Full Title:

Genie: An interactive real-time simulation for teaching genetic drift

Authors and affiliations:

Andreina I. Castillo[1], Ben H. Roos[2], Michael S. Rosenberg[2], Reed A. Cartwright[2,3]*, and Melissa A. Wilson Sayres[2,3,4]*

1. School of Plant Sciences, The University of Arizona, Tucson, Arizona, 85721 USA
2. School of Life Sciences, Arizona State University, Tempe, Arizona, 85287-4501 USA
3. The Biodesign Institute, Arizona State University, Tempe, Arizona, 85287-5301 USA
4. Center for Evolution and Medicine, Arizona State University, Tempe, Arizona, 85287-4501 USA

* To whom correspondence may be addressed.
Email: melissa.wilsonsayres@asu.edu; cartwright@asu.edu

Keywords:

genetic drift, simulations, education,

IRB protocol:

STUDY00003707

Abstract

Background: One of the most challenging topics for students in Evolution courses is that of non-adaptive evolution, particularly genetic drift. Novel teaching techniques and software have been implemented to facilitate student understanding of this and other evolution related topics; nonetheless, some of them still present certain disadvantages. Here we introduce Genie, a web-based application designed to demonstrate population genetics and evolutionary concepts. **Results:** We used Genie as a tool to teach 203 students taking Arizona State University's Evolution course. Students freely used Genie during recitation session after having learned about genetic drift and other mechanism of non-adaptive evolution during lectures. Student performance and comprehension of genetic drift, and other evolutionary concepts, was tested with the Genetic Drift Inventory before and after using Genie. We found that Genie was an efficient tool for teaching genetic drift, mutation, the effects of barrier formation, and gene flow, across a variety of student demographics. Specifically, we found that with our implementation of Genie, students had significantly improved understanding of concepts such as: changes in alleles frequencies due to genetic drift and the difference between adaptive and non-adaptive evolutionary mechanisms. **Conclusions:** We believe that the easy usage, creativity, and real-time nature of Genie makes it an accessible tool for both teachers and students learning non-adaptive evolution, as well as a means for student development of creative and critical thinking. Genie is freely available (https://cartwrig.ht/apps/genie/) and can be easily accessed across different operating systems.

Background

Though frequently considered one of the most important evolutionary mechanisms, natural selection represents only one of many evolutionary forces that can act on a population. Evolution, or the change in allele frequencies over time, also occurs via several non-adaptive evolutionary processes such as gene flow (Ellstrand and Rieseberg 2016; Morjan and Rieseberg 2004), random mutation (Barton 2010; Sniegowski et al. 2000), and genetic drift (Andrews et al. 2012). Typically, teaching students about non-adaptive evolutionary forces can be difficult, particularly on the instance of genetic drift.

Regardless of a student's background and chosen major, the majority of college students have difficulties learning and retaining fundamental science concepts (Alters and Nelson 2002), leading to a decrease in comprehension of more advanced concepts that build on the basic ones. While teaching advanced scientific concepts is in itself challenging, an equally urgent task is aiding students to develop creative and critical thinking skills essential for science-based majors.

Instructor-provided situations that allow students to discuss, challenge, and test the adequacy of a concept have been shown to be effective in science teaching and learning (Slavich and Zimbardo 2012). Consequently, tools that facilitate free exploration of evolutionary concepts, including a short introduction by the instructor, are particularly useful. Numerous programs have been designed as stand-alone software to fill this niche in regards to evolutionary concepts (Hoban et al. 2012). Many of them focus in exploring a single evolutionary force, *e.g.* mutation (Haubold et al. 2010) or migration (Arenas and Posada 2007). However, the majority of this software typically requires some degree of installation and may not be compatible with multiple operating systems. Alternatively, web development technologies, and in particular the

71 programming language JavaScript, provide unique opportunities for creating computationally-
72 rich browser-based educational tools that are accessible across operating systems.

73 Here we developed a web application (Genie) designed to demonstrate several population
74 genetics and evolutionary concepts including genetic drift, gene flow, and random mutation. This
75 application conducts a real-time simulation of the change in allele frequencies in a finite
76 population of spatially isolated individuals. Using colors, the application allows students to
77 visualize changes in population over time and understand how those visual changes translate to
78 fluctuations in allele frequency, and eventually, fixation/loss of an allele.

79 We find that this web-based software is accessible to students and increases knowledge
80 of genetic drift concepts, as tested using a genetic drift inventory (Price et al. 2014). These types
81 of assessments have proven to be useful in capturing student's understanding of other complex
82 evolutionary concepts in the past (Perez et al. 2013). The Genie software requires no startup
83 other than navigating to a web page, thus making the use of programmed stochastic simulations
84 to demonstrate the concept of genetic drift practical and accessible to both educators and
85 students.

86

87 **Methods**

88 *Genie simulation program*

89 Genie (https://cartwrig.ht/apps/genie/) is a web-based, stochastic simulation app written
90 in JavaScript. The simulation uses a spatially explicit Moran Model (Nei et al. 1976) to describe
91 a finite population of 1,024 individuals on a 32 by 32 grid. Each individual is haploid with a
92 single locus. The locus mutates according to the infinite alleles model (Nei et al. 1976). Genie
93 works as follows:

- *Population Initialization.* The simulation begins when a population is randomly initialized according to Hoppe's Urn (Perez et al. 2013). Briefly, the population is created one individual at a time, and each individual either carries a new, unique allele or is a copy of a previously created individual. The probabilities that individual *i* carries a new allele is $\theta/(\theta+i-1)$ and the probability that the individual carries a copy allele is $(i-1)/(\theta+i-1)$, where $\theta = 2N\mu$, N is the population size and μ is the mutation rate. Each individual that carries a copy allele is chosen uniformly from the previously initialized individuals. As a default, at initialization a $\mu=0.001$ is selected to ensure diversity within the initial population, but then the mutation rate each generation is 0 unless otherwise modified by the user.
- *Algorithm.* At each step of the simulation, a randomly selected individual dies, leaving its corresponding cell momentarily empty. A parent allele is then randomly selected from the eight immediate neighboring cells (both adjacent and diagonal). Cells on the edges and corners of the simulation have fewer neighbors than internal cells, causing a small edge effect. The probability that a new individual will have the same allele as its parent its $1-\mu$, and the probability that an individual has a new, unique allele is μ. Each "generation" consists of 2000 death/birth steps after which the population is redrawn in the visualization window.
- *Running*. The application contains four components: a grid, where the population is displayed (Fig. 1-1a); a control panel, where users can manipulate the simulation's mutation parameter (Fig. 1-1b); an upper graph, where users can see the number of alleles in the population at any given time (Fig. 1-1c); and a lower graph, where users can see the frequency of different alleles at any given time (Fig. 1-1d). Each initial allele is assigned one of 18 basic colors, while each mutant allele is assigned one of six neon colors. Unless changed, the default

mutation rate while the simulation is running is zero. A single button allows users to toggle between starting the simulation or pausing it. A reset button allows users to restart and reinitialize the simulation at any point.

- *Barriers*. Users also have the ability to create a barrier in the population grid. To do so, users can alter a cell (by clicking on it) or alter a set of cells (by clicking and dragging the cursor to select multiple cells). When a barrier is created, the color associated with the cell changes to black. Barriers act neither as parent cells (they are never replicated) nor die to be subsequently replaced. Thus, for each created barrier cell the total population size declines by one. By building barriers, users can construct physical constraints that restrict the movement of alleles between subpopulations. Barriers can be used to create subpopulations of different size and shape, as well as to study the effects of corridors on gene flow. Barriers can be removed by clicking on the chosen cell(s) a second time; this will set the cell color to white and designate the cell as unoccupied. Neighboring cells will replicate into unoccupied cells; unoccupied cells cannot serve as a parent of a neighboring cell.
- *Forced Mutation*. Users can force a mutation to occur in a manner similar to creating barriers. Cells can be mutated by holding the SHIFT button while clicking the cell, or while clicking and dragging the cursor across several cells. Forcing a mutation immediately creates a new, unique allele in each of the chosen cell(s).
- *Graphs*. Two graphs are displayed to the right of the grid as described above. Both graphs update in real time as the simulation runs.

Assessment

139 The impact of Genie as a tool for teaching concepts of genetic drift was evaluated using the
140 Genetic Drift Inventory (Price et al. 2014). All research was reviewed and approved by Arizona
141 State University's IRB protocol STUDY00003707.

142 The inventory was used without changes in pre-and post-lesson assessments. The pre-
143 lesson assessment was posted online on Blackboard two days before the class lesson (recitation).
144 Students were asked to answer all questions individually by 3:00 pm the day of the in-class
145 activity (described below, Recitation activity). All students were allowed the same amount of
146 time to complete the assessment.

147 Students were divided as follows: (a) by recitation start times (3:00 pm, 4:30 pm, 6:00
148 pm, and 7:00 pm); and (b) by Teaching Assistants (TA) pairs. Each recitation was taught by one
149 lead TA and one assistant TA; henceforth, the TA pairs will be referred as TA pair 1 and TA pair
150 2. Overall, the class was divided into 8 groups of roughly equal size. No more than 48 students
151 were allowed to participate per recitation session. The recitation was co-designed by AC, MR,
152 and MAW. The post-lesson assessment was posted on Blackboard at 9:00 pm after the last
153 recitation session ended. Students had two days to individually complete the assessment (same
154 amount of time as the pre-lesson assessment).

155 At the end of the semester, students were given the opportunity to opt-in to the study of
156 their pre- and post-lesson assessments (considered as homework for the entire class) and final
157 course grade. In the present study we report pre-lesson scores, post-lesson scores, and final
158 scores in the course. In addition, students were also requested to report: gender, first generation
159 college student status, race/ethnicity, and whether they had taken a genetics course (BIO340) at
160 the same institution (Additional file 1).

Welch Two Sample t-tests were used to evaluate the statistical significance of the change in pre- and post-lesson assessment scores within demographic classes, and recitation sections. All statistical analyses and the associated figures (Fig. 2, Additional files 2 and 3) can be regenerated using custom-made R scripts (Additional file 1). All data and code used in these analyses can be found at DOI: 10.5281/zenodo.1158033.

Recitation activity

The basic concept of non-adaptive evolution, and specifically of genetic drift, was illustrated for all sessions at the start of the recitation class. Then, the basic features, display, and usability of the Genie software were explained to students. Questions designed to facilitate student discussion and interpretation of Genie simulation results were provided alongside images of the Genie output (Fig. 1 1-4). The recitation slides (Additional file 4) were made available to all students after all recitation sessions concluded (9:00 pm). Overall, four activities were conducted in all recitation sessions:

Activity 1: Defaults parameters/settings.

In the first activity, students were instructed to run Genie without modifying any parameters or creating any barriers. As the number of generations increased, students kept track of the changes in the number of alleles in the population and the allele frequencies. Students made conjectures on the distribution of haplotypes in the population by tracking variations in the colors patterns (alleles) shown in the population grid. The mutation rate was not modified; however, students were instructed to be on the lookout for any new alleles arising at any point of the simulation. The simulation ran until one allele reached fixation, students were instructed to keep track of the generation at which this occurred.

185

186 *Activity 2: Effects of absolute barriers on genetic drift and gene flow.*

187 The second activity introduced the concept of barrier formation in the population grid.
188 This activity allowed students to identify the effects of genetic drift simultaneously with those of
189 population isolation. The simulation was re-started and students were instructed to create two
190 barriers reaching opposite borders of the population grid (one horizontal and one vertical). This
191 setup resulted in four completely isolated populations of roughly equal size (Additional file 4.
192 No modifications in the mutation rate were introduced. Students kept track of variations in the
193 colors patterns (alleles) shown in the population grid, and changes in number of alleles and allele
194 frequency in the overall population. Additionally, students kept track of the allele number and
195 distribution of alleles in each of the four independent sections/populations. The simulation
196 continued until one allele became fixed in each subsection/subpopulation.

197 After one allele became fixed in each of the four subsections/subpopulations, students
198 were instructed to pause the simulation and create a corridor by removing part of the barrier
199 between two or more sub-areas, and then unpause the simulation. Students kept track of changes
200 in number of alleles and allele frequency, as well as the movement of alleles between connected
201 sections/populations. The simulation ran until one allele became fixed between the sections with
202 barriers removed. The number of generations for an allele to become fixed amongst independent
203 sections/populations was recorded.

204

205 *Activity 3: Effects of partial barriers and corridors on genetic drift and gene flow.*

206 The third activity was designed to further explore the effects of barrier formation in
207 genetic drift. Students were instructed to restart the simulation and create barriers that entirely

separated the population grid into four sections of roughly equal size. Before the simulation started, students formed a corridor by removing a portion of the barriers (Additional file 4). This setting allowed for gene flow to occur between sections/populations from the beginning of the simulation and before any allele reached fixation. Students tracked the changes in number and allele frequency between: (1) completely isolated sections/populations; and (2) sections/populations connected by the corridor. Students were instructed to compare the flow of alleles across the corridor with that observed in Activity 2. Additionally, students also recorded the number of generation until fixation was reached in connected and isolated areas. The mutation rate was not modified in this activity.

Activity 4: Effects of mutation rate on genetic drift

The fourth activity centered in evaluating the effects of changes in mutation rate along those of genetic drift. Students were instructed to restart the simulation, increase the mutation rate, and take note of the changes in the population grid and accompanying graphs. Alternatively, students were instructed to perform the activity while markedly reducing the mutation rate. Students also kept track of changes in the number of alleles, allele frequency, and the number of generations until the point of fixation of a single allele. No barriers were created on the population grid.

In-class interpretation

After the entire recitation section completed the four main activities, students were allowed to freely explore other potential outcomes of genetic drift. Students freely modified the population landscape by creating various types of barriers and/or changing the mutation rate. To better guide students into examining important genetic drift concepts, and to better enhance

discussion among class members, a series of suggestions activities/questions were provided (Additional file 4). The suggested prompts included:

- Evaluate the effects of creating barriers of different size and shape.
- Assess the effects of genetic drift on different population sizes.
- Discern the effects of genetic drift on allele diversity within a single population, and between isolated populations.
- Observe the effects that creating corridors with different size and shapes have on gene flow.
- Evaluate the effects of creating corridors and barriers at different points of the simulation.
- Track the effects of modifying the mutation rate at different points of the simulation.

Results

Increased understanding of genetic drift across demographics

The number of respondents in each demographic/classification assessed in this study is reported (Fig. 3). Out of 22 questions in the genetic drift inventory, the mean correct answer increased significantly from 14.18 on the pre-lesson assessment to 16.46 on the post-lesson assessment (Fig. 2a; Methods). In fact, we see significant (p-value <0.05) improvements in understanding of genetic drift concepts across all classifications from the pre-lesson to the post-lesson assessments (Fig. 2b-f). When we look across demographics, we observe that there is no significant difference in pre-lesson or post-lesson performance by gender (Fig. 2b; Additional files 2 and 3). On the other hand, while there were some differences on either the post-lesson or pre-lesson assessments by first generation status or by race/ethnicity as recorded here (Fig. 2b

and 2c; Additional files 2 and 3), the most significant difference occurred between students with different final grades in the overall course (Fig. 2e; Additional files 2 and 3).

Additionally, we observed improvement in understanding genetic drift concepts in all recitation sessions, with the exception of TA Pair's 1 7:30pm class (Fig. 2f; Additional files 2 and 3). This was the recitation session with the fewest number of students in the class (Fig. 3f).

Question by question breakup

The top three questions with improved student outcomes were questions 13, 3, and 17 (Table 1). Questions 13 and 3 both evaluated concepts related to the fixation of alleles and loss of alleles via genetic drift. In contrast, Q17 asked if one allele (or feature) would increase in the population due to genetic drift. Many students also improved their scores on questions 4, 5, 6, 11, and 12. These questions assessed different aspects of genetic drift and natural selection as unique evolutionary processes with specific outcomes (Q5, Q6, and Q12); as well as the significance of isolated and small populations on the fixation of traits (Q4 and Q11). It is worth noting that some students switched their answer from correct to incorrect on Q19 (new mutation occurring during genetic drift), Q21 (gene flow aiding on the spread of a disadvantageous trait), and Q16 (chance and selection playing a role in some, but not all, generations), suggesting that there is room for improvement in our simulation or instructions relating to that simulation (Table 1).

Discussion

Increased understanding of genetic drift across demographics

In this lesson, we used Genie to improve student understanding of non-adaptive evolutionary mechanisms. By using Genie, students were able to observe: (a) changes in allele frequencies through time, and (b) variation in the number of alleles within a population. Both aspects were simulated using our web-based dynamic computer application. In-class activities (Additional file 4) were developed with the objective of illustrating the change in allele frequencies solely as the product of genetic drift, gene flow, or mutation. In addition, students were allowed to freely explore the Genie software, coming up with and developing their own activities to explore genetic drift related concepts. Furthermore, students were encouraged to follow activities tailored to evaluate the effects of genetic drift in combination with those of barrier formation and change of the mutation rate (Additional file 4).

The most significant differences were observed amongst students with different final course grades (Figs. 2e and 3d), however, all students improved their score regardless of their letter grade, showing that Genie was effective in aiding students with various performance levels on the class (Fig. 2e). On the other hand, we found that the efficiency of Genie varied slightly across all the demographic classifications included in our study (Fig. 2b-d); particularly between first and non-first generation students (Fig. 2c; Additional file 2 and 3). While these slight variations should be considered in future classes, they do not seem to indicate a differential effectiveness of Genie as a teaching tool. Finally, we found that all recitation sessions showed some level of improvement in understanding of genetic drift and related concepts (Fig. 2f). Moreover, classes taught by different teams (TA pairs 1 and 2) did not show significant differences in comprehension of class concepts (Fig. 2f), suggesting that the overall course design may be successful when taught by other instructors. Only one recitation session (TA Pair 1 7:30pm class) did not show significant improvement in the post-lesson assessment; however,

since this session was comprised by the fewest number of students (Fig. 3f), it is possible that the lack of significance in our results is related to the low sample size.

Question by question breakup

Genie was particularly effective in helping students understand concepts related to the loss of alleles due to genetic drift (Table 1; Q3 and Q13), and concepts related with the change in allele frequency occurring via mechanisms other than natural selection (Table 1; Q17). These results suggest that our lessons especially helped increase students' understanding of the following two concepts: (1) loss of alleles occurring due to genetic drift, and (2) that allele frequencies can change independently of natural selection. This is likely a result of Genie's capabilities to generate a dynamic simulation of the variations in allele frequencies coordinated to the changes in the population grid. Students also improved their understanding of genetic drift and natural selection as two different evolutionary processes after using Genie (Table 1; Q6). Moreover, student capacity to define the distinct effects of natural selection and genetic drift on isolated (Table 1; Q11 and Q12) and reduced size populations (Table 1; Q4) also improved. It is likely that these concepts were better grasped thanks to the free-hand nature of barrier formation provided by Genie. Specifically, students freely explored the effects of complete or partial population isolation at different stages of the simulation (Additional file 4); hence, they were able to fully discover the effect of population size and different levels of population isolation on the strength of genetic drift.

Overall, we believe that the intuitive and free modification of the population grid, with little to no hard-coded numbers, is one of the most powerful features of Genie. This feature permitted students to explore genetic drift and related concepts to their own pace, design their

own experiments to test their hypotheses, and discuss their results among peers. Thus, Genie was effective not only in teaching students the concept related with genetic drift, but also in providing a mean for them to hone their creative thinking and reasoning skills.

We did not improve student understanding of all concepts related to genetic drift. Upon introspection we propose that this is due to our lecture design and unlikely to be due to the nature of Genie; regardless, further testing is required. Specifically, we found that students' comprehension on the role of novel mutations on a population was lowered after our lesson (Table 1, Q19). It is possible that this is an unintended consequence of the in-class activities designed to modify the mutation rate (See Methods - Activity 4). Briefly, by increasing the mutation rate students observed new alleles arising on the population and potentially reaching fixation, this might have been misinterpreted as new mutations arising due to genetic drift and not due to the change of the mutation rate itself. In order to address this issue, the effects of changing the mutation rate on allele diversity should be explored in more detail in future classes. In particular, we will develop discussion questions that clarify that the changes in mutation rate occur independently from those of genetic drift.

In addition, students did not have an increased understanding of natural selection across generations (Table 1; Q16); mainly, students were confused about the number of generations in which natural selection and random chance act. While this misconception cannot be addressed on the Genie simulation itself, it should be included in future pre-recitation activities as a pre-emptive discussion. Finally, student grasp of concepts related to the change in frequency of disadvantageous traits, as a result of genetic drift and gene flow (Table 1; Q21), also decreased after our lesson. This is likely the result of students not understanding that genetic drift and natural selection are different evolutionary mechanisms, or not understanding their combined

effects on a single population. To address this issue, open discussion questions on the interaction between evolutionary mechanisms will be incorporated at the end of future recitations.

Genie compared to other software

There are numerous software packages capable of generating genetic drift simulations; many of them include an ample array of parameters to be modified by the user (http://evolution.gs.washington.edu/popgen/popg.html). Nonetheless, more often than not, these programs need to be locally installed, can be difficult to execute across diverse platforms and operating systems, and might be negatively affected by hardware limitations and system updates. Since these issues can be avoided using web-based platforms, there have been numerous online tools developed for teaching purposes.

An ample set of web-based genetic drift simulators have been created by diverse groups and can be found publicly available online. While each of these may be an efficient teaching tool in their own regard, they each have certain shortcomings compared to Genie. For one, most genetic drift simulators display a static model of allele frequency variation (http://www.biology.arizona.edu/evolution/act/drift/drift.html). We believe that these static images make it difficult for students to grasp genetic drift as a random and ongoing process, and thus, a dynamic display such as that provided by Genie should make for a more effective teaching tool. Contrary to other web-based simulators (https://cartwrig.ht/apps/redlynx/), Genie focuses on genetic drift as the main acting evolutionary force. While evolution of real populations is the product of combined factors, students who are being introduced to non-adaptive evolution might have difficulty understanding the complexity of these interactions, and might prefer to focus their attention on the more familiar effects of natural selection. By aiming

370 the teaching experience only to non-adaptive evolutionary forces with Genie, students can fully
371 appreciate how evolution occurs in the absence of natural selection. There are other online
372 simulators that also provide a dynamic interface (http://phyletica.org/teaching/drift-simulator/).
373 This is an improvement in capturing the unique patterns observed in genetic drift; nonetheless,
374 these simulators often display variation of a single allele. Maintaining focus on a single allele can
375 be a major limitation in showing the role of genetic drift on allele diversity, a concept that most
376 students have difficulty grasping. Moreover, such representations are inaccurate in modeling the
377 effects on genetic drift in real populations, which are—with the exception of clonal
378 populations—likely composed of multiple alleles. As a result, we consider Genie's capacity to
379 dynamically and simultaneously simulate multiple alleles a significant feature compared to other
380 online teaching tools.
381
382 **Conclusion**
383 Genie is a unique tool to facilitate the demonstration of the concepts of genetic drift,
384 population isolation, gene flow, and genetic mutation to a large and diverse group of students.
385 Additionally, Genie's implementation in JavaScript allows it to be run from virtually any modern
386 computer and smart-phone, giving students the ability to use the tool on their own to either
387 explore these mechanisms or to complete assignments. Moreover, because the tool provides few
388 options for students to adjust input parameters, assignments need not include lengthy tutorials or
389 instructions. The primary feature of Genie is the dynamic visualization of population and non-
390 adaptive evolutionary mechanisms aimed to improve understanding of challenging biological
391 notions. Furthermore, by having students develop and come up with ways to test their own
392 hypotheses, Genie provides an easy and engaging tool for future scientists to practice and

develop their critical thinking without having to create specific in-class activities for this purpose. Overall, we believe that Genie is an effective tool for teaching genetic drift and related concepts, as well as for developing comprehensive scientific skills.

Table 1. Distribution of student's answers and change by question.

Question	Changed to		Differences	No change	
	Correct	Incorrect		Correct	Incorrect
Q13	86	9	77	90	18
Q3	53	7	46	133	10
Q17	58	12	46	84	49
Q12	60	22	38	85	36
Q5	51	15	36	99	38
Q6	53	19	34	62	69
Q11	50	17	33	119	17
Q4	44	13	31	118	28
Q8	46	22	24	64	71
Q20	40	18	22	123	22
Q18	29	8	21	154	12
Q2	41	27	14	96	39
Q14	46	32	14	64	61
Q7	32	19	13	142	10
Q15	26	15	11	153	9
Q10	28	18	10	136	21
Q1	22	14	8	159	8
Q9	28	21	7	146	8
Q22	30	30	0	98	45
Q19	36	39	-3	100	28
Q21	24	28	-4	131	20
Q16	28	43	-15	74	58

Distribution of student's answers and change by question. Student's answers showed different levels of improvement or deterioration of concepts' understanding across questions and topics.

Figure Legends

Figure 1. Genie's layout is intuitive and easy to use. 1-4 Shows a time-lapse of the Genie simulation (starting with the initial population up until generation ~335). **a.** Population grid showing the diversity of alleles in the population; **b.** Control panel with which users can alter the mutation rate to be used during the simulation; **c.** Number of unique alleles in the population over time; **d.** Allele frequencies of each unique allele in the population over time.

Figure 2. Student's assessment scores improved after teaching a genetic drift recitation class with Genie. a. Overall pre- and post-lesson assessment scores in the entire class, **b.** Students divided by reported gender; **c.** Student divided by reported first generation in college; **d.** Student's divided by ethnicity (non-white students have been grouped in the POC category); **e.** Student's divided by final letter grade; **f.** Students divided by recitation session in TA pair 1 and 2 sessions. Error bars represent standard error of the mean for each described group.

Figure 3. Sample adequately represents distinct demographics. a. Students divided by reported gender; **b.** Student divided by reported first generation in college; **c.** Student's divided by ethnicity (non-white students have been grouped in the POC category); **d.** Student's divided by final letter grade; **e.** Students divided by recitation session in TA pair 1 sessions; **f.** Students divided by recitation session in TA pair 2 sessions.

Additional files descriptions

Additional file 1. ReadMe of code for replicating analysis. All code for replicating analysis is available here and on GitHub, along with the de-identified data used for analysis.

Additional file 2. Score variation of pre and post-lesson assignment. Within group variations are represented by pre and post-lesson p-values, and their differences, between all evaluated demographic groups.

Additional file 3. Mean score variation of pre and post-lesson assignment. Within group variations are represented by pre and post-lesson mean values. Significant differences within demographic groups are represented by p-values.

Additional file 4. Slides provided during the recitation activity.

Declarations

- Ethics approval and consent to participate

IRB protocol approval: STUDY00003707

- Availability of data and materials

Full data set and R code used in the analysis are available at DOI: 10.5281/zenodo.1158033
Genie is freely available at https://cartwrig.ht/apps/genie/

- Competing interests

The authors declare no competing interests.

- Funding

Startup from the School of Life Sciences and the Biodesign Institute to MAW.

- Authors' contributions

AC and MAW wrote the manuscript and performed data analysis. The recitation was co-designed by AC, MR, and MAW. Genie was developed by BHR and RAC. AC, MR, RAC, and MAW edited the manuscript.

- Acknowledgements

The authors would like to thank to the students of the BIO345 Evolution course at Arizona State University for participating in this study.

References

Alters BJ, Nelson CE. Teaching evolution in higher education. Evolution (N. Y). 2002;56(10):1891–901.

Andrews TM, Price RM, Mead LS, McElhinny TL, Thanukos A, Perez KE, et al. Biology undergraduates' misconceptions about genetic drift. CBE Life Sci. Educ. 2012;11(3):248–59.

Arenas M, Posada D. Recodon: Coalescent simulation of coding DNA sequences with recombination, migration and demography. BMC Bioinformatics. 2007;8(1):458.

Barton NH. Mutation and the evolution of recombination. Philos. Trans. R. Soc. B Biol. Sci. 2010;365(1544):1281–94.

Ellstrand NC, Rieseberg LH. When gene flow really matters: gene flow in applied evolutionary biology. Evol. Appl. 2016;9(7):833–6.

Haubold B, Pfaffelhuber P, Lynch M. MlRho - A program for estimating the population mutation and recombination rates from shotgunsequenced diploid genomes. Mol. Ecol. 2010;19(SUPPL. 1):277–84.

Hoban S, Bertorelle G, Gaggiotti OE. Computer simulations: tools for population and evolutionary genetics. Nat. Rev. Genet. Nature Publishing Group; 2012;13(2):110–22.

Morjan CL, Rieseberg LH. How species evolve collectively: Implications of gene flow and selection for the spread of advantageous alleles. Mol. Ecol. 2004;13(6):1341–56.

Nei M, Chakraborty R, Fuerst PA. Infinite allele model with varying mutation rate. Genetics. 1976;73(11):4164–8.

Perez KE, Hiatt A, Davis GK, Trujillo C, French DP, Terry M, et al. The evodevoci: A concept inventory for gauging students' understanding of evolutionary developmental biology. CBE Life Sci. Educ. 2013;12(4):665–75.

Price RM, Andrews TC, McElhinny TL, Mead LS, Abraham JK, Thanukos A, et al. The genetic drift inventory: A tool formeasuring what advanced undergraduates havemastered about genetic drift. CBE Life Sci. Educ. 2014;13(1):65–75.

Slavich GM, Zimbardo PG. Transformational Teaching: Theoretical Underpinnings, Basic Principles, and Core Methods. Educ. Psychol. Rev. 2012;24(4):569–608.

Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. The evolution of mutation rates: Separating causes from consequences. BioEssays. 2000;22(12):1057–66.