

1 **Full Title:**

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

3

4 **Authors and affiliations:**

5 Andreina I. Castillo¹, Ben H. Roos², Michael S. Rosenberg², Reed A. Cartwright^{2,3*}, and Melissa
6 A. Wilson Sayres^{2,3,4*}

7

8 1. School of Plant Sciences, The University of Arizona, Tucson, Arizona, 85721 USA

9 2. School of Life Sciences, Arizona State University, Tempe, Arizona, 85287-4501 USA

10 3. The Biodesign Institute, Arizona State University, Tempe, Arizona, 85287-5301 USA

11 4. Center for Evolution and Medicine, Arizona State University, Tempe, Arizona, 85287-4501
12 USA

13

14

15

16 * To whom correspondence may be addressed.

17 Email: melissa.wilsonsayres@asu.edu; cartwright@asu.edu

18

19

20 **Keywords:**

21 genetic drift, simulations, education,

22

23 **IRB protocol:**

24 STUDY00003707

25

26 **Abstract**

27 **Background:** One of the most challenging topics for students in Evolution courses is that
28 of non-adaptive evolution, particularly genetic drift. Novel teaching techniques and software
29 have been implemented to facilitate student understanding of this and other evolution related
30 topics; nonetheless, some of them still present certain disadvantages. Here we introduce Genie, a
31 web-based application designed to demonstrate population genetics and evolutionary concepts.

32 **Results:** We used Genie as a tool to teach 203 students taking Arizona State University's
33 Evolution course. Students freely used Genie during recitation session after having learned about
34 genetic drift and other mechanism of non-adaptive evolution during lectures. Student
35 performance and comprehension of genetic drift, and other evolutionary concepts, was tested
36 with the Genetic Drift Inventory before and after using Genie. We found that Genie was an
37 efficient tool for teaching genetic drift, mutation, the effects of barrier formation, and gene flow,
38 across a variety of student demographics. Specifically, we found that with our implementation of
39 Genie, students had significantly improved understanding of concepts such as: changes in alleles
40 frequencies due to genetic drift and the difference between adaptive and non-adaptive
41 evolutionary mechanisms. **Conclusions:** We believe that the easy usage, creativity, and real-time
42 nature of Genie makes it an accessible tool for both teachers and students learning non-adaptive
43 evolution, as well as a means for student development of creative and critical thinking.
44 Genie is freely available (<https://cartwrig.ht/apps/genie/>) and can be easily accessed across
45 different operating systems.

46

47

48

49 **Background**

50 Though frequently considered one of the most important evolutionary mechanisms,
51 natural selection represents only one of many evolutionary forces that can act on a population.
52 Evolution, or the change in allele frequencies over time, also occurs via several non-adaptive
53 evolutionary processes such as gene flow (Ellstrand and Rieseberg 2016; Morjan and Rieseberg
54 2004), random mutation (Barton 2010; Sniegowski et al. 2000), and genetic drift (Andrews et al.
55 2012). Typically, teaching students about non-adaptive evolutionary forces can be difficult,
56 particularly on the instance of genetic drift.

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

62 Instructor-provided situations that allow students to discuss, challenge, and test the
63 adequacy of a concept have been shown to be effective in science teaching and learning (Slavich
64 and Zimbardo 2012). Consequently, tools that facilitate free exploration of evolutionary
65 concepts, including a short introduction by the instructor, are particularly useful. Numerous
66 programs have been designed as stand-alone software to fill this niche in regards to evolutionary
67 concepts (Hoban et al. 2012). Many of them focus in exploring a single evolutionary force, *e.g.*
68 mutation (Haubold et al. 2010) or migration (Arenas and Posada 2007). However, the majority of
69 this software typically requires some degree of installation and may not be compatible with
70 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:

- 94 ● *Population Initialization.* The simulation begins when a population is randomly initialized
95 according to Hoppe's Urn (Perez et al. 2013). Briefly, the population is created one
96 individual at a time, and each individual either carries a new, unique allele or is a copy of a
97 previously created individual. The probabilities that individual i carries a new allele is
98 $\theta/(\theta+i-1)$ and the probability that the individual carries a copy allele is $(i-1)/(\theta+i-1)$, where θ
99 $= 2N\mu$, N is the population size and μ is the mutation rate. Each individual that carries a copy
100 allele is chosen uniformly from the previously initialized individuals. As a default, at
101 initialization a $\mu=0.001$ is selected to ensure diversity within the initial population, but then
102 the mutation rate each generation is 0 unless otherwise modified by the user.
- 103 ● *Algorithm.* At each step of the simulation, a randomly selected individual dies, leaving its
104 corresponding cell momentarily empty. A parent allele is then randomly selected from the
105 eight immediate neighboring cells (both adjacent and diagonal). Cells on the edges and
106 corners of the simulation have fewer neighbors than internal cells, causing a small edge
107 effect. The probability that a new individual will have the same allele as its parent is $1-\mu$,
108 and the probability that an individual has a new, unique allele is μ . Each "generation"
109 consists of 2000 death/birth steps after which the population is redrawn in the visualization
110 window.
- 111 ● *Running.* The application contains four components: a grid, where the population is displayed
112 (Fig. 1-1a); a control panel, where users can manipulate the simulation's mutation parameter
113 (Fig. 1-1b); an upper graph, where users can see the number of alleles in the population at
114 any given time (Fig. 1-1c); and a lower graph, where users can see the frequency of different
115 alleles at any given time (Fig. 1-1d). Each initial allele is assigned one of 18 basic colors,
116 while each mutant allele is assigned one of six neon colors. Unless changed, the default

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

120 ● *Barriers*. Users also have the ability to create a barrier in the population grid. To do so, users
121 can alter a cell (by clicking on it) or alter a set of cells (by clicking and dragging the cursor to
122 select multiple cells). When a barrier is created, the color associated with the cell changes to
123 black. Barriers act neither as parent cells (they are never replicated) nor die to be
124 subsequently replaced. Thus, for each created barrier cell the total population size declines by
125 one. By building barriers, users can construct physical constraints that restrict the movement
126 of alleles between subpopulations. Barriers can be used to create subpopulations of different
127 size and shape, as well as to study the effects of corridors on gene flow. Barriers can be
128 removed by clicking on the chosen cell(s) a second time; this will set the cell color to white
129 and designate the cell as unoccupied. Neighboring cells will replicate into unoccupied cells;
130 unoccupied cells cannot serve as a parent of a neighboring cell.

131 ● *Forced Mutation*. Users can force a mutation to occur in a manner similar to creating
132 barriers. Cells can be mutated by holding the SHIFT button while clicking the cell, or while
133 clicking and dragging the cursor across several cells. Forcing a mutation immediately creates
134 a new, unique allele in each of the chosen cell(s).

135 ● *Graphs*. Two graphs are displayed to the right of the grid as described above. Both graphs
136 update in real time as the simulation runs.

137

138 *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).

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

166
167 *Recitation activity*

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

175
176 *Activity 1: Defaults parameters/settings.*

177 In the first activity, students were instructed to run Genie without modifying any
178 parameters or creating any barriers. As the number of generations increased, students kept track
179 of the changes in the number of alleles in the population and the allele frequencies. Students
180 made conjectures on the distribution of haplotypes in the population by tracking variations in the
181 colors patterns (alleles) shown in the population grid. The mutation rate was not modified;
182 however, students were instructed to be on the lookout for any new alleles arising at any point of
183 the simulation. The simulation ran until one allele reached fixation, students were instructed to
184 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

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

217

218 *Activity 4: Effects of mutation rate on genetic drift*

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

226

227 *In-class interpretation*

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

232 discussion among class members, a series of suggestions activities/questions were provided

233 (Additional file 4). The suggested prompts included:

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

242

243

244 **Results**

245 *Increased understanding of genetic drift across demographics*

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

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

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

260

261 *Question by question breakup*

262 The top three questions with improved student outcomes were questions 13, 3, and 17
263 (Table 1). Questions 13 and 3 both evaluated concepts related to the fixation of alleles and loss
264 of alleles via genetic drift. In contrast, Q17 asked if one allele (or feature) would increase in the
265 population due to genetic drift. Many students also improved their scores on questions 4, 5, 6,
266 11, and 12. These questions assessed different aspects of genetic drift and natural selection as
267 unique evolutionary processes with specific outcomes (Q5, Q6, and Q12); as well as the
268 significance of isolated and small populations on the fixation of traits (Q4 and Q11). It is worth
269 noting that some students switched their answer from correct to incorrect on Q19 (new mutation
270 occurring during genetic drift), Q21 (gene flow aiding on the spread of a disadvantageous trait),
271 and Q16 (chance and selection playing a role in some, but not all, generations), suggesting that
272 there is room for improvement in our simulation or instructions relating to that simulation (Table
273 1).

274

275

276 **Discussion**

277 *Increased understanding of genetic drift across demographics*

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

288 The most significant differences were observed amongst students with different final
289 course grades (Figs. 2e and 3d), however, all students improved their score regardless of their
290 letter grade, showing that Genie was effective in aiding students with various performance levels
291 on the class (Fig. 2e). On the other hand, we found that the efficiency of Genie varied slightly
292 across all the demographic classifications included in our study (Fig. 2b-d); particularly between
293 first and non-first generation students (Fig. 2c; Additional file 2 and 3). While these slight
294 variations should be considered in future classes, they do not seem to indicate a differential
295 effectiveness of Genie as a teaching tool. Finally, we found that all recitation sessions showed
296 some level of improvement in understanding of genetic drift and related concepts (Fig. 2f).
297 Moreover, classes taught by different teams (TA pairs 1 and 2) did not show significant
298 differences in comprehension of class concepts (Fig. 2f), suggesting that the overall course
299 design may be successful when taught by other instructors. Only one recitation session (TA Pair
300 1 7:30pm class) did not show significant improvement in the post-lesson assessment; however,

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

303

304 *Question by question breakup*

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

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

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

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

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

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

349

350 *Genie compared to other software*

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

358 An ample set of web-based genetic drift simulators have been created by diverse groups
359 and can be found publicly available online. While each of these may be an efficient teaching tool
360 in their own regard, they each have certain shortcomings compared to Genie. For one, most
361 genetic drift simulators display a static model of allele frequency variation
362 (<http://www.biology.arizona.edu/evolution/act/drift/drift.html>). We believe that these static
363 images make it difficult for students to grasp genetic drift as a random and ongoing process, and
364 thus, a dynamic display such as that provided by Genie should make for a more effective
365 teaching tool. Contrary to other web-based simulators (<https://cartwrig.ht/apps/redlynx/>), Genie
366 focuses on genetic drift as the main acting evolutionary force. While evolution of real
367 populations is the product of combined factors, students who are being introduced to non-
368 adaptive evolution might have difficulty understanding the complexity of these interactions, and
369 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

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

396
 397
 398
 399

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

400 Distribution of student’s answers and change by question. Student’s answers showed different
 401 levels of improvement or deterioration of concepts’ understanding across questions and topics.
 402
 403

404 **Figure Legends**

405

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

411

412 **Figure 2. Student's assessment scores improved after teaching a genetic drift recitation**
413 **class with Genie.** **a.** Overall pre- and post-lesson assessment scores in the entire class, **b.**
414 Students divided by reported gender; **c.** Student divided by reported first generation in college; **d.**
415 Student's divided by ethnicity (non-white students have been grouped in the POC category); **e.**
416 Student's divided by final letter grade; **f.** Students divided by recitation session in TA pair 1 and
417 2 sessions. Error bars represent standard error of the mean for each described group.

418

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

424

425 **Additional files descriptions**

426

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

429

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

433

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

437

438 **Additional file 4. Slides provided during the recitation activity.**

439 **Declarations**

- 440 • Ethics approval and consent to participate

441 IRB protocol approval: STUDY00003707

442

- 443 • Availability of data and materials

444 Full data set and R code used in the analysis are available at DOI: 10.5281/zenodo.1158033

445 Genie is freely available at <https://cartwrig.ht/apps/genie/>

446

- 447 • Competing interests

448 The authors declare no competing interests.

449

- 450 • Funding

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

452

- 453 • Authors' contributions

454 AC and MAW wrote the manuscript and performed data analysis. The recitation was co-

455 designed by AC, MR, and MAW. Genie was developed by BHR and RAC. AC, MR, RAC, and

456 MAW edited the manuscript.

457

- 458 • Acknowledgements

459 The authors would like to thank to the students of the BIO345 Evolution course at Arizona State

460 University for participating in this study.

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

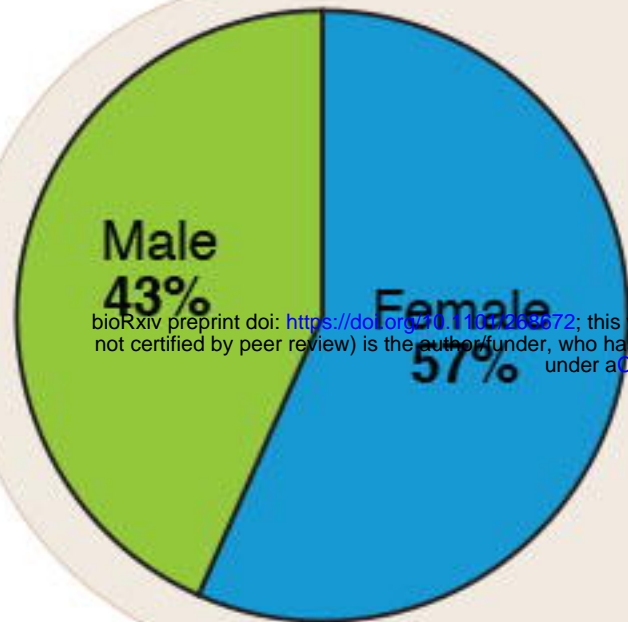
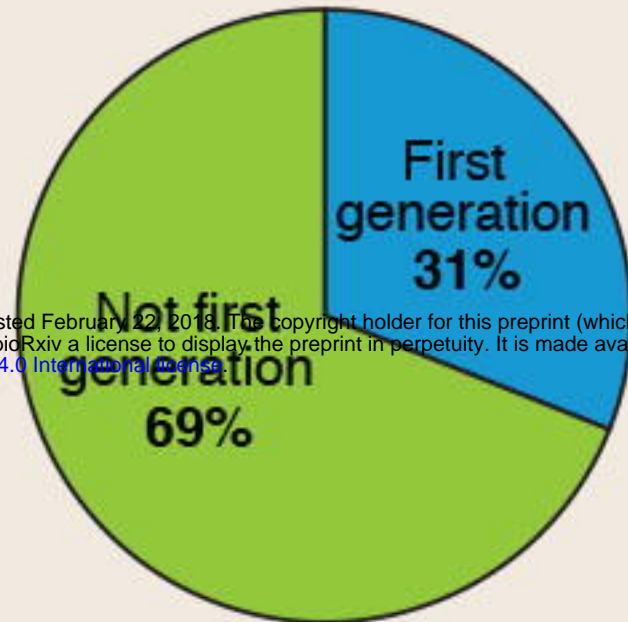
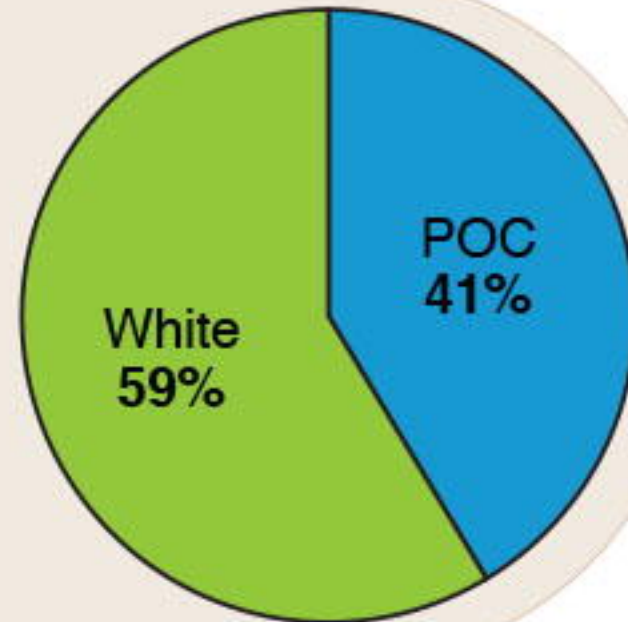
476

477

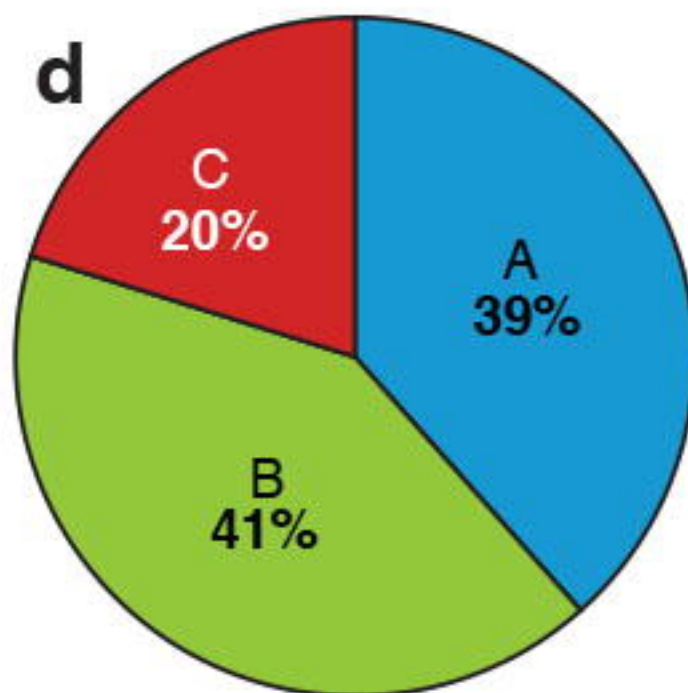
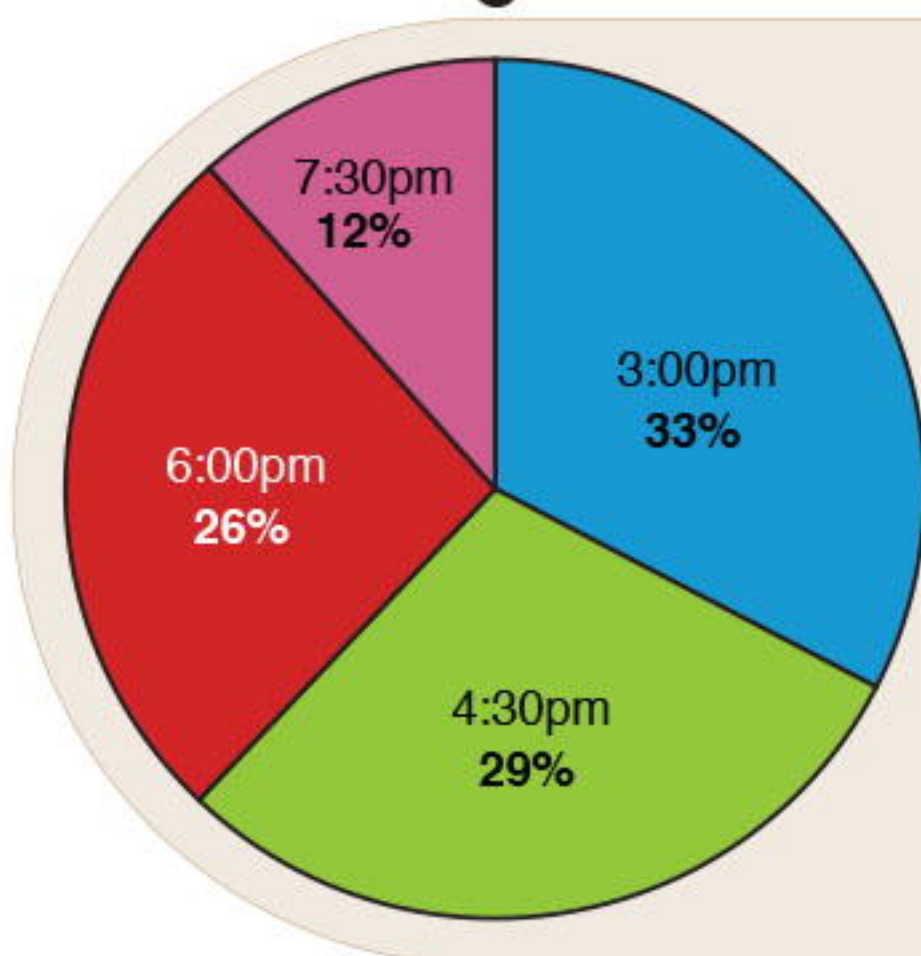
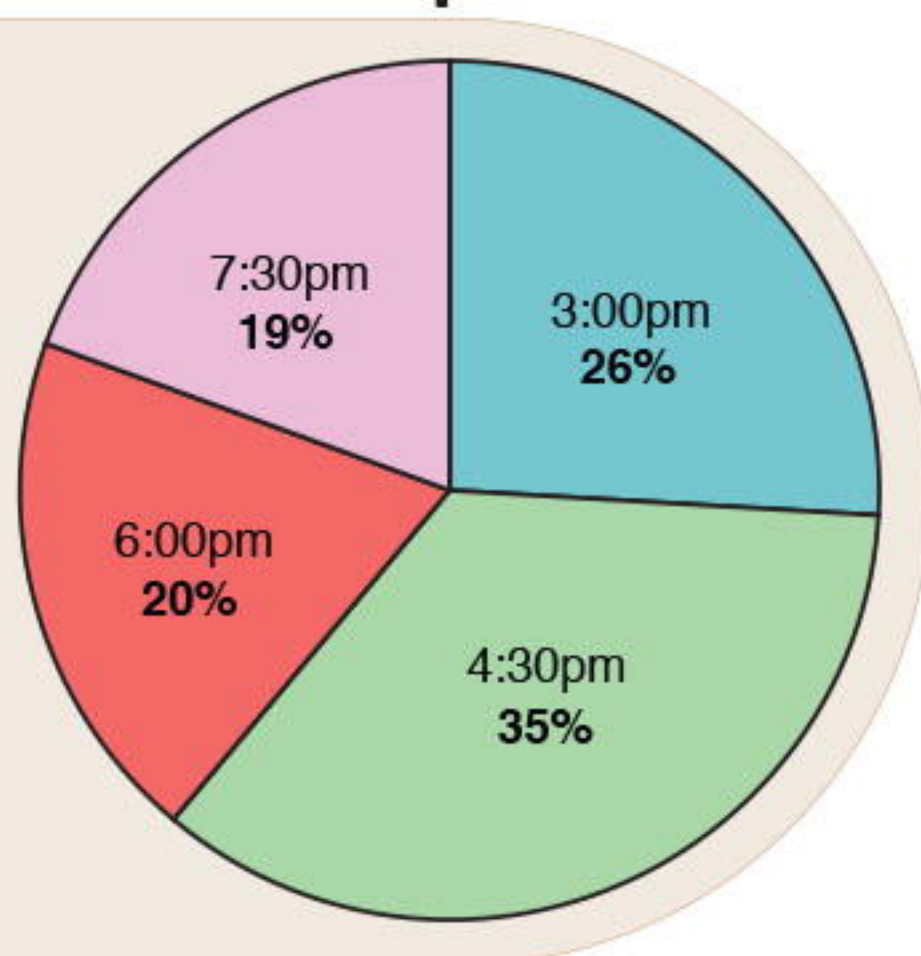
478

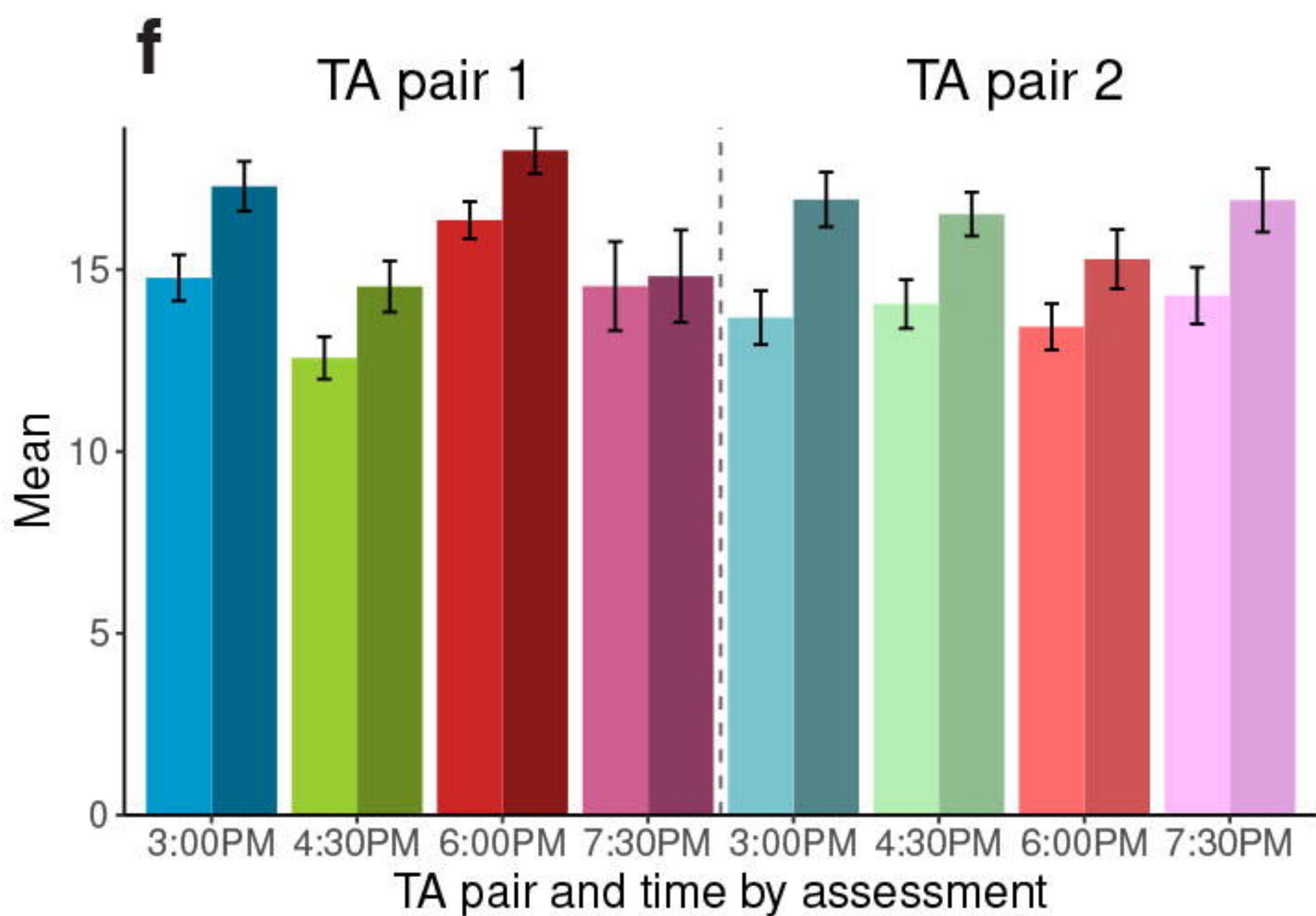
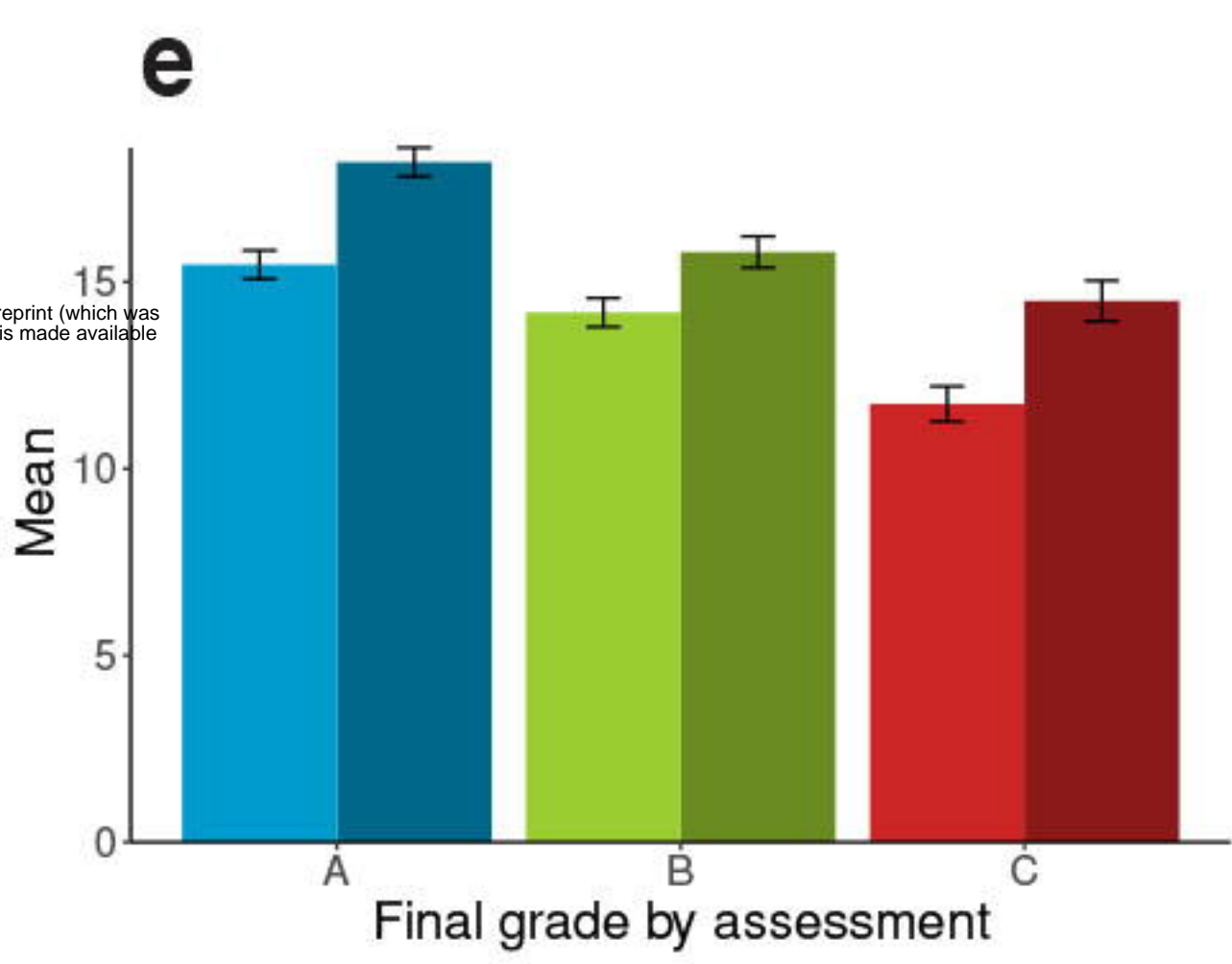
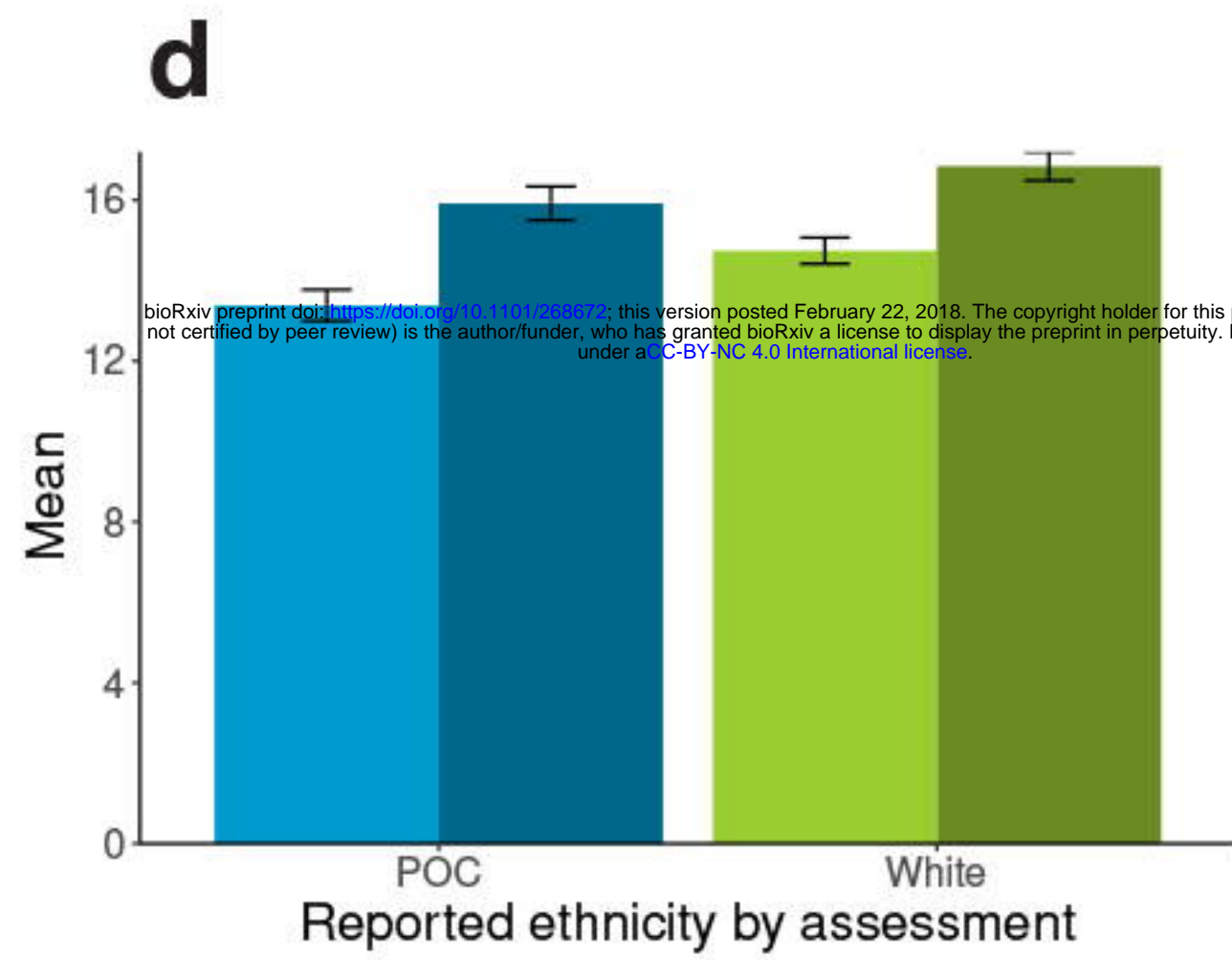
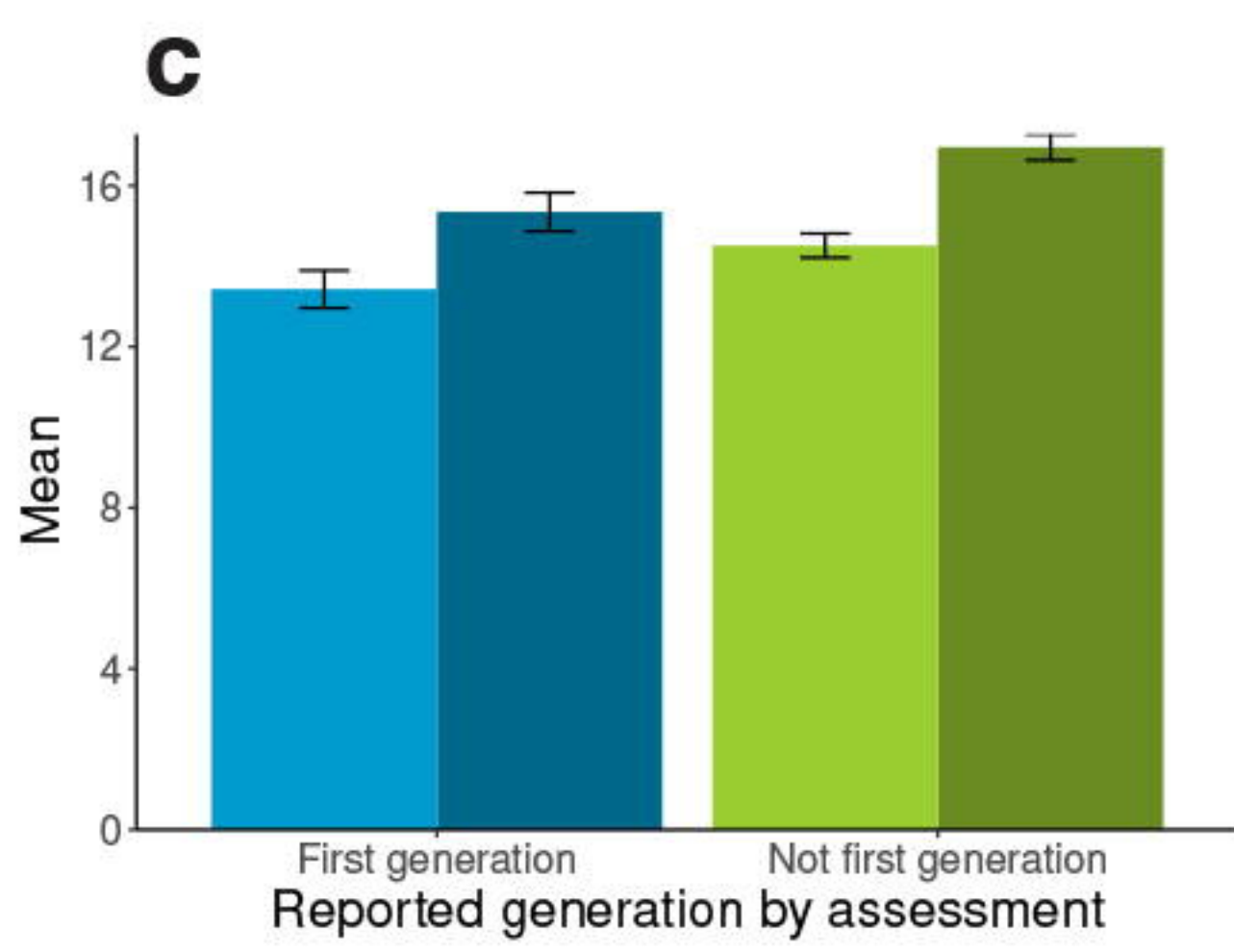
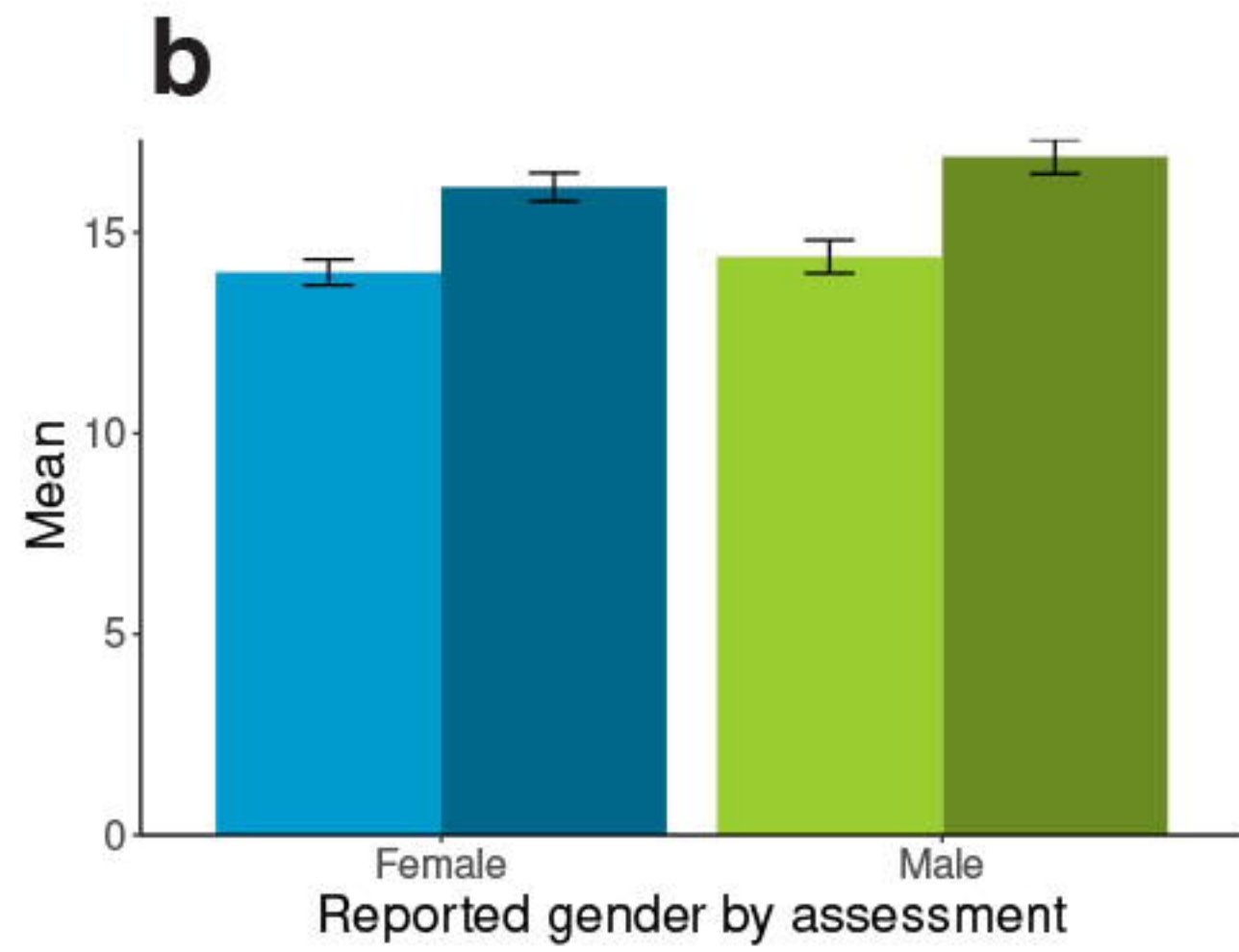
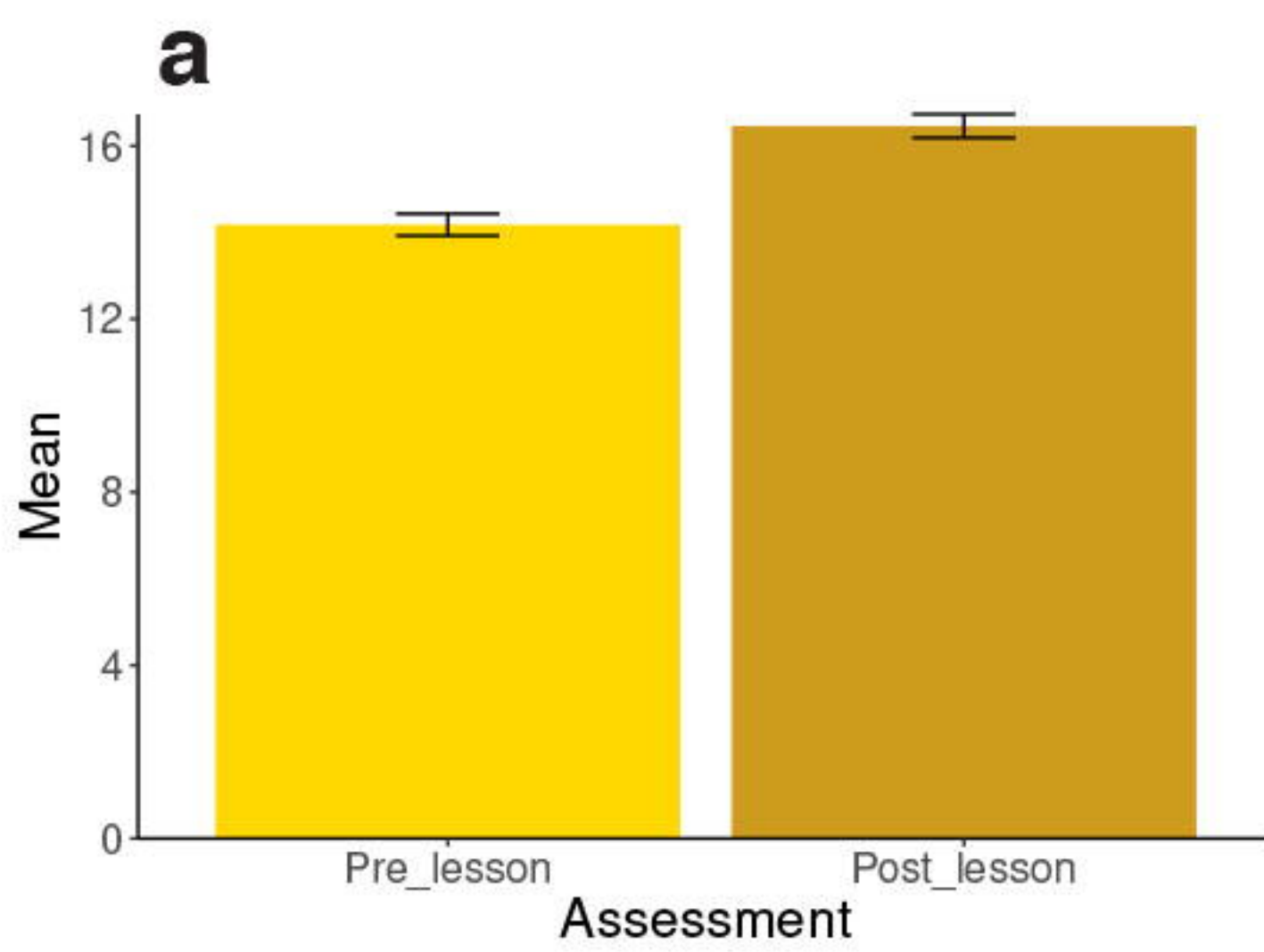
479 **References**

- 480
- 481 Alters BJ, Nelson CE. Teaching evolution in higher education. *Evolution* (N. Y).
482 2002;56(10):1891–901.
- 483 Andrews TM, Price RM, Mead LS, McElhinny TL, Thanukos A, Perez KE, et al. Biology
484 undergraduates' misconceptions about genetic drift. *CBE Life Sci. Educ.* 2012;11(3):248–59.
- 485 Arenas M, Posada D. Recodon: Coalescent simulation of coding DNA sequences with
486 recombination, migration and demography. *BMC Bioinformatics.* 2007;8(1):458.
- 487 Barton NH. Mutation and the evolution of recombination. *Philos. Trans. R. Soc. B Biol. Sci.*
488 2010;365(1544):1281–94.
- 489 Ellstrand NC, Rieseberg LH. When gene flow really matters: gene flow in applied evolutionary
490 biology. *Evol. Appl.* 2016;9(7):833–6.
- 491 Haubold B, Pfaffelhuber P, Lynch M. MIRho - A program for estimating the population
492 mutation and recombination rates from shotgunsequenced diploid genomes. *Mol. Ecol.*
493 2010;19(SUPPL. 1):277–84.
- 494 Hoban S, Bertorelle G, Gaggiotti OE. Computer simulations: tools for population and
495 evolutionary genetics. *Nat. Rev. Genet.* Nature Publishing Group; 2012;13(2):110–22.
- 496 Morjan CL, Rieseberg LH. How species evolve collectively: Implications of gene flow and
497 selection for the spread of advantageous alleles. *Mol. Ecol.* 2004;13(6):1341–56.
- 498 Nei M, Chakraborty R, Fuerst PA. Infinite allele model with varying mutation rate. *Genetics.*
499 1976;73(11):4164–8.
- 500 Perez KE, Hiatt A, Davis GK, Trujillo C, French DP, Terry M, et al. The evodevoci: A concept
501 inventory for gauging students' understanding of evolutionary developmental biology. *CBE Life*
502 *Sci. Educ.* 2013;12(4):665–75.
- 503 Price RM, Andrews TC, McElhinny TL, Mead LS, Abraham JK, Thanukos A, et al. The genetic
504 drift inventory: A tool formeasuring what advanced undergraduates havemastered about genetic
505 drift. *CBE Life Sci. Educ.* 2014;13(1):65–75.
- 506 Slavich GM, Zimbaro PG. Transformational Teaching: Theoretical Underpinnings, Basic
507 Principles, and Core Methods. *Educ. Psychol. Rev.* 2012;24(4):569–608.
- 508 Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. The evolution of mutation rates: Separating
509 causes from consequences. *BioEssays.* 2000;22(12):1057–66.

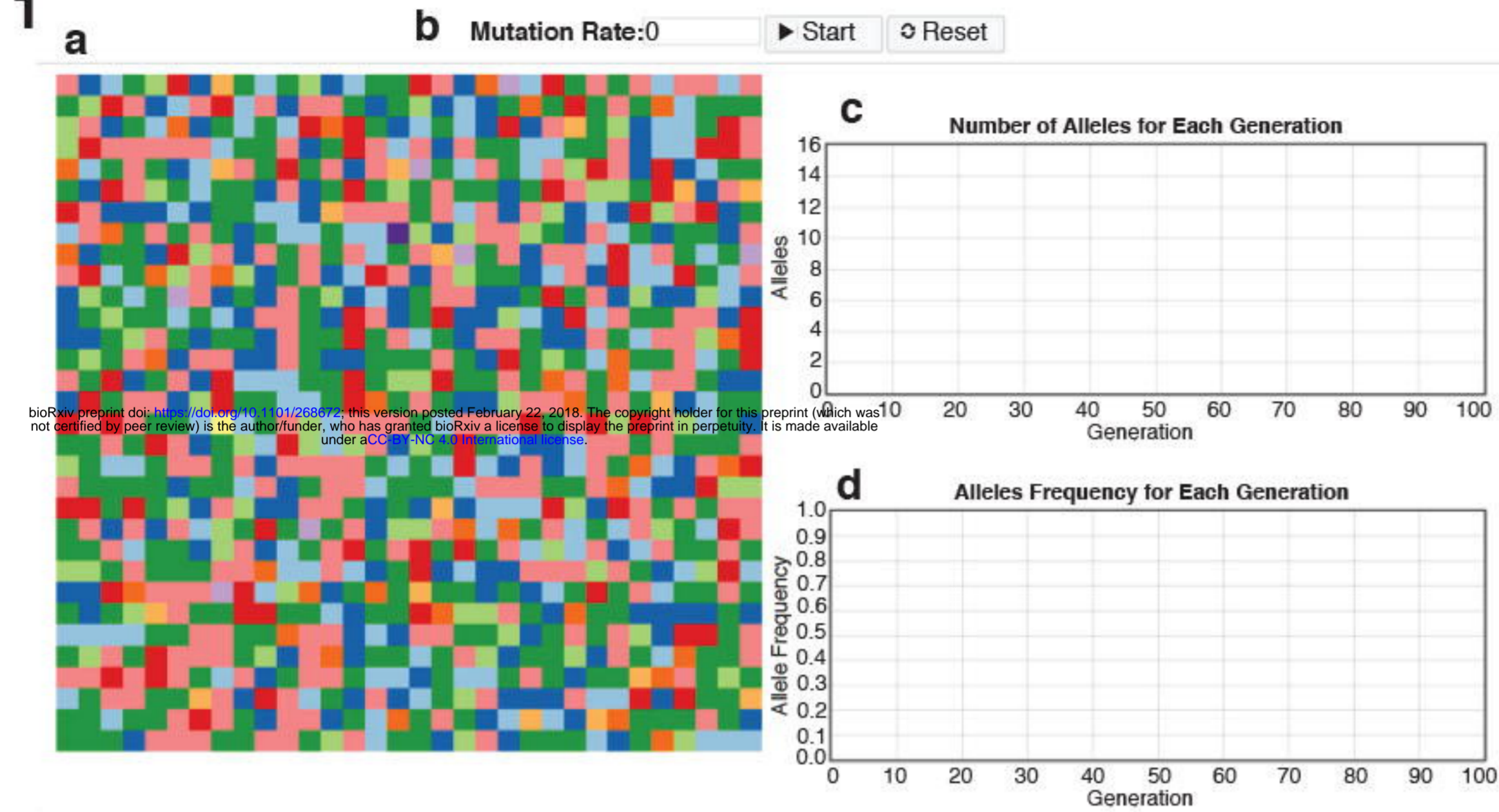
a**b****c**

bioRxiv preprint doi: <https://doi.org/10.1101/200872>; this version posted February 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

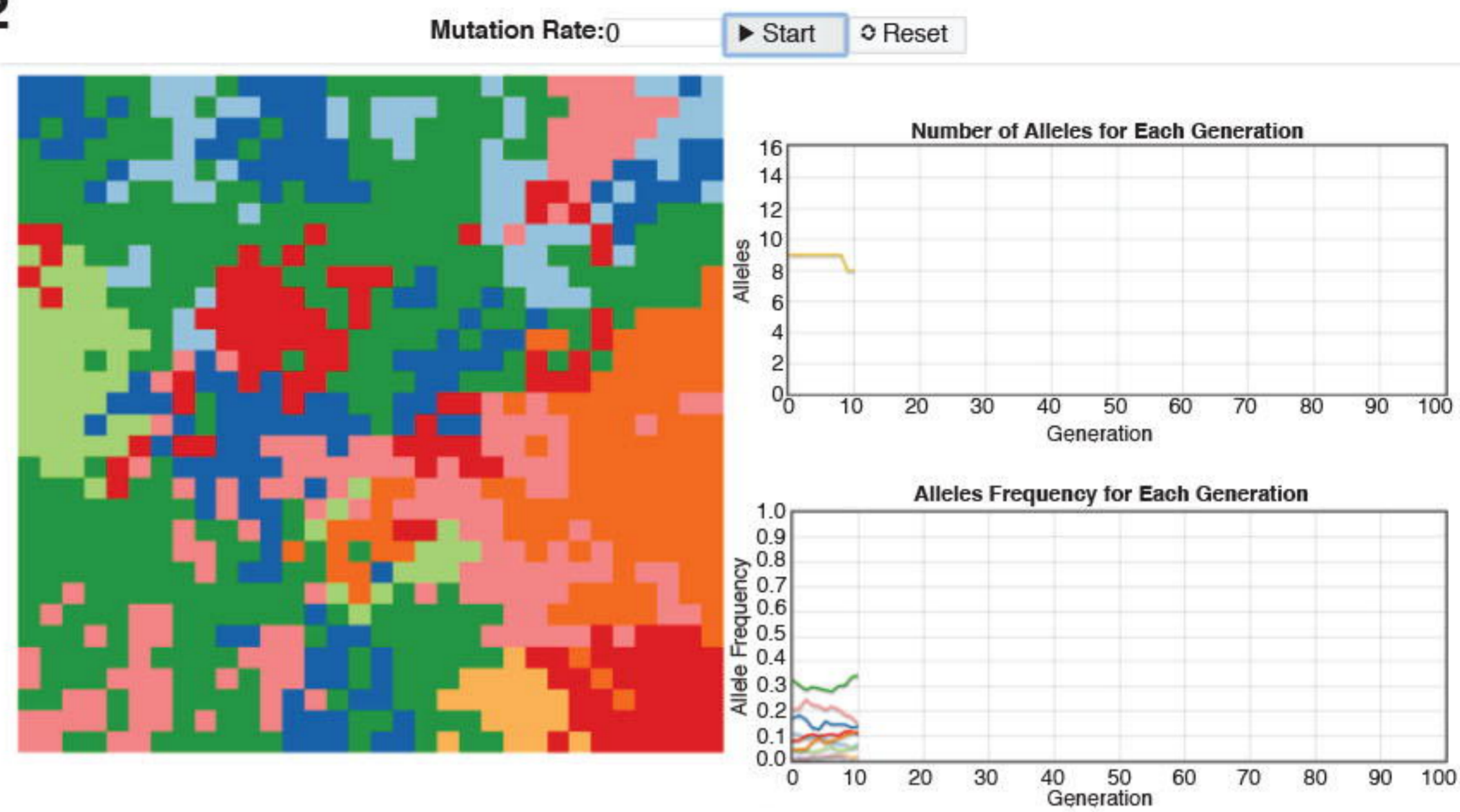
d**e****f**



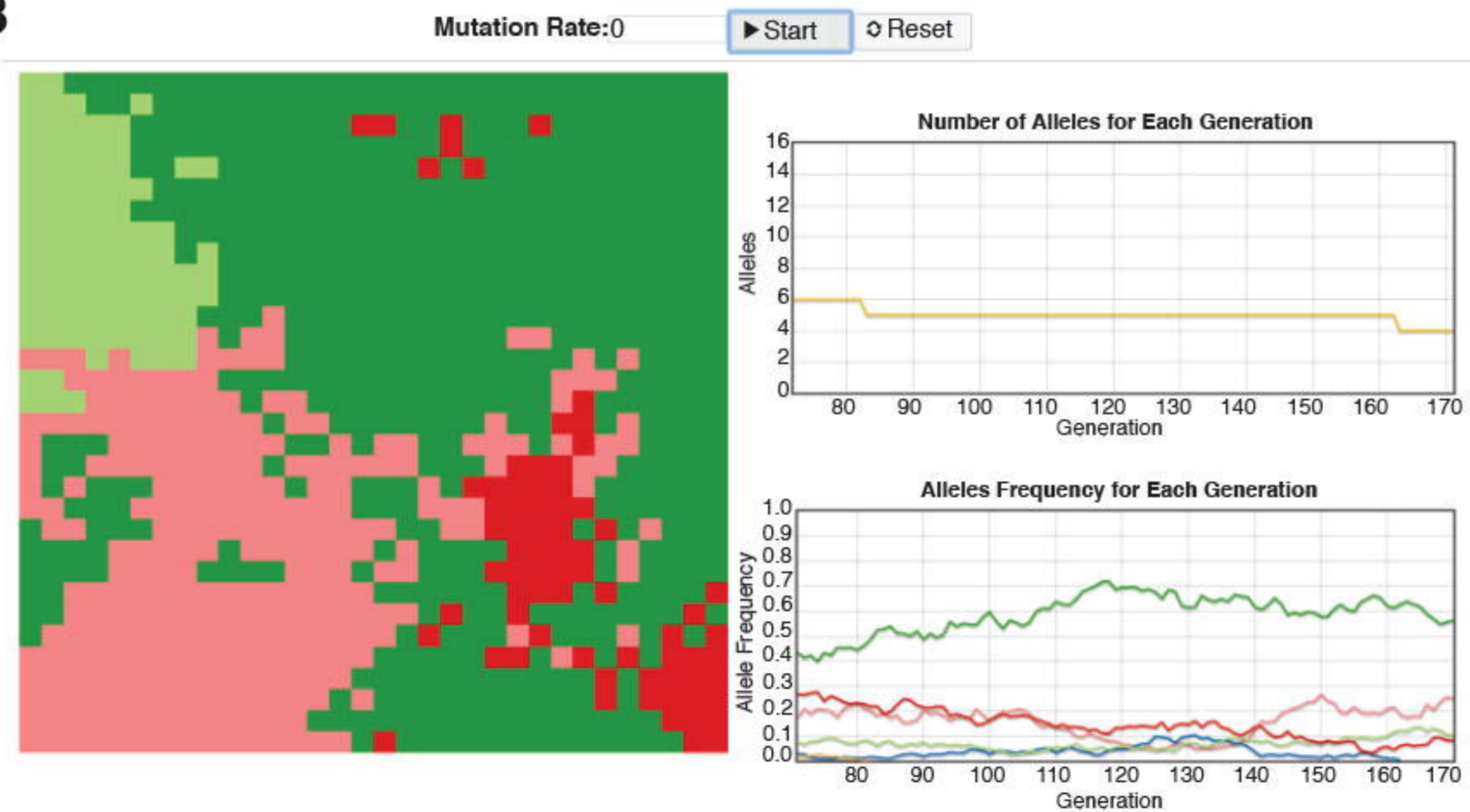
1



2



3



4

