] however, it requires users create accounts and grow comfortable navigating a complex, multi window environment, which has limitations when used as a one-time classroom simulation. To our knowledge, neither the cell collective nor the PhET simulation creators have published assessment(s) on the effect of their lac operon simulations on conceptual learning.

In this article, we describe an easy to use web-based simulation of the lac operon that was designed and developed by a team of advanced undergraduates in a bioinformatics projects course at the Seattle University. This electronic resource is free, does not require a login, and consists of a simple interactive web interface accessible from a range of devices, including smartphones. We selected the lac operon for simulation development because it supports core concepts in genetics education,[11] appears in almost all Advanced Placement and undergraduate genetics textbooks, and is frequently used in genetics or molecular biology courses as a way to introduce students to gene regulation and metabolic pathways. Since many students first encounter, the lac operon via traditional instruction we tested whether the LacOp simulation enhances student learning when used as a complement to conventional instruction practices. We also assessed the students' attitudes toward the simulation and their perception of value. Preliminary analyses show that there may be an enhancement in conceptual learning and that students perceived LacOp as a valuable tool.

2 | METHODS

2.1 | Development of LacOp simulation

All undergraduate Cellular and Molecular biology majors at the Seattle University enroll in a bioinformatics course focused on the key tools (e.g., BLAST, protein viewers, etc.)[12] and basic programming skills using the Python language. Every 2 years, we offer a subsequent projects course for those who wish to further develop their coding skills in small collaborative teams of three to four students. Each team envisions and builds a novel biological simulation(s) to assist life sciences education. This approach combines student's expertise as new learners of life science with their enthusiasm to create a project of their own conception while practicing coding skills. Some of these simulations become independent projects and persist beyond the course, growing in complexity and potential utility to a broader set of students. The LacOp simulation is one such project.

The LacOp simulation models the biochemical lactose pathway and the various mutations that can occur within the genes of the pathway. It is a learning tool designed for advanced high school and undergraduate genetic and molecular biology students to help them understand the lactose operon, and it lays a foundation for students to interpret scientific assumptions using graphical data.

2.2 | Technical details of the LacOP simulation

The development of the LacOp simulation can be split into two main phases. The first phase was the development of the backend program that does all the biological simulation. For the backend, all scripts were written in Python 2.7 without any external Python library dependencies, largely on desktop and laptop machines. The LacOp tool was modeled using Michaelis–Menton equations with approximations.[13] Beta-galactosidase rate constants were derived from UniProt (https://www.uniprot.org/uniprot/P00722). In doing so, we consulted published work which computationally modeled the lac operon[14,15] adapting those aspects which were most important for this educational use. For the second stage, we developed a web application to allow general access to the simulation. To aid in the creation of a web application and server management, we chose to use the micro framework known as FLASK (https://flask.palletsprojects.com), a set of tools within the Python programming language, for its versatility and ease of use. FLASK allowed us to host a webpage that would display information from the backend scripts as well as collect information from the web page and pass it back to the backend programs. The WTForms Python library was used to help control user input to the webpage. The pygal library (www.pygal.org) was chosen due to its flexibility and clarity when displaying the information from the backend in a graphical model within an html page. Finally, to host our webpage we chose Elastic Beanstalk as a platform from Amazon Web Services (https://aws.amazon.com/elasticbeanstalk/). Elastic Beanstalk is a platform offered as a service that automatically manages the underlying systems on a server and allows for a ready to go platform to host a web application. Together, this set of tools allows accessibility from a wide range of operating systems platforms. Importantly, this combination of tools seems likely to be resilient and functional in the face of evolving hardware and software, which is always

an issue with electronic resources. LacOp can be accessed and used via a web browser from computers and web-enabled phones.

2.3 | The target audience for the LacOp simulation

We identified the target audience for the simulation as any group of students engaged in learning gene regulation with access to the Internet, especially undergraduate students and advanced high school students. Although the simulation grew organically based on its designers' interests and can be used in a range of ways, we designed the accompanying tutorial (see supplementary materials and also at http://supergenetics.org/lacop) to support the learning objectives measured by the Lac Operon Concept Inventory (LOCI).[16] This 8-page tutorial guides students to perform digital experiments with the simulation and prompts them with questions that require students to interpret the simulations' graphical results by providing written answers. While the full text of the tutorial is available at http://supergenetics.org/lacop, an overview of the tutorial is available in Section 3.

2.4 | Simulation use and assessment in the classroom

Students in the Fall 2019 intermediate genetics course (BIO 233) at St. Olaf College participated in a remote site pilot study on the use of the LacOp simulation in a classroom laboratory environment. This course is typically taken by sophomores or juniors, and it is a required course for the biology major, the biomolecular science concentration and for some pre-health students. The course enrolls approximately 150 students per academic year; there were 58 students enrolled in the Fall 2019 section. The course consists of three 55-minute class sessions and one 3-hour lab each week. The LacOp simulation was used during one of the three-hour lab sections, but it could also be used as a take-home exercise.

To measure the simulation's value in enhancing traditional classroom experiences, students had a 55-min class prior to the LacOp activity. During this class, students were introduced to general features of gene regulation including concepts such as positive and negative gene regulation and effector molecules. They were also given introductory information about operons as well as the function of the *lacI*, *lacY*, and *lacZ* gene products and the negative regulation of the lac operon. Students then attended one of three different three-hour labs that occurred that afternoon, the next morning or the following afternoon (4, 23, or 28 h, respectively, after the end of the class). To assess their knowledge of key concepts before engaging with the simulation, students completed the Lac Operon Concept Inventory (LOCI).[16] The LOCI is a 12-question multiple-choice assessment tool, whose learning objectives are (1) identify and understand the role of the structure and components of the lac operon (Questions 1–4), (2) when given particular cellular conditions, accurately predict whether or not gene expression will occur (Questions 5–6), and (3) when given particular mutations to the lac operon, predict affected outcomes of gene expression (Questions 7–12). To mimic a situation in which students might use the simulator on their own outside of class time, students were given minimal instruction on completing the LacOp simulation and encouraged to follow the instructions in the LacOp tutorial while working at their own pace. After completion of the LacOp tutorial and submitting written answers to the questions therein, students once again completed the LOCI, this time as a post-test of their conceptual knowledge.

Five weeks after completing the LacOp simulation, students were asked for their opinions on the LacOp simulation. Attitudes toward the simulation were measured by adapting a previously reported survey of likeability, usability, credibility, and acceptability of a biological simulation.[7]

The Seattle University IRB has reviewed this research and determined the study to be exempt from IRB review in accordance with federal regulation criteria. St. Olaf College's IRB accepted Seattle University's IRB exemption.

3 | RESULTS AND DISCUSSION

3.1 | The LacOp simulation and tutorial overview

The LacOp simulation helps students learn *E. coli* lactose metabolism and regulation, common topics in most genetics textbooks.[17] Within the simulation, a user typically configures an *E. coli* cell by choosing a set of initial sugar substrates and their concentrations and the cell's genotype at key loci. Having done so, the user launches the dynamic simulation. Upon completion, the simulation generates a graphic showing the concentration(s) of a set of key molecules over time (Figure 1).

The written LacOp Tutorial (See supplementary material, or the link in the online simulation), which is about eight pages long when printed, guides students through the use of the simulation. It begins by introducing students to the simulations with a simple,

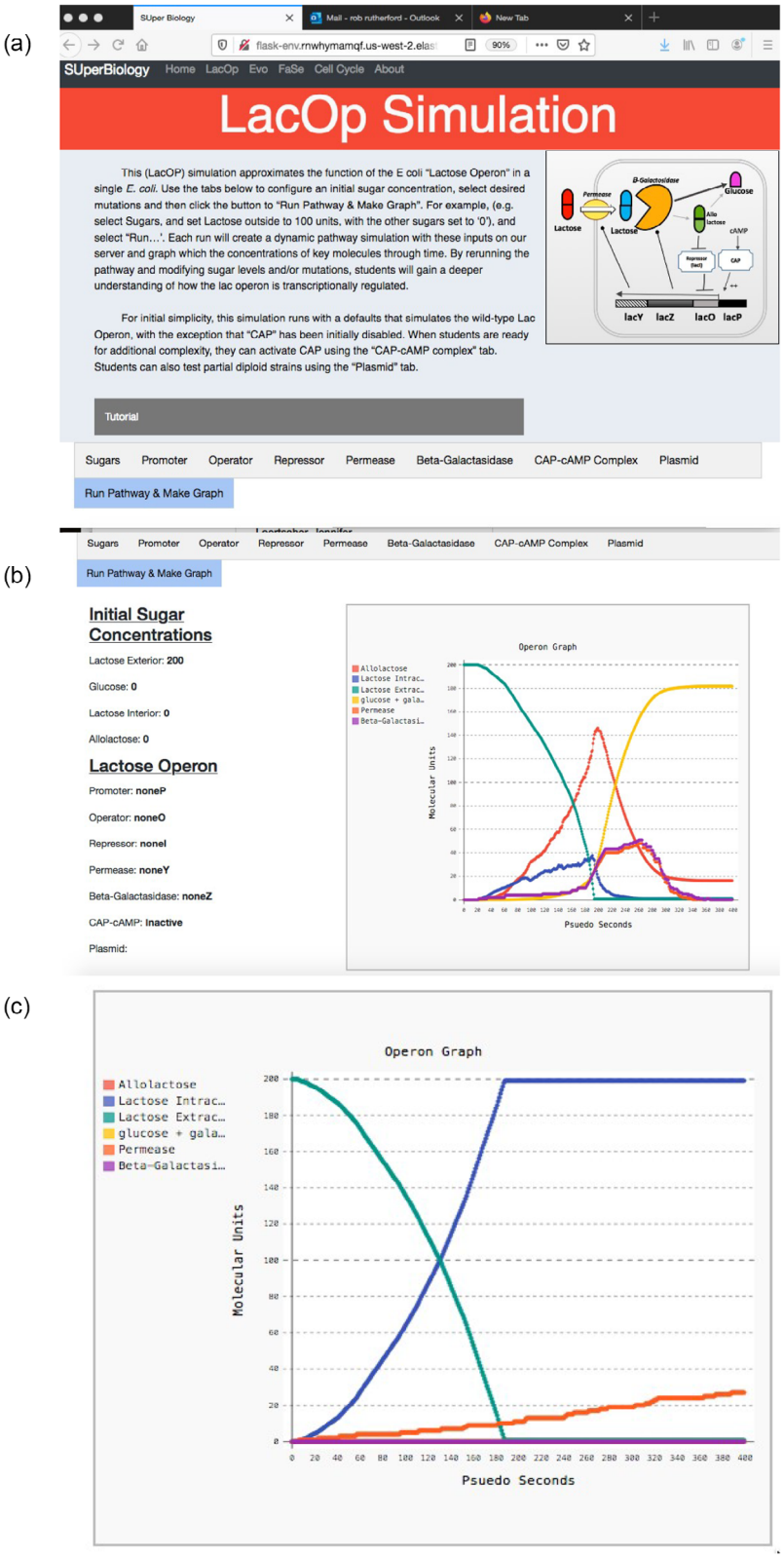

FIGURE 1 A browser view of LacOP interface and output. (a). Shows what the user sees when accessing the LacOp simulation with a simplified metabolic pathway for lactose utilization. (b). LacOp output showing wild type conditions and the lactose operon products. (c). A graph of the same initial concentrations as in B, except that the cell lacks a functional *lacZ* gene that encodes ß-galactosidase. Note that the blue line represents intracellular lactose which accumulates because without lacZ, lactose cannot be metabolized into glucose and galactose

metabolically informative user interface (Figure 1a) with preconfigured default conditions, and invites students to press the "Run Pathway and Make Graph" button. Doing so runs the simulation, which generates graphical output as shown in Figure 1b. These results are dynamic, for example, Figure 1c shows the output of the same conditions but with a loss of function mutation to *lacZ* (*lacZ*$^{-}$). The tutorial then directs students to note how extracellular lactose is moved inside the cell, converted to allolactose and ultimately to glucose over time. They are

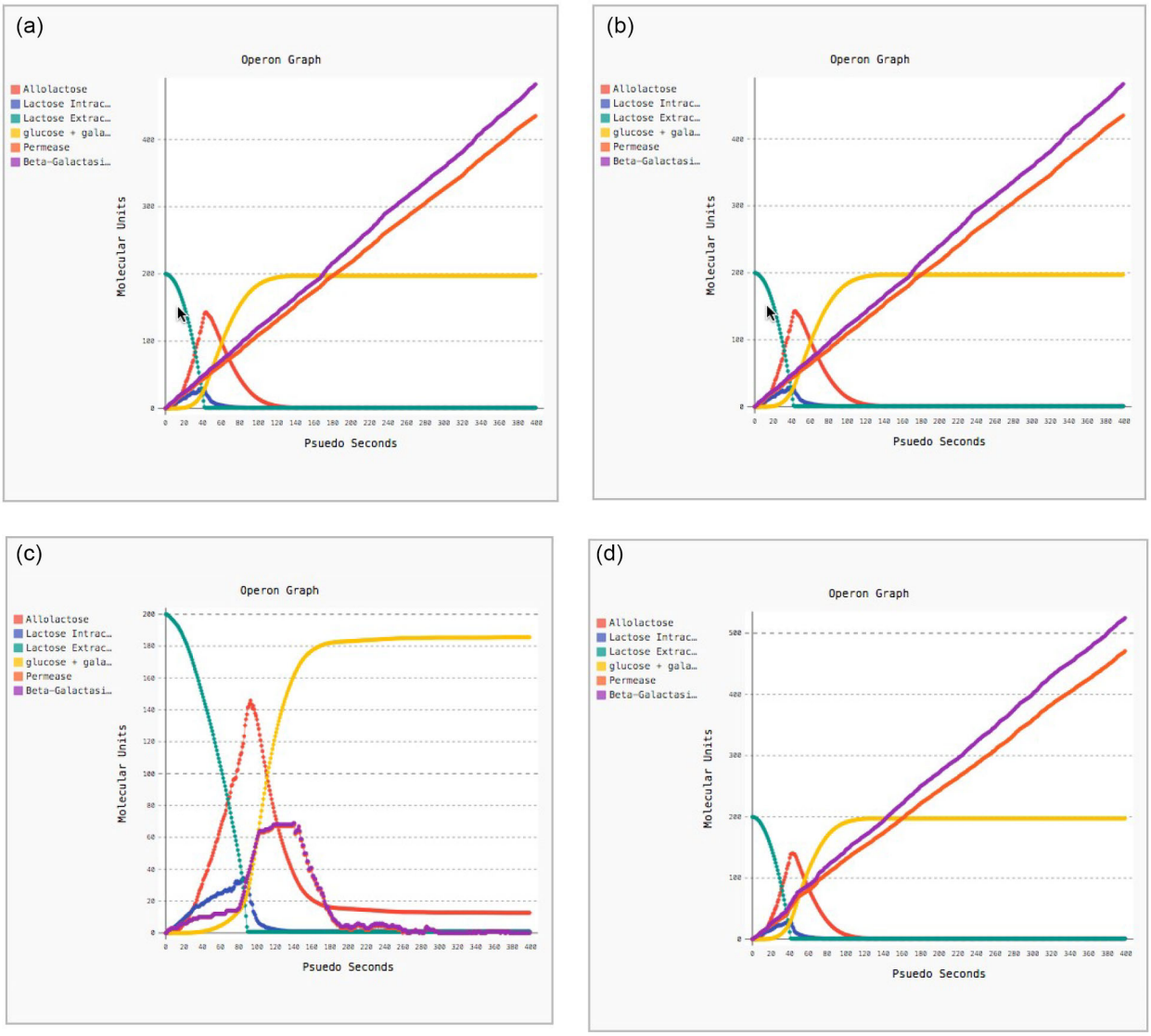

FIGURE 2 LacOp results from various regulatory mutants. (a). Lac repressor (lacI$^-$)mutant. Note the rising levels of ß-galactosidase (purple line) and permease(red line) which indicate constitutive expression. (b). Lac constitutive operator mutation. (c). Phenotype of a repressor mutant with a functional copy of the repressor locus "in trans" on a plasmid. (d). Phenotype of a constitutive operator mutant with a wild-type operator "in trans" on a plasmid. By working through these scenarios, students can simulate performing the classic "cis-trans" test described in many genetics textbooks

also encouraged to note the levels of the two proteins shown (permease and β-galactosidase). To deepen their understanding, the tutorial then directs students to rerun the simulation with a loss of function mutation of the promoter, by clicking on the promoter tab and selecting the *lacP*$^-$ checkbox, and interpret the results.

Students are then guided to examine regulatory mutants and run the simulation with a series of such mutations. Students test a repressor loss of function mutation, *lacI*$^-$, (Figure 2a), then a constitutive operator mutation, *lacO*^{C} (Figure 2b). Since the phenotypes and graphical output of these two mutants are identical, the tutorial then challenges students to find a way of discriminating between the two mutations. The LacOp simulation also supports the creation of partial diploid experiments. Students find that a wild-type copy of a transcription repressor (*lacI*) can restore regulation, even if outside the chromosomal locus on a plasmid (Figure 2c). By contrast a wild-type copy of the repressor binding site (*lacO*) cannot overcome the constitutive expression from a chromosomal *lacO*^{C} allele. In this way, the simulation encourages students to construct and perform their own "cis-trans" test, a key experiment that underpins the initial discovery of transcription regulators and their binding sites.[18] In a last experiment to test their understanding of gene

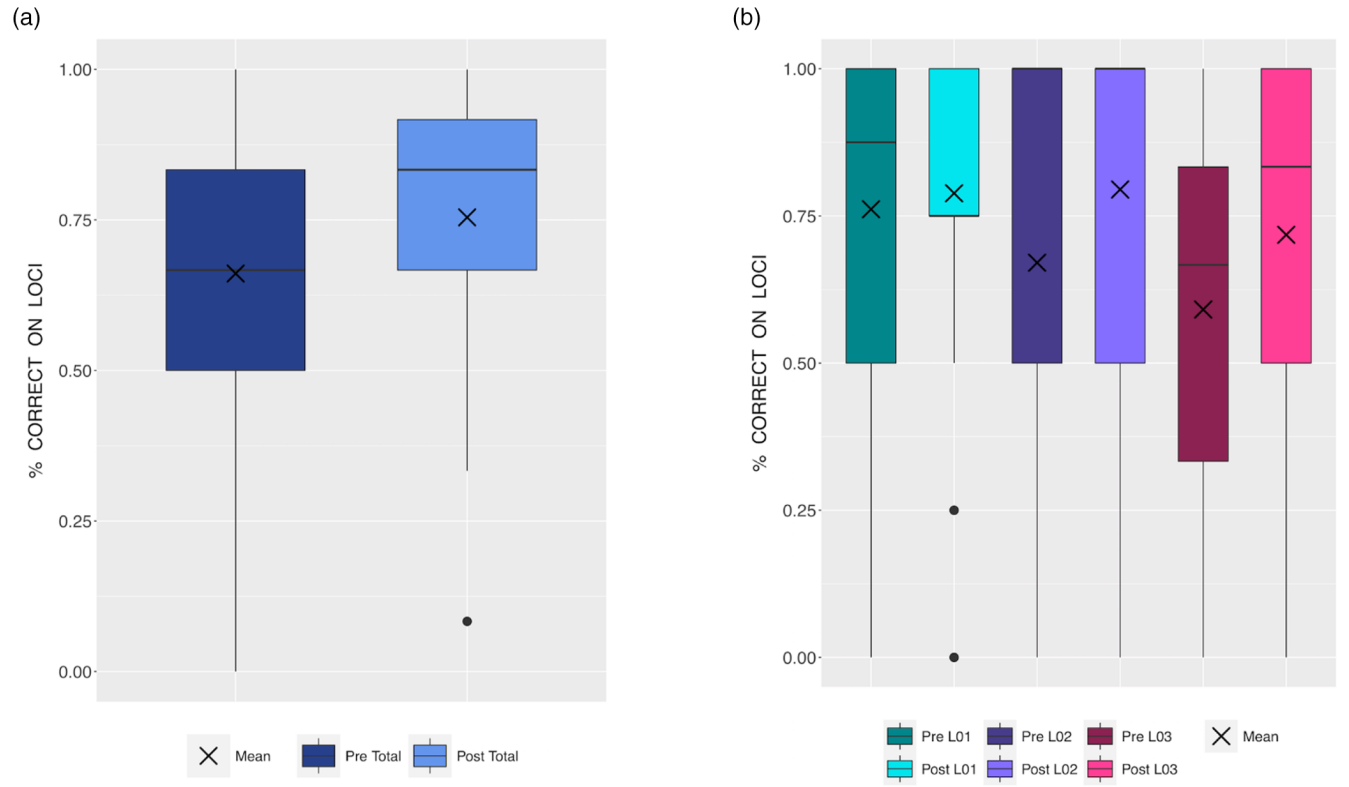

FIGURE 3 Conceptual knowledge pre and post LacOp simulation. Panel (a) shows student performance overall on the LOCI before and after completing the simulation. Panel (b) results by learning objective. LO1, knowledge of operon structure and its components; LO2, predicting outcomes of various cellular conditions; LO3, understanding the effect of known mutations. The "X" represents the mean score. Figure made by ggPlot package in R

regulation, students are asked to predict the phenotype of a cell possessing two seemingly opposite mutations: a constitutive allele of operator $lacO^C$ and the $lacI^S$ "Super Repressor." They then use the simulation to test this hypothesis, and then report and interpret the results in a written answer.

Although many students complete the tutorial at this point, it also contains some additional exercises for expanded learning. One exercise introduces positive regulators by exploring modified genotypes of the catabolite activator protein (CAP). A second invites students to find a way to discriminate between superficially similar phenotypes of promoter ($lacP^-$) and permease ($lacY^-$) mutants. Clever students can discover that this task is best completed by testing each mutant with intermediate compounds (such as intracellular lactose, thereby bypassing the need for a permease), a method similar to the classic experiments of Beadle and Tatum used to develop the "One Gene One Enzyme" Hypothesis.[17] In summary, the LacOp simulation is flexible, and students could use the simulation to simulate a wide range of possible scenarios that deepen their understanding of gene regulation and metabolism.

3.2 | Assessment of student learning

Although the simulation was rewarding to build, the most important question is whether it enhances the learning of students who use it. To answer this question, we used the Lac Operon Concept Inventory (LOCI)[16] to measure students' conceptual understanding of the lac operon immediately before and then after use of the LacOP simulation. Gains after use of the simulation can be seen broadly across the LOCI, which showed, on average, an 8.3% increase in students selecting the correct answer (Figure 3a). These increases were broadly reflected across all three learning objectives (Figure 3b). For example, we observe a 3.9% increase in correct answers for LO1 (identify and understand the role of the structure and components of the lac operon), a 10.9% increase on LO2 (when given particular cellular conditions, accurately predict whether or not gene expression will occur), and a 10.4% increase on LO3 (when given particular mutations to the lac operon, predict affected outcomes of gene expression). The greatest percentage gains were observed after using the simulation in the two most advanced learning objectives, LO2 and LO3, which

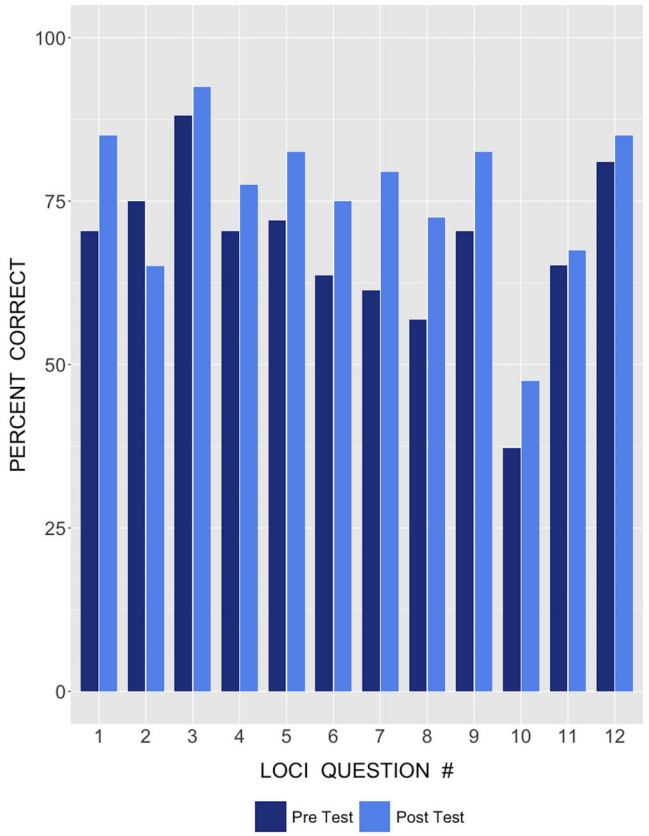

FIGURE 4 Conceptual knowledge pre and post LacOp simulation by individual question on the LOCI.[16] Questions 1–4 support learning Objective 1, Questions 5–6 support learning Objective 2 and Questions 7–12 support learning Objective 3. Figure made by ggPlot package in R

ask students to make predictions. Although the material in LO2 had been covered in class, students were not previously exposed to the material covered in LO3, as these topics were not covered during the 55-min class prior to the students' participation in the simulation. This suggests that the simulation might work synergistically with traditional instruction. Taken together, these data suggest the LacOp simulation improves the conceptual understanding of students on all learning objectives of the LOCI.

To understand the differences between the gains on each learning objective, we examined performance on each question (Figure 4). Students improved their score on 11 of the 12 questions. The one question in the Concept Inventory in which the simulation appeared to negatively affect student performance was Question 2, which states: "Which of the following does NOT encode for a protein that is directly involved in metabolizing lactose? a. *lacZ* b. *lacY* c. *lacI* d. *lacA*." We observed a 10% reduction in the correct answer, C. The most common error was to misidentify "*lacA*" instead of "*lacI*" as being "not....directly involved in metabolizing lactose." Since *lacZ* and *lacY* are given primary emphasis in most textbooks, and the simulation was built on the information in these textbooks, we did not initially include discussion of *lacA* in our LacOp simulation. We think it is plausible that by omitting *lacA* from the simulation for simplicity, students with a less clear sense of what constitutes metabolism may have chosen *lacA* simply because it did not occur in the simulation. In response to this assessment data, we have added more full discussion of the *lacA* to the tutorial, emphasizing this protein's role in metabolism and *lacI*'s role in regulation.

3.3 | Student attitudes toward the simulation

Although there are standardized tools to measure student engagement in biological education in general (e.g.,[19,20]), we could not find any published validated instrument to specifically measure engagement with a single simulation. Therefore, to assess student attitudes toward LacOP, we adapted a previously published survey first designed for a single molecular biology simulation for AP high school students.[7] The survey results (Figure 5) show generally favorable attitudes toward the LacOp simulation. For example, 77% of students who used the simulation reported they "liked" using it, compared to 18% who disliked using it. Most students (65%) found it easy or very easy to know what to do; though some students found it "kind of hard" (32%) or "hard" (2%) to know what to do. Because we wanted to pilot the use of the simulation as adaptable as either an in-class or take-home experience, in this first trial, we did not provide students with any verbal instructions or demonstration of the simulation. Either of these interventions could have an impact on student evaluations of the difficulty of use of the simulation. Furthermore, the subset(s) of students who expressed challenges in using the simulation may be uncomfortable with the complexity of the lac operon, biochemical pathways, or have a lower comfort with digital simulations or graphical representation of results.

Importantly, 96% of students felt that the results of the simulation can be trusted, which suggests that the simulation's performance reliably matched their prior expectations of the lac operon learned from their textbook and lecture. When asked to compare the simulation to other classroom activities, 22% of students ranked the simulation as "more fun," 47% ranked it "as much fun" and 31% said that it was "less fun" than other classroom activities.

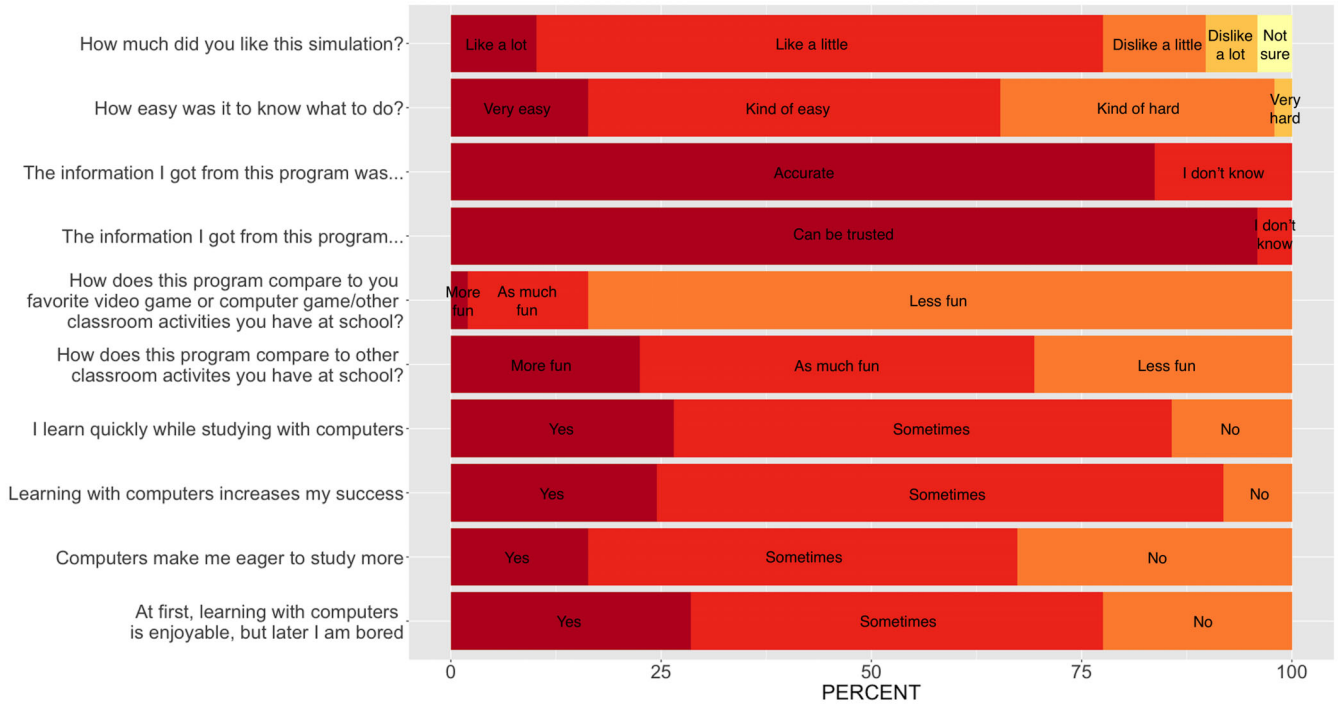

FIGURE 5 Student impressions of the simulation in the areas of usability attitudes toward computers. These questions were faithfully adapted from prior educational research on the value of biological simulations .[7] Figure made by ggPlot package in R

4 | CONCLUSION

The LacOp simulation engages students in many of the examples of core competencies specifically highlighted in Vision and Change[4] including hypothesis testing, the evaluation of experimental evidence, interpreting graphs, computational modeling of dynamic systems, managing and analyzing large datasets, applying physical laws to biological dynamics and the chemistry of biological systems, and it supports engagement with core concepts in genetics education.[11] LacOP has many features deemed "best practices" by a critical review of past electronic science simulations[21]: notably that is designed to be used as a supplement to traditional instruction, encourages reflection, and has high-quality support structures such as the tutorial and web interface. For this reason, and because of our own findings presented in this study, we believe it can be useful to those learning about gene regulation.

The LacOp simulation is free, can be easily from used on any device with internet connectivity and a browser. Using the tutorial that accompanies the simulation, students can test a range of mutants and metabolic conditions by clicking just a few buttons. These simple and straightforward features make the LacOp simulation a suitable companion to courses taught both online and in-person. With the recent SARS-CoV-2 global pandemic, many educational professionals found themselves looking for ways to engage their students remotely, and this simulation is one additional tool available to educators and students.

ACCESSING MATERIALS

Materials are available online at http://supergenetics.org/lacop.

ACKNOWLEDGMENTS

The authors would like to thank the other students enrolled in BIOL3770 who contributed their talent(s) to this work: Anna Miller, Joshua Stanfield, Tudi Li, Shawn Herring and Dylan Gnatz, as well as the SU CMOL and biochemistry faculty (Michelle DuBois, Glenn Yasuda, Mark Jordan, Michael Zanis, Brett Kaiser, Jennifer Loertscher, Kristin Hultgren) who provided peer review and resource support. We would like to thank students in the Fall 2019 Bio233 class at St. Olaf College for participating in the study. We would also like to thank Alexa Clemmons PhD, of the University of Washington and the broader faculty of UW-BERG for constructive guidance in educational methodology and manuscript preparation.

CONFLICTS OF INTEREST

Although these authors were involved in developing and producing this software, it is a free online resource, and the authors have no commercial interest in this software.

ORCID

Robert Rutherford https://orcid.org/0000-0002-5756-8073

REFERENCES

1. Freeman S, Eddy SJ, McDonough M, Smith MK, Okoroafor N, Jordt H, et al. Active learning boosts performance in STEM courses. Proc Natl Acad Sci. 2014;111:8410–5. https://doi.org/10.1073/pnas.1319030111
2. Mintz S. Maintaining the campus experience in the time of social distancing. Inside Higher Educ. 2020. https://www.insidehighered.com/blogs/higher-ed-gamma/maintaining-campus-experience-time-social-distancing
3. NRC. A framework for K-12 science education: practices, cross-cutting concepts and Core ideas. Washington, DC: National Academies Press; 2012.
4. American Association for the Advancement of Science. Vision and change in undergraduate biology education: a call to action. Washington, DC: AAAS; 2011.
5. White B, Bolker ED. Interactive computer simulations of genetics, biochemistry and molecular biology. Biochem Mol Biol Educ. 2008;36:77–84. https://doi.org/10.1002/bmb.20152
6. Gibbins S, Sosabowski M, Cunningham J. Evaluation of a web-based resource to support a molecular biology practical class—does computer-aided learning really work? Biochem. Mol. Biol. Educ. 2002;31:353–5.
7. Shegog R, Lazarus M, Murray KG, Diamond PM, Sessions N, Zsigmond E. Virtual transgenics: using a molecular biology simulation to impact student academic achievement and attitudes. Res Sci Educ. 2012;42:875–90. https://doi.org/10.1007/s11165-011-9216-7
8. https://phet.colorado.edu/sims/cheerpj/gene-network/latest/gene-network.html?simulation=gene-machine-lac-operon
9. Helikar T, Cutucache CE, Dahlquist LM, Herek TA, Larson JJ, Rogers JA. Integrating interactive computational modeling in biology curricula. PLoS Comput Biol. 2015;11:e1004131. https://doi.org/10.1371/journal.pcbi.1004131
10. Crowther A, Bergan-Roller HE, Galt NJ, Appleby L, Dauer JT, Helikar T. Discovering prokaryotic gene regulation by building and investigating a computational model of the lac operon. CourseSource. 2019;6. https://doi.org/10.24918/cs.2019.4
11. Genetics Society of America (2015) Genetics Learning Framework. https://genetics-gsa.org/education/genetics-learning-framework/
12. Libeskind-Hadas R, Bush E. Computing for biologists. New York: Cambridge University Press; 2014.
13. Stryer L, Berg J, Tymoczko J, Gatto G. Biochemistry. 9th ed. New York: MacMillan; 2019.
14. Santillán M. Bistable behavior in a model of the lac operon in *Escherichia coli* with variable growth rate. Biophys J. 2008;94:2065–81. https://doi.org/10.1529/biophysj.107.118026
15. Robeva R, Davies R, Hodge T, Enyedi A. Mathematical biology modules based on modern molecular biology and modern discrete mathematics. CBE Life Sci Educ. 2010;9:227–40. https://doi.org/10.1187/cbe.10-03-0019
16. Stefanski KM, Gardner GE, Seipelt-Thiemann RL. Development of a Lac operon concept inventory (LOCI). CBE Life Sci Educ. 2016;15:1–11. https://doi.org/10.1187/cbe.15-07-0162
17. Pierce BA. Genetics: a conceptual approach. 6th ed. New York: W.H.Freeman; 2016.
18. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol. 1961;3:318–56.
19. Kuh GD. Assessing what really matters to student learning inside the national survey of student engagement. Change. 2001;33:10–7.
20. Hopper MK. Assessment and comparison of student engagement in a variety of physiology courses. Adv Physiol Educ. 2016;40:70–8. https://doi.org/10.1152/advan.00129.2015
21. Smetana RLB. Computational simulations to support science instruction and learning: a critical review of the literature. Int J Sci Educ. 2012;34:1337–70.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Charczenko R, McMahon M, Kandl K, Rutherford R. LacOp: A free web-based lac operon simulation that enhances student learning of gene regulation concepts. Biochem Mol Biol Educ. 2022;50(4):360–8. https://doi.org/10.1002/bmb.21638