The Power of Novelty: How Mitochondrial Diversity and Video Games Can Shape the Way We Think About Evolution

By

Jacob David Botello

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy in Biological Sciences

Auburn, Alabama
August 9, 2025

Keywords: Evolution, Video Games, Quantitative Genetics, Mitonuclear, Biology Education, Sexual Selection

Approved by

Matthew Wolak, Chair, Associate Professor of Biology
Cissy Ballen, Associate Professor of Biology
Wendy Hood, Professor of Biology and Curator of Mammals
Geoffrey Hill, William P. Molette Professor of Biology and Curator of Birds

Abstract

In this dissertation, I explore questions related to mitonuclear ecology, sexual selection, the maintenance of genetic variation, and biology education. In my first two chapters, I use cybrid laboratory populations of *Callosobruchus maculatus*, a seed beetle (introgressed genomes originating from disparate ancestral sources in southern India and Burkina Faso) to investigate questions related to the interplay between the mitochondrial and nuclear genomes. In chapter 1, I test whether the genomic interplay between the mitochondrial and nuclear genomes facilitates the maintenance of genetic variation. I find that when two mitochondrial haplotypes are present in a population, that population maintains greater stability in additive genetic variance than populations with only a single mitochondrial haplotype present. In my second chapter, I investigate the effects of interactions between the mitochondrial genome and the nuclear genome of eukaryotic organisms and how the interplay between the two genomes influences mate choice preferences in populations of these beetles. I find that *C. maculatus* demonstrate a preference for individuals whose mitochondrial haplotype originates from the same population as their own. In my third and final chapter, I investigate the effectiveness of video games as a teaching tool for use in undergraduate courses aiming to teach evolution, and to this end, I develop a video game designed to accurately represent evolution by natural selection. I test this tool's efficacy by evaluating more than 900 Auburn University undergraduate students' knowledge of evolution, as measured through a validated concept inventory, before and after thirty minutes of gameplay. I find that students do significantly

improve their understanding of evolution on average after playing as compared to students who played either a simple (evolution-free) version of the same game or no game at all over those same thirty minutes. In sum, I ask novel questions about how mitochondrial diversity influences the evolutionary trajectories of populations, and I test new approaches to teaching the all-important subject of evolution.

Artificial Intelligence (AI) Use Disclosure Statement

I utilized AI in the following ways that relate to the research conducted in this dissertation: (1) to troubleshoot coding issues, especially those related to syntax or specific argument usages within certain packages, functions, or methods necessary for analysis and figure production in R and C#, and (2) as a search engine for the titles, authors, and years of some studies which I then investigated and cited independently. The language, analyses, and research present in this dissertation are my own original work (except where explicitly attributed to others).

Acknowledgements

I want to thank each member of my committee individually. Without their tireless efforts, I could not have completed three projects far too colossal to belong entirely to me. I want to thank Wendy Hood for accompanying (and indulging) me on a journey of many wild ideas and for somehow perfectly balancing herself upon razor's edge between challenging and supporting me in just the ways I needed. I cannot imagine a living human being more capable of infusing everyone around her with *joie de vivre* than Cissy Ballen, even in the bleakest and most challenging times, and without her ceaseless enthusiasm for my education-driven work, I am certain I would not have had the confidence to keep making the video games that are now at the center of my upcoming opportunities. I want to thank Geoffrey Hill for his advice, endless excitement for the mitonuclear work, and for our unplanned office chats—they were one of the great highlights of my academic career and genuinely made me believe that I belong in science. Most of all, I want to sincerely acknowledge the contributions of my advisor, Matthew Wolak. He possesses the rare, perhaps singular, expertise necessary to have served as a mentor to a frazzled graduate student wading into the terrifyingly complex world of quantitative genetics. He did what many advisors would have refused and allowed me to stretch out of the familiar and into the research that fits me best. Most critically, he sustained the (near infinite) patience required to guide me to this destination. I want to thank my parents for their unwavering belief in my ability to accomplish the work below, and I want to thank the innumerable friends and family who have cheered me along over these many years. Without the

immeasurable artistic talents of my sister, Leah Tennery, the warlak would never have come to be, and so we (and the many populations of warlaks that have come and gone) all owe her a great debt of gratitude. I want especially to single out Annie Culver for her empathy and support during the most difficult stretches of this process and Joe Mikolaj for encouraging me from my first to my very last day as a PhD student. I'm grateful to Josh Hall for becoming a lifelong friend while showing me the best herping spots in Alabama, and to Abby Beatty for putting her entire world aside time and again to help with anything I might possibly need. Special thanks to Tonia Schwartz for her truly invaluable instruction and assistance in all the molecular methods so arcane to me and to Randy Klabacka for his critical insights into the mitonuclear world. More special thanks still to Stephen "Duncy" Duncan, Alex Decaulp, Eric Botello, Betsy Corona, Chris Norris, Sydney Wayne, Kaylene Yamada, Natalie Harris, Natasha Grabau, and Madeline Corbeau. I'd like to thank Murry Gans for giving me my first research opportunities, without which many doors may never have opened for me. Finally, I would be remiss (and would deserve a yowling) if I neglected to mention and thank Ma'am, the fluffiest and most affectionate cat that a person could hope to call a companion on this long and winding journey.

I could not—would not—have done this without each of you.

Table of Contents

List of Tables

List of Figures

Introduction

An Introduction to Evolution

The diversification of all life as we know it proceeds through the process of evolution in which traits and qualities are inherited across generations of organisms through expression via deoxyribonucleic acid, or DNA. The central dogma of molecular biology states that DNA gives rise to RNA, RNA produces proteins, and those proteins are folded and arranged into all the components that make up a given organism (Crick 1970). Transcription and translation produce all life as we know it. The observable qualities of an organism constitute its phenotype, and these phenotypes can vary wildly across individuals, but the heritable portion of this variation is ultimately derived from changes to DNA that occur through errors in replication which lead to nucleotide substitutions we name mutations (Dobzhansky 1937). Though these mutations are the primary source of all the observable genetic diversity that exists today, many mechanisms (explored more in depth in chapter 1) exist which preserve the standing variation produced by mutation. However, what we observe in the present is only a small subset of the genetic variation possible because each generation these diverse individuals, in competition for resources of many kinds, may be more or less likely to survive and reproduce (the measure of success in these areas we often refer to as 'fitness') than others as a result of their differences—a process we call selection (Fisher 1930; Dobzhansky 1937). Thus, evolution by natural selection occurs when differences among individuals leads to differential fitness that changes the frequencies of alleles between one generation and the next.

For a population to adaptively evolve in response to natural selection, though, there must be genetic variation present in the population upon which selection can act, and this presents a long-standing problem in evolutionary biology. Theory predicts natural selection should rapidly deplete standing genetic variation (Fisher 1930; Barton and Turelli 1989; Barton and Keightley 2002), especially for fitness and its components. Empirical tests, however, often find substantial genetic variation for fitness (often measured as additive genetic variance, or VA, Hendry et al. 2018) in wild populations. Biologists have proposed many mechanisms to explain this persistence in heritable variation, such as antagonistic pleiotropy, through which an allele that causes deleterious effects in one trait might produce a net benefit to an organism by advantaging that organism in a second trait, effectively producing balancing selection (Hughes and Burleson 2000), or intralocus sexual conflict, a circumstance in which optimal trait values differ between sexes, and a given allele's effects on fitness might differ in sign and magnitude based on the sex in which it is expressed (Arnqvist and Rowe 2005). None of these explanations, though, fully explain the amount of standing genetic variation measured.

Chapter 1: What Maintains Genetic Variation?

In chapter 1, I propose a novel solution to this question by considering a critical component of every eukaryotic cell, the mitochondria, not only in terms of its well-celebrated energy production but also in terms of its capacity for cooperation—and dysfunction—when it comes to its interactions with the nuclear genome. Although biologists have studied mitochondria and their maternally-inherited, non-recombining genomes for more than a century since their naming by Altmann (1894), only since the turn

of the millennium has the mitochondrial interplay with the nuclear genome been brought into major focus (Rand et al. 2004; Dowling et al. 2007, but see Burton 1992a). Across the inner mitochondrial membrane, the process of oxidative phosphorylation (OXPHOS) progresses as electrons are guided through five complexes that make up the electron transport system, producing the majority of the ATP that powers most energy-requiring operations performed by eukaryotes (Barrientos et al. 2009a). Although this process takes place entirely within a mitochondrion, the subunits that compose these five complexes are co-encoded by both the mitochondrial and nuclear genomes. Cooperation between the two genomes entails maximizing the production of ATP and minimizing the production of any harmful byproducts, and this cooperation can occur between proteins encoded by different genomes, between proteins and DNA, and between proteins and RNA. The results of dysfunction through the production of incompatible electron transport system subunits by the two encoding genomes can cause the production of harmful free radicals, a reduction in ATP output, and in more extreme cases of incompatibility between the two genomes, complete organismal system failures or cell death (Rand et al. 2004; Dowling et al. 2007; Lane 2011). The evolutionary stakes of mitonuclear matching, then, must be very high, as an organism's fitness—its contribution to the next generation—is heavily dependent upon its ability to compete across numerous factors where energetic demands are functionally impossible to meet without mitochondrial ATP production (Lane 2011). These mutation-driven incompatibilities might be inescapable in asexual species whose nuclear and mitochondrial genes are co-transmitted to offspring. In sexually reproducing species, however, although the mitochondria's vertical transmission renders it incapable of

changing genetically except through mutation, the nuclear genomes of most eukaryotes are capable of recombination and produce novel allelic combinations through mating. Therefore, populations can adaptively respond over multiple generations to compensate for such dysfunction, potentially pairing nuclear alleles of superior compatibility to a given mitochondrial genotype (Hill 2015).

I propose that this critical interplay between the two genomes presents a potential explanation for the maintenance of genetic variation in nature. When only a single mitochondrial haplotype is present in a population, selection should eliminate any deleterious alleles that do not exhibit antagonistic pleiotropy and that contribute to mitonuclear dysfunction over time. When multiple mitochondrial types exist, however, nuclear alleles might be associated with higher or lower fitness depending on the mitochondrial backdrop against which they are expressed. An allele's contribution to fitness could change sign or magnitude based on its context, preserving it in the population.

Chapter 2: What Drives Mate Choice?

In my second chapter, I investigate whether the long-observed choosiness of individuals selecting mates is driven by an organism's interest in pairing with a mate whose mitonuclear complement is compatible with the chooser. Mate choice represents one individual's preference for certain individuals over others as potential mates (Darwin 1871). Since its proposal by Darwin in the late 1800s, many biologists have suggested that signal and preference are evolutionarily entangled and have hypothesized endlessly about the

nature of what these signals may indicate. These hypotheses explore a range of mechanisms—for instance, that preference for traits represents a runaway process through which (typically female) preferences for a trait result in a feedback-loop that produces initially arbitrary continuous phenotypic exaggeration (Fisher 1930). Another hypothesis underscores an argument that extreme expressions of traits, and especially ornaments, serve as indications of quality because they produce a handicap too extreme for a lower-quality individual to bear (Zahavi 1975). Others argue that these preferences indicate signals for 'good genes' likely to confer some benefit to the organisms carrying them, such as resistance to ecologically relevant pathogens likely to endanger offspring (Hamilton and Zuk 1982). The lattermost of these, the 'good genes' hypothesis, has been well explored theoretically but has yielded little empirical evidence (Achorn and Rosenthal 2020) despite its popularity. In 2013, Hill and Johnson proposed the mitonuclear compatibility hypothesis of sexual selection, arguing that there are few features more common and essential to all eukaryotes than mitochondria, and thus the costs of possessing low-quality mitochondria are extreme, and the benefits of efficient mitochondria are great. Because the mitochondria are co-encoded by the mitochondria genome and the nuclear genome, the stakes of these two genomes producing a functioning product are great, and similarly, as mates each donate half of their mitochondria-encoding nuclear (N-mt) alleles, the stakes of finding a compatible mate are high indeed. Therefore, mates must assort based on mitonuclear compatibility to maximize mitochondrial function (or at least avoid mitochondrial dysfunction) in their offspring. Whether individuals should prefer a nuclear match, a mitochondrial match, or a mitonuclear match, though, remains

unknown. In my second chapter, I test the mitonuclear compatibility hypothesis, further proposing that the compatibility of OXPHOS genes drives not only the evolution of ornaments but also the evolution of preference for compatible mates.

Chapter 3: If Evolution is so Critical a Subject, How Can We Most Effectively Teach It?

Portraying evolution by natural selection requires consideration of factors ranging from the instantaneous reactions occurring at the molecular and cellular levels to the phenotypic composition of a whole organism manifested at the intersection of its genetics and environment. The scope of evolution by natural selection stretches from the interactions that occur between the endlessly complex individuals that constitute populations to processes that take place across an entire planet (at least one that we know of) over a geological timescale, with species colliding and intersecting to further diversify and shape themselves and one another. Yet, as Theodosius Dobzhansky famously contended, without evolution, nothing in biology makes any sense (1973). Fully grasping the mechanics of this diversifying force and the possibilities it conceals has major implications for the eradication of dangerous pathogens and evading the accidental production of new ones. Comprehension of evolution can similarly affect the way we live and reproduce, our stewardship of the planet, the length and quality of our lives, and even the cures of our ailments—but grasping its full scope is a lot to ask of a student (or an educator). It is no surprise, then, that students routinely struggle with the subject, and such gains as are made are often later lost to either misconceptions or confusion (Anderson et al. 2002; Gregory 2009; Abenes and Caballes 2020; Archila et al. 2024). Moreover, in some environments, evolution meets strong opposition by those who argue that it undermines

their religious beliefs or that its adoption represents a political position or even causes them to contemplate their own existence or mortality in ways that can be frightening (Yasri and Mancy 2014; Barnes et al. 2021; Newall and Reiss 2023). How, then, can these complex but essential ideas be communicated to each new generation of learners? Although other effective approaches, such as developing cultural competencies, have been well-studied (see work by Elizabeth Barnes, for example, Barnes and Brownell 2017; Barnes et al. 2017, 2020b), I make the case for a new but familiar approach to tackling this issue. I argue that video games offer a medium through which students can experience (or simulate) the full scope of evolution, observe its processes, and, as an added benefit, reduce stereotype threat, the gender gap, and social anxiety in the process (Annetta 2008; Annetta et al. 2009; Franceschini et al. 2022; Checa-Romero and Gimenez-Lozano 2025; Lavalle et al. 2025). In my third and final chapter, I plan, develop, and test at a large state university a video game designed for the expressed purpose of teaching evolution by natural selection simply by representing it accurately in a familiar, 2D platformer format.

Study System

Callosobruchus maculatus, a seed beetle, is a widespread equatorial pest known to parasitize legumes, preferentially those in the *Vigna* genus in the wild (Fox et al. 2004; Craig Stillwell and Fox 2009). Seed beetles have established footholds in North and South America, Asia, and Africa, though Africa hosts the largest number of populations (Kebe et al. 2017). *C. maculatus* has been shown to display considerable mitochondrial diversity, boasting a significant number of unique mitochondrial haplotypes (Kebe et al. 2017) both within and across populations.

Figure 1: *Callosobruchus maculatus* (male).

Seed beetles are semelparous with females laying between approximately 30 and 120 eggs in a lifetime (which typically lasts just over a month, Fox 1993a). Females adapted to mung beans (*Vigna radiata*) will generally lay a single egg per host (bean) until no fresh beans are available, at which point they will lay on already-parasitized beans (Messina 1990). Both males and females nearly always die long before their offspring emerge as adults, producing non-overlapping generations within populations (Fox 1993b). Eggs hatch within one week and larvae begin consuming the host bean, burrowing into the interior as they develop. Offspring undergo metamorphosis within 3-4 weeks of hatching, emerging from beans as sexually mature adults. Male larvae tend to metamorphose earlier than females, emerging at a lower mass on average. Variation in development time has also been linked to variation in size of the host bean, average temperature, and day/night cycle.

In instances of multiple eggs sharing a single bean, host size may influence larval strategies (Messina 1990; Fox 1993b,a). Though larval strategies are the result of a confluence of other factors, larvae laid on smaller host species (e.g. mung beans) tend to develop into contest competitors, killing other larvae also developing within the same bean, whereas eggs laid on larger species of beans (e.g. cowpeas, *Vigna unguiculata*) tend to produce scramble competitors, associated with accelerated development time and reduced adult mass at emergence (Messina 1990). Adult seed beetles are short-lived, typically surviving only 1-3 weeks. Adults are facultatively aphagous, living only on the resources they accumulate as larvae, though adults have been shown to experience greater longevities when fed sugar water when offered in experimental settings (Fox 1993b,a).

The Wolak lab currently houses two distinct populations of *C. maculatus* derived from the stock culture populations of the Fox lab at the University of Kentucky (Kebe et al. 2017), one of which ancestrally originated in Burkina Faso (BF) and is adapted to using cowpeas as a host, and another which ancestrally originated in southern India (SI) and is adapted to using mung beans.

Stock Rearing Protocols

We house stock culture seed beetles in large, plastic boxes (approximately 30cm x 20cm x 6cm) which we seal with a thin, breathable cloth to secure beetles inside while allowing air circulation. Nested within each large box is a smaller, interior plastic box (approximately 15cm x 5cm x 4cm) filled with 750mL of either unparasitized (hereafter 'fresh') mung beans

(SI boxes) or fresh cowpeas (BF boxes)—all beans used are organic and frozen below 0° C for at least forty-eight hours (and then thawed) before use. Adults emerge from parasitized beans, mate freely, and lay eggs on fresh beans within the small boxes. Once either the vast majority of the adults have died or approximately ten days have passed since the majority of the adults have emerged, we discard all adults (dead and the remaining living individuals) and parasitized beans outside the interior box. We then sample 225mL of the 750mL newly parasitized beans from the interior dish and place them into a clean, large box, nesting within it 750mL of fresh seeds in a clean, small box.

We repeat this procedure once every generation (approximately 23-26 days) for both BF and SI stock populations. Rearing protocols for these two populations differ only in that SI boxes are stocked with mung beans and BF boxes are stocked with cowpeas. Our stock populations of BF and SI beetles are repeated across three such stock boxes (i.e. three for SI and three for BF) and all kept on a fourteen-hour day, ten-hour night cycle at 27° C. These populations house between approximately 5,000 and 10,000 individuals per generation (with more individuals appearing among populations reared on cowpeas rather than mung beans).

Every six months (approximately 7-8 generations), we combine all parasitized beans from all stock boxes within a population (e.g. BF) to mix and redistribute them among the three new boxes to maximize genetic diversity among stock culture boxes.

When managing or cultivating smaller populations, instead of using the large plastic bins described above as the exterior bin, we use the smaller plastic bins mentioned earlier

(approximately 15cm x 5cm x 4cm), providing a large, 60mm petri dish (which holds approximately 60g of beans) in the center filled with the appropriate beans for a given population. As with the larger boxes, we outfit these smaller population containers with a lid and breathable cloth to allow air circulation. Unless otherwise specified, we sampled 100% of the beans from the petri dish to serve as the new generation of adults (rather than the 1/3rd typically sampled in large populations).

Cybrid Line Development

We developed two types of cybrid lines (i.e. which possess a mitochondrial genotype and an introgressed nuclear genotype): those with African nuclear alleles and southern Indian mitochondria, or the "spice" type cybrid lines, and those with southern Indian nuclear alleles and African mitochondrial alleles, or the "candy" type cybrid lines. To form the spice lines, we initially haphazardly sampled >100 mung beans with eggs laid on them from a southern Indian (SI) stock culture of *C. maculatus* seed beetles in May 2020. We housed each of these beans singly in a 35mm petri dish and, once the adults had emerged from beans, we randomly selected forty female individuals to serve as mitochondrial line founders. We monogamously mated each of these females to a haphazardly sampled male from a stock culture of *C. maculatus* from Burkina Faso (BF) over at least thirty minutes in a 60mm petri dish and allowed each female to lay eggs by providing her with 2.4g of mung beans within a unique 60mm petri dish. We expect that this mating event reduced the average number of southern Indian nuclear alleles present in parents to those in offspring by 50%. In the following (and each subsequent) generation, we collected 3-5 virgin, adult females from each line's pool of eggs on beans and group-mated them to 3-5

haphazardly sampled adult males from the BF stock culture, preserving mitochondrial haplotypes of offspring but reducing, on average, SI nuclear alleles among offspring by 50%. We repeated this process for more than 30 generations, reducing the average expected abundance of SI nuclear alleles in the descendent population by 99.99%. We made a second set of cybrid lines, the "candy" lines, by following the same procedure exactly but by pairing BF females with SI males and ultimately producing a cybrid line composed of SI nuclear alleles but BF mitochondrial haplotypes.

Chapter 1: Diversity Keeps Us Together: How Mitochondrial Diversity Stabilizes Additive Genetic Variance in Populations of Seed Beetles

Introduction

Evolutionary theory predicts that both stabilizing and directional selection will deplete adaptive genetic variation in a given trait within a population (Barton and Turelli 1989; Barton and Keightley 2002). An individual's fitness, which can be broadly defined as the reproductive contribution an individual makes over its lifetime to its population (Stearns 1989; Morrissey et al. 2012; Shaw and Shaw 2013; Queller 2017; Grafen 2018; Hendry et al. 2018; Shaw 2018) determines its representation in the next generation, and selection is expected to favor alleles associated with high-fitness phenotypes. These alleles are expected to increase in frequency until they reach fixation, eliminating allelic variation at a given locus (Lewontin 1974). Conversely, alleles consistently associated with low-fitness phenotypes are likely to be selected against and, over time and under persistent selection, these alleles are predicted to eventually disappear from the population.

As more and more alleles are purged from the population through selection and other alleles move to fixation, theory predicts that the population-level additive genetic variance—a measure of genetic variation which represents the proportion of phenotypic variance that can be attributed to differences in average allelic effects—will decline (Barton et al. 2016). Studies, however, consistently find that substantial population-level additive genetic variance for traits—including key life history traits and even fitness itself—

perplexingly remains in wild populations even after many generations of selection (Hendry et al. 2018).

A number of mechanisms have been proposed to explain the preservation of such variation, including migration, variation in the strength of stabilizing selection and/or balancing selection, sexual selection, sexual conflict, frequency dependence, dominance reversals, genotype-by-environment interactions, and high mutation rates (Barton and Keightley 2002; Radwan 2008; Kazancioğlu and Arnqvist 2014; Brisson 2018; Grieshop and Arnqvist 2018; Connallon and Chenoweth 2019); however, the substantial balance of additive genetic variance in many populations is incompletely explained even by these mechanisms (Turelli and Barton 2004; Bürger 2005). What sustains the abundance of additive genetic variance observed in natural populations remains a long-standing, unanswered question in evolutionary biology (Fisher 1930; Connallon and Chenoweth 2019).

A novel potential solution to this question involves the epistatic interplay between the mitochondrial genome and the nuclear genome. Selection acts directly on phenotypes rather than alleles themselves, and so the same allele may potentially be associated with differing levels of fitness depending on the backdrop in which the allele is expressed (Arnqvist et al. 2014). The mitochondrial background is one such context that may dramatically influence the level of fitness associated with a given organism and all the nuclear alleles it carries (Hill et al. 2019). Although mitochondria possess unique genomes vertically transmitted from females to offspring, the subunits of the five mitochondrial

complexes which power the electron transport system (ETS) are encoded by both the mitochondrial genome and the nuclear genome (Anderson et al. 1981; Rand et al. 2004). These nuclear genes and others that encode mitochondrial components, so-called N-mt genes, can directly influence the performance of the mitochondria, and by extension, the entire organism which the mitochondria powers. For many years, mitochondria were assumed to evolve only under a neutral model (Ballard and Kreitman 1995), but studies since the turn of the millennium (Arnqvist et al. 2010; Kazancioğlu and Arnqvist 2014; Pennisi 2016; Wolff et al. 2016; Hill 2017, 2020; Havird et al. 2019; Hill et al. 2019; Keaney et al. 2020; Dowling and Wolff 2023) have highlighted the impact of not only non-neutral mitochondrial effects on traits but the effects of the interactions between the mitochondrial and nuclear genome, or mitonuclear interactions (sometimes referred to as cytonuclear when mitochondrial effects are difficult to disentangle from other factors within the cell).

A substantial body of work in recent years has focused on the effects of interactions between the nuclear and mitochondrial genomes on individual level and population level performance. The interplay between the mitochondrial genome and the nuclear genes that co-encode mitochondria—particularly those involved in assembling the subunits of the electron transport system—is integrally related to mitochondrial ATP production, and thus an organism's mitonuclear compatibility affects its potential to make energy. This serves as the primary energy source for most physiological processes in eukaryotes (Van Der Giezen 2011; Lane and Martin 2015; Barreto et al. 2018), and is consequently linked to differences among individuals in longevity, offspring production, size, mating success, and

competitive success against conspecifics (Burton 1990b; Gomulkiewicz and Holt 1995; Lynch 1996; Edmands and Burton 1999a; Lane and Martin 2015; Blier et al. 2019; Hill et al. 2019; Hill 2020; Koch et al. 2020; Rank et al. 2020; Weaver et al. 2022). Different mitonuclear combinations have been shown to meaningfully influence the evolutionary trajectories of populations of yeast (e.g. *Saccharomyces cerevisiae*, Biot-Pelletier et al. 2023) under strong, experimental selection for cellular respiration (*S. cerevisiae* are facultative anaerobes and can exist without mitochondrial cellular respiration).

The consequences of dysfunctional mitochondria affect most processes involved in survival and reproduction, and so the tight-knit relationship between the mitochondria and nuclear genomes appears to promote selection on coevolutionary compatibility between the two (Weaver et al. 2022). The mitochondrial genome does not recombine and is transmitted directly from mothers to their offspring. Unlike traits chiefly encoded by the nuclear genome in which rapid evolution of quantitative traits (Gomulkiewicz and Holt 1995) can quickly occur through rapid adaptation from one generation to the next, selection is generally forced to act upon the entire mitochondrial haplotype rather than specific alleles within it because the mitochondrial genome is inherited in its entirety. Moreover, vertically transmitted genomes are subject to Mueller's ratchet (Lynch 1996) and much more likely to accumulate deleterious mutations that go to fixation, ultimately producing less functional mitochondria, the negative effects of which can be especially biased toward males due to selection operating exclusively on mitochondria in one sex (the "Mother's Curse" hypothesis (Keaney et al. 2020; Leeflang et al. 2021). Hill (2020) recently proposed the mitonuclear compensatory coevolution hypothesis, which suggests that over

time, substitutions in the nuclear genome that encode for components of mitochondrial subunits offset the dysfunction produced by the accumulation of deleterious mutations in the non-recombining mitochondrial genome, offering a novel explanation for how mitochondria remain efficient even under pressure from Mueller's ratchet. The mitonuclear compensatory coevolution hypothesis may offer insights into another major standing question in evolutionary biology, namely, what maintains the substantial genetic variation consistently measured in key fitness components (and fitness itself) in wild populations?

This interaction between the nuclear and mitochondrial genomes may affect how much nuclear genetic trait variation persists within populations because different nuclear alleles can be associated with greater or lesser individual fitness depending on the mitochondrial background in which they are expressed (Burton 1990b; Edmands and Burton 1999b; Blier et al. 2019; Hill et al. 2019). If only a single mitochondrial haplotype is present within a population, mitochondria-encoding nuclear alleles which are incompatible with this mt-haplotype will be selected against and possibly eventually purged from the population. If more mitochondrial haplotypes are present, however, more compatible mitonuclear combinations may be revealed as nuclear alleles which pair poorly with one mitochondrial haplotype may pair favorably with another. Although the interactions between the mitochondria and nuclear alleles are strictly epistatic, these interactions could serve to preserve additive genetic variance (or VA) by potentially creating associations between the nuclear alleles involved (which might otherwise be selected against) and high-fitness phenotypes. We recognize also that, although the specific

number of nuclear alleles involved in OXPHOS may vary by species, these alleles likely constitute only a very small portion of the complete nuclear genome, their effects are almost certainly pleiotropic and therefore likely to affect many traits given the suite of processes powered by ATP. The more functional consequences of mitonuclear compatibility at the level of the individual have been studied. However, what remains an open question is the extent to which these functional consequences of mitonuclear interactions may impact evolutionary processes (Dowling and Wolff 2023, but see Biot-Pelletier et al. 2023).

A diversity of genetically distinct mitochondrial haplotypes in a population opens the door to more high-fitness mito-nuclear combinations, potentially slowing the rate at which selection and drift can eliminate nuclear alleles associated with these combinations. However, whether the number of genetically distinct mitochondrial haplotypes present in a population directly influences the rate at which nuclear genetic variation changes over multiple generations has never been empirically tested.

Here, we experimentally test the role of mitochondrial diversity in preserving additive genetic variation in life history traits, which relate to an organism's reproduction and survival, and thereby are closely linked with fitness. We hypothesize that epistatic interactions between the mitochondrial and nuclear genomes influence the rate at which nuclear genetic variation is lost because selection is less effective at eliminating nuclear alleles from a population when they can be associated with both high and low fitness phenotypes. To test this hypothesis, we measured and contrasted how the additive genetic

variance in several fitness components of *C. maculatus* changed in populations with a single mitochondrial haplotype as compared with those populations in which two genetically distinct mitochondrial haplotypes were present. We predict that the loss over time of genetic variation, measured as additive genetic variance, in key traits will be significantly greater in populations with a single mitochondrial haplotype than in those with two mitochondrial haplotypes. Support for our prediction requires first that significant differences between timepoints emerge in either single-haplotype populations, double-haplotype populations, or both. If additive genetic variance changes faster in populations with one haplotype than in populations with two haplotypes (or if one type of line, single or double, does not change but the other does), we support the assertion that the number of mt-haplotypes present affects the rate at which additive genetic variance is depleted.

Methods

Experimental Design

To test this prediction, we used cybrid populations of *Callosobruchus maculatus* descended ancestrally from BF and SI wild populations (see the Cybrid Line Development, Study System sections of the Introduction to the dissertation). In brief, we cultivated cybrid lines in which individuals possessed BF nuclear alleles and SI mitochondrial genomes. We initially produced forty such lines. However, only twelve continued producing viable offspring long enough to reach full introgression and mitochondrial sequencing fifteen generations after initiation. Unless otherwise noted, we reared all populations and individuals according to stock population protocols (see the Stock Rearing Protocols in the Introduction to the dissertation).

Mitochondrial Sequencing

We used nanopore adaptive sequencing to obtain the full mitochondrial genome sequences of individuals belonging to each extant cybrid line. We extracted tissue from seven organisms per line using a tissuelyzer, and followed the standard QIAGen DNEasy protocol for DNA extraction using the QIAGen DNEasy kit. We then followed DNA library preparation and sequencing protocols for the nanopore minion using the Native Barcoding Kit 24 V14 (SQK-NBD114.24). We used GENEIOUS to identify genetic differences in mitochondrial genomes based on sequence data.

Creating Populations for Additive Genetic Variance Estimation

To measure additive genetic variance in the suite of traits mentioned (see Trait Measurement for details), we cultivated large populations seeded with offspring from the two chosen cybrid lineages that differed most in mitochondrial nucleotide base pair differences. We sampled approximately 150 female individuals from each line and monogamously paired them with haphazardly sampled BF males from the lab stock populations (see Study System) in 60mm petri dishes with 2.4g of mung beans present. After mating was complete, we removed males and allowed females to lay eggs freely on beans for twenty-four hours. After removing females from dishes, we gathered and gently shuffled the beans from all pairings (approximately 360 mL by volume) of a given line for sixty seconds in a clean plastic bin and then split them into two containers per line (containers: 15cm width 15cm length 4cm height). We produced the single-haplotype populations by seeding a small, plastic bin (see Stock Rearing Protocols for more details) with two-thirds of the beans collected from that line's pairings (approximately 240 mL by volume). We combined the remaining third from both the Relish and Pesto lines in a separate, small plastic bin to form the two-haplotype population, or "RandP" (i.e. "Relish and Pesto"). We allowed these populations to live and breed freely for five generations following the laboratory protocols for our stock culture maintenance (see Introduction, Stock Rearing) to allow for adjustment to the new environment, after which point we took steps to maximize nuclear genetic diversity in each population. To do this, we sampled more than 200 beans (with eggs) from each population and separated them into microcentrifuge tubes to eclose singly. Once those individuals emerged as adults, we

sexed them and then mated all of those we identified as female to a unique BF stock population male. To maximize additive genetic variance and equalize representation across both single-haplotype populations and the double-haplotype population, we selected 108 of the Pesto/BF pairings and the Relish/BF pairings and represented them evenly in their respective single-haplotype populations. To found both the final Relish and Pesto populations, we deposited ten eggs with beans on them from each of 108 line-specific pairings (1080 beans total to found each population), and to found "RandP", we sampled five beans with eggs on them from all 108 Pesto pairings and five beans with eggs on them from the 108 Relish pairings (Figure 2), evenly representing the number of individuals founding each population (i.e. 1080). We allowed these individuals to emerge and mate freely, and we considered their offspring to be "generation 0" (to be sampled for additive genetic variance estimation as "timepoint 1"). The above description represents the protocol for producing a single replicate, and we repeated the entire process for a second replicate, beginning from the sampling process, one generation later (i.e. replicate one and replicate two were sampled to produce the breeding design for timepoint one approximately 26 days offset from one another).

We allowed unsampled individuals from generation zero to remain in the population boxes, emerge, and mate freely. We similarly allowed all individuals from subsequent generations until generation 9 (timepoint 2) to live and mate normally in the population box environment, and we refreshed the supply of mung beans once per generation, discarding already-parasitized beans and deceased adults in accordance with stock rearing protocols (see Stock Rearing Protocols in the Introduction section of the dissertation). At generation

9, we repeated the sampling process (from generation zero) once more exactly as described above.

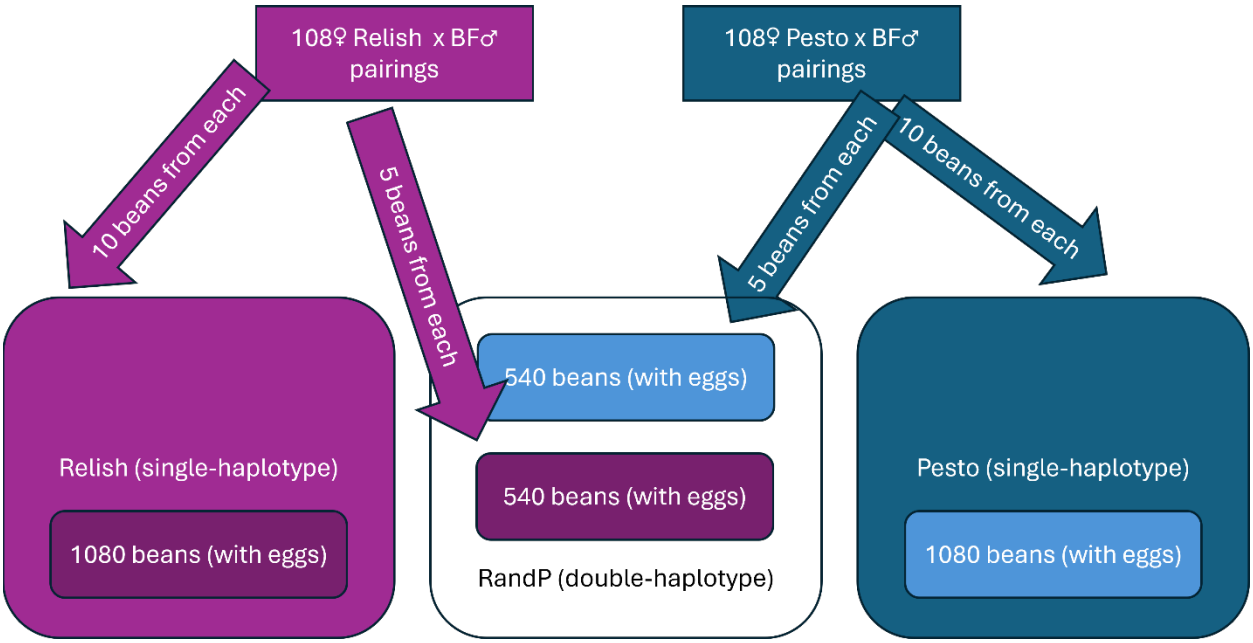


Figure 2: Division of beans with eggs laid upon them by population.

We seeded the red box (left) with individuals from the Relish cybrid line, and the blue box (right) with individuals from the Pesto line. We seeded the central box (white) with an equal number of Pesto and Relish individuals which were full siblings to those in the Pesto and Relish boxes.

Additive Genetic Variance Estimation

We estimated additive genetic variance for adult longevity, total longevity, and mass at emergence for all populations in both replicates at two timepoints. To do this at timepoint one (generation zero) and at timepoint two (generation nine) for each replicate, we haphazardly sampled more than 200 individual beans with eggs from each population after all the adults had died, placed each bean with eggs on it into its own 35mm petri dish for larval development and pupation, sexed them after emergence and measured their masses. We then randomly selected twenty individuals to form the grandparent (GP) generation of a double-first cousin breeding design (Fairbairn and Roff [2006] for each

population at each timepoint, see Figure 2). This design structure offers several advantages over others (e.g. the half-sibling design) in that it produces many kinds of relationships, aside from the titular double-first cousins, useful for partitioning phenotypic similarity among relatives into the portion due to additive genetic variance, including (non-exhaustively) full-siblings, half-siblings, parents, cousins, aunts, uncles, and grandparents (see Meyer 2008 for extensive detailing of the design).

Power Analysis

We ran a power analysis to determine whether a half-sibling or double-first cousin breeding design would better provide the necessary structure to successfully estimate two, timepoint-specific VA values with error values sufficiently small to allow us to identify a significant difference between them. We selected 0.1 to be the minimum realistic expected difference in VA across timepoints given lab-specific experience with estimation of VA of these traits in our stock populations of *C. maculatus*. Using the nadiv package (Wolak 2012), we simulated either half-sib or double-first cousin pedigrees for two timepoints, assigning individuals phenotypes drawn randomly from a normal distribution with a sensible and likely trait mean (e.g. 5mg for mass at emergence) and variance of 1, and we assigned individuals breeding values drawn from a normal distribution with a mean of zero and a variance of our additive genetic variance value of interest. We set the initial additive genetic variance values of interest and the secondary timepoint additive genetic variance values of interest based on the change in additive genetic variance we wanted to test (e.g. two additive genetic variance values of interest of 0.4 and 0.1 could test an expected drop in VA of 0.3). After simulating populations, we estimated an additive genetic

variance component for each of the simulated populations using ASReml. We then ran a likelihood ratio test with a chi-squared distribution and single degree of freedom to determine whether a model that evaluated a single additive genetic variance estimate which combined both timepoints was a significantly better fit than a model that estimated the two timepoints' VA values independently. Note that we held phenotypic variance stable at 1 in all cases to simulate different possible narrow-sense heritability values (narrow-sense heritability [h^2] = additive genetic variance [VA] / total phenotypic variance [VP]). We repeated this process, simulating varying population sizes and structures one-thousand times each and counting the proportion of instances over which the likelihood ratio test yielded a p-value of less than or equal to 0.05 (see a subset of simulation parameters and p-values in Table 1), selecting 0.85 as our minimum threshold of statistical power for identifying changes in VA.

Planned Breeding Design

We determined that the optimal population structure necessary to successfully distinguish a change in additive genetic variance of 0.1 from timepoint one to timepoint two (given population size and logistical constraints) called for a double-first cousin breeding design constituted of the following: for each of five blocks, where a block is functionally a replicate within the breeding design, we mated four pairs of unrelated grandparent (GP) individuals to produce full-sib offspring per parent (P) generation family. Two sons from each P generation full-sib family each mated with one unique daughter (i.e. females mate just once per lifetime) from each of the other (unrelated) P generation full-sib families. This produces, among many other kinds of relationships, double-first-cousin offspring in the F1

generation. Our specific structure yielded an estimated total pedigree size (one per replicate per haplotype per timepoint, resulting in 12 pedigrees total) of at least 560 measured individuals. We phenotyped at least five individuals in each full-sib family of the GP, P, and F1 generations (with the exception of total longevity in the population-sampled GP individuals because we did not know the date upon which they were laid as eggs).

Table 1: Sample power analysis values for a double-first cousin breeding design (selected structure in gray).

VA Change	Blocks	Grandparent (GP) Pairings (per block)	P-Gen Full-Sib Sires Per Family (per block)	P-Gen Families With Sires (per block)	Offspring Per Family	Proportion <= 0.05
0.1	4	5	2	2	5	0.85
0.1	4	5	4	3	6	0.85
0.1	4	4	3	2	9	0.86
0.1	4	4	3	4	15	0.86
0.1	4	5	2	2	9	0.87
0.1	4	5	3	4	6	0.87
0.1	4	4	3	3	9	0.88
0.1	4	4	4	3	5	0.88
0.1	4	4	4	4	8	0.88
0.1	4	5	2	3	8	0.88
0.1	4	5	2	4	15	0.88
0.1	4	5	3	4	7	0.88
0.1	4	5	3	4	18	0.88
0.1	4	4	3	4	11	0.89
0.1	4	4	3	4	16	0.89
0.1	4	5	3	3	5	0.89
0.1	4	5	4	2	9	0.89
0.1	4	5	2	4	11	0.9
0.1	4	5	4	2	10	0.9
0.1	4	4	3	4	12	0.91
0.1	4	5	3	3	17	0.91

Realized Breeding Design

In timepoint one, we faced the limitation that males often died after multiple matings but not sufficient matings to meet the structural requirements laid out by the breeding scheme. We therefore amended the structure slightly. Within each block, instead of mating a pair of full-sibling males from four of five families each to one unique sister from each of the four other within-block families, we mated a pair of full-sibling males from all five within-block families each to unique full-sibling females from three other families, therefore creating the same kind of unique relationships (e.g. cousins, double-first cousins, aunts, uncles, etc.) across all families represented within a block. In timepoint two, males were better able to persist through multiple matings and we adhered to (or exceeded by phenotyping more individuals or producing more pairings) the intended structure provided by the power analysis. We selected individuals to participate in the breeding design by using a random number generator to choose from among full-siblings within the families represented.

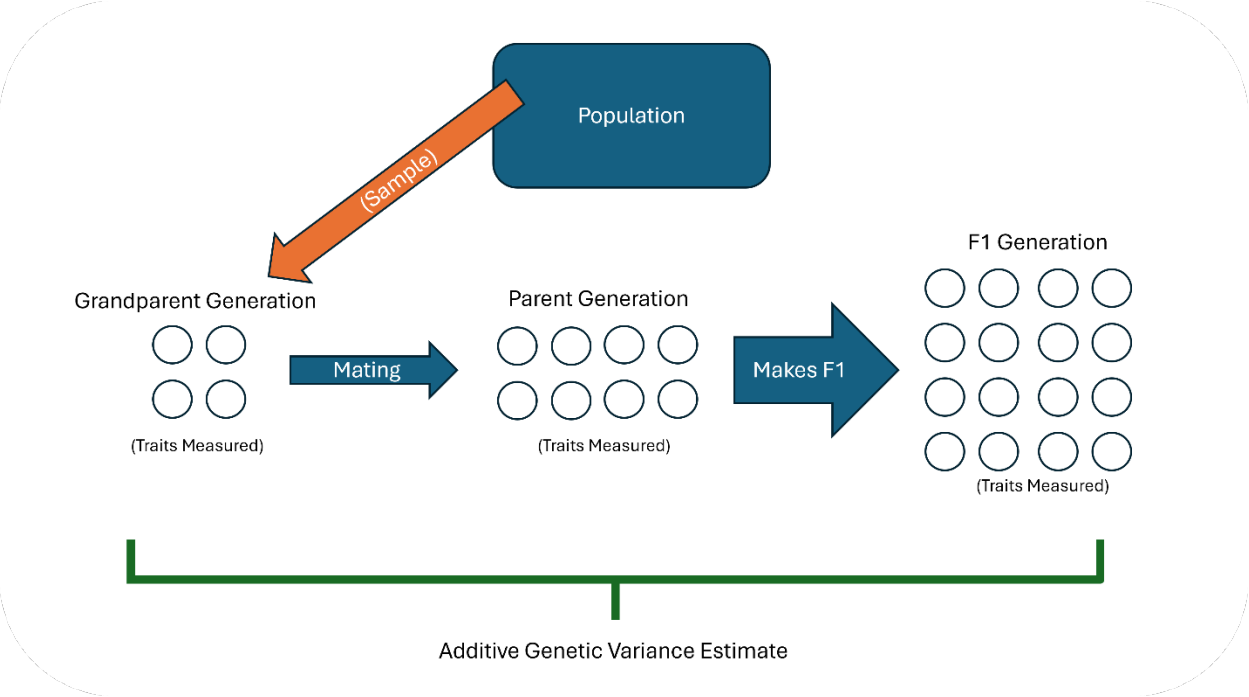


Figure 3: VA Estimation using a double first cousin (DFC) design.

During the breeding designs implemented at each timepoint, we measured the following traits in each experimental population: mass at adult emergence in mg—we weighed the beetles on a microbalance on the day of eclosion, adult longevity, which spanned the days between eclosion and death, and total longevity, or days elapsed between when that beetle was laid as an egg to until its death in a petri dish. These are all key life history traits and fitness components in *Callosobruchus maculatus* (Kawecki 1995). We recorded emergence, sexed, and measured mass in all adults within twenty-four hours of eclosure and recorded deaths within one day of an individual's demise.

Statistical Methods

To measure VA in each population, we implemented a quantitative genetic linear mixed model, or animal model (Henderson 1950; Kruuk 2004; Wilson et al. 2010) using Asreml-R (Version 4.2, Butler et al. 2007), treating mass at emergence, adult longevity, and total longevity as normally-distributed continuous response variables (modeled independently). The animal model uses the matrix inverse of an A-matrix, a genetic relationship matrix that measures the expected identity-by-descent of individuals within a pedigree, combined with an additive genetic variance/covariance matrix (i.e. a G-Matrix), to measure the phenotypic similarity among relatives that can be attributed to shared additive genetic effects, thus estimating the additive genetic variance within the population of their shared ancestors (Henderson 1976, 1986; Meyer 1989, 1991; Wilson et al. 2010). In each model, we included generation, timepoint, and their interaction as categorical fixed effects and random effects of individual identity linked to the inverse A-matrix to estimate VA and of dam identity to estimate variance due to maternal common environment effects. Additionally, for mass at emergence models, we included categorical fixed effects of the balance used to weigh beetles as well as random effects of measurer identity. Additionally, we estimated timepoint specific additive genetic variance for each population within a replicate by fitting two separate identity terms: one for the first timepoint (generation zero) and one for the second timepoint (generation nine). We included these as separate, independent sets of random effects, each with their own pedigree-based inverse A-matrix, given that the timepoint-specific pedigrees were independent of one another and therefore there is no information to estimate an additive genetic covariance between timepoints

(though individuals across pedigrees may possibly have been related, we did not collect any data to link them). We similarly estimated timepoint-specific residual variances. However, in most cases we only estimated a single dam variance across timepoints due to model convergence issues that arose when attempting to estimate timepoint-specific dam variances. The convergence issues were caused by one or both time-point specific dam variances became fixed at effectively zero and so were inestimable and prevented models from converging.

To determine whether a significant change in VA had occurred between timepoint one and timepoint two, we employed the same methodology as in the power analysis. We employed the approach suggested by Shaw (Shaw 1991) and ran a likelihood ratio test (LRT) with a chi-square distribution and one degree of freedom comparing each model above (i.e. time-point specific additive genetic variances estimated) to a reduced model in which a single additive genetic variance was estimated across both timepoints. The reduced model collapses the pedigrees of both timepoints into a single pedigree and inverse A-matrix, but retains timepoint-specific residual variances.

We measured the change in VA as $VA_{TP1} - VA_{TP2}$ where a positive difference suggests a decrease in VA and a negative difference suggests an increase. We present the timepoint-specific VA estimates, differences, and standard errors as estimated by ASReml as well as the likelihood ratio test statistic and its corresponding degrees of freedom and p-value.

Results

Selection of Mitochondrial Lineages for VA Measurement

We used nanopore adaptive sequencing to obtain the full mitochondrial genome sequences of individuals belonging to each extant cybrid line. We extracted tissue from seven organisms within a line using a tissuelyzer, and followed the standard QIAGen DNEasy protocol for DNA extraction using the QIAGen DNEasy kit. We then followed DNA library preparation and sequencing protocols for the nanopore minion R10 flow cell (FLO-MIN112, Nanopore) using the Native Barcoding Kit 24 V14 (SQK-NBD114.24), ensuring first that we had secured a volume of at least 400ng of DNA present for each line by assessing the concentrations of DNA present in 28 mL of solution using an Invitrogen Qubit 3.0 Fluorometer. We ran these libraries for 35 hours on the minION flow cell. We then moved the raw minION data to Easley High Performance Computer for bioinformatic processing and mitochondrial genome assembly. First, we summarized the data with NANOPLOT and high-accuracy base-called using Guppy. We assessed the quality of sequence data using FASTQC and split reads into separate files by barcodes, removed barcodes and adapters using PORECHOP, and then reassessed for quality with FASTQC. For each strain, we mapped reads to the reference seed beetle mitochondrial genome (NCBI accession MF960125), using MINIMAP2 and sorted the sam files which we then converted to bam files and assessed coverage using SAMTOOLS. We created the initial draft assembly of the mitochondrial genome sequence for each strain using SAMTOOLS consensus and polished this assembly using MEDAKA. We assessed the polished assemblies for quality using QUAST.

We imported the polished mitochondrial genome assemblies into Geneious and aligned them. We removed highly fragmented assemblies from the alignment. We assessed pairwise genetic variation between the assemblies.

Matrix: % Identity Decimal Places: 3 Style: Heatmap and Numbers

	MF960125	BC14_med...	BC20_med...	BC21_med...	BC23_med...	BC15_med...	BC16_med...	BC13_med...	BC17_med...	BC22_med...	none_meda...
MF960125		98.798%	98.711%	98.711%	98.719%	98.711%	98.705%	98.705%	98.662%	98.694%	98.719%
BC14_medaka_cons.fa	98.798%		98.999%	98.999%	99.007%	98.999%	98.993%	98.993%	98.950%	98.982%	99.007%
BC20_medaka_cons.fa	98.711%	98.999%		100%	99.992%	100%	99.978%	99.978%	99.951%	99.967%	99.992%
BC21_medaka_cons.fa	98.711%	98.999%	100%		99.992%	100%	99.978%	99.978%	99.951%	99.967%	99.992%
BC23_medaka_cons.fa	98.719%	99.007%	99.992%	99.992%		99.992%	99.986%	99.986%	99.943%	99.974%	100%
BC15_medaka_cons.fa	98.711%	98.999%	100%	100%	99.992%		99.978%	99.978%	99.951%	99.967%	99.992%
BC16_medaka_cons.fa	98.705%	98.993%	99.978%	99.978%	99.986%	99.978%		99.972%	99.935%	99.961%	99.986%
BC13_medaka_cons.fa	98.705%	98.993%	99.978%	99.978%	99.986%	99.978%	99.972%		99.951%	99.967%	99.986%
BC17_medaka_cons.fa	98.662%	98.950%	99.951%	99.951%	99.943%	99.951%	99.935%	99.951%		99.923%	99.943%
BC22_medaka_cons.fa	98.694%	98.982%	99.967%	99.967%	99.974%	99.967%	99.961%	99.967%	99.923%		99.974%
none_medaka_cons.fa	98.719%	99.007%	99.992%	99.992%	100%	99.992%	99.986%	99.986%	99.943%	99.974%	

Figure 4: Sequence differences by mt-haplotype

Nucleotide base pair percent sequence differences by mitochondrial haplotype (Compared with the NCBI *Callosobruchus maculatus* reference genome (Zhang et al. 2018). “Pesto” corresponds to BC14, and “Relish” corresponds to BC17.

Additive Genetic Variance Estimates

We measured the additive genetic variance (VA) in mass at emergence, adult longevity, and total longevity of each population by sex. We estimated VA independently for each population (haplotype x replicate) but included both timepoint one (generation 0) and timepoint two (generation 9) for a given population in each model. In all pedigrees, we were able to produce the necessary 560 individuals in the required structure to be necessary to estimate statistically significant changes in VA from timepoint one to timepoint two of at least 0.1 given the results of the power analysis (see Table 1). Given the parameters of the power analysis simulation, the given sample size and breeding design produced sufficient power to detect differences in VA of 0.1 or greater while balancing the logistical constraints

imposed by achieving such a sample size for three populations and once and with two generations of overlap during breeding designs between all six populations of the two replicates.

Table 2: Number of phenotyped individuals by population, replicate, and timepoint. “r1” and “r2” indicate replicates 1 and 2 respectively.

Population	Timepoint 1 (Generation 0)	Timepoint 2 (Generation 9)
Relish (r1)	n = 1173	n = 1005
Pesto (r1)	n = 1154	n = 1051
RandP (r1)	n = 1189	n = 1041
Relish (r2)	n = 1088	n = 1055
Pesto (r2)	n = 1070	n = 1096
RandP (r2)	n = 1081	n = 1077

Females

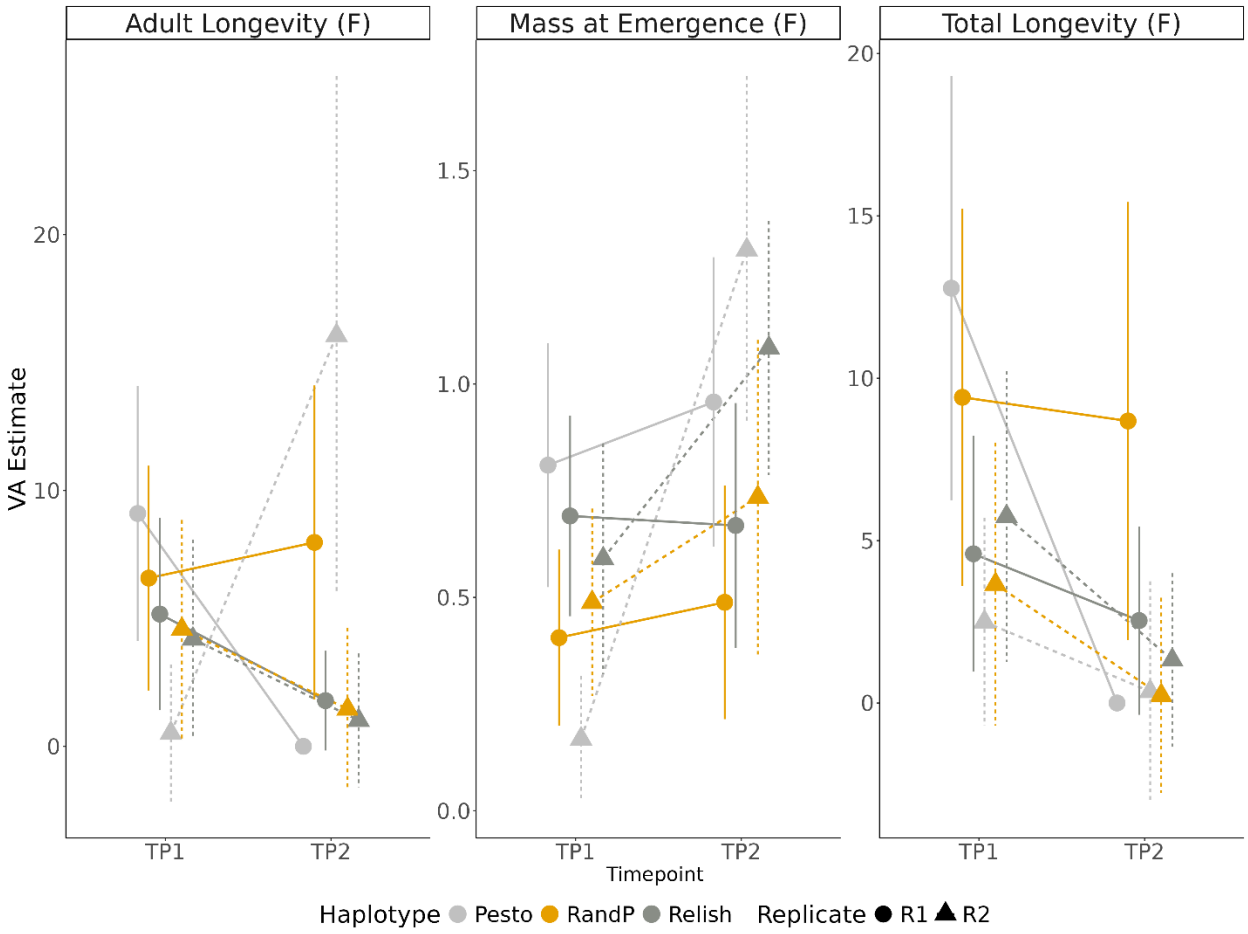


Figure 5: Female (F) Additive Genetic Variance (VA), by Timepoint (TP1 and TP2) in Adult Longevity, Mass at Emergence, and Total Longevity.

Error bars represent the estimate +/- twice the standard error. Solid lines and circles represent replicate one (R1), and dotted lines and triangles represent replicate two (R2). Missing error bars on Pesto, replicate 1 adult and total longevities are due to the model's inability to successfully estimate error due to boundary values.

For females, in the double-haplotype populations (RandP), we found no significant changes in VA of any trait in either replicate. However, for both single-haplotype lines, we found strong evidence that VA in mass at emergence rose for females in replicate two

(Pesto line ΔVA = +1.146, p = 1.5e-06, Relish line ΔVA = +0.494, p = 4.91e-02). We found no other evidence that female mass at emergence VA values changed between timepoint one (generation zero) and timepoint two (generation nine). We found very strong evidence that VA in adult longevity fell for females in the Pesto replicate one line (ΔVA = -9.101, p = 7.52e-06) and moderate evidence that it rose significantly in the Pesto replicate two line (ΔVA = +15.548, p = 2.61e-03). The data showed weak evidence that VA fell for Relish replicate one (ΔVA = -3.383, p = 8.84e-02). We did not find any evidence that VA changed for Relish replicate two. We found no evidence of either RandP (double-haplotype) replicate line changing in VA for adult longevity. We found very strong evidence that VA in total longevity fell for females in Pesto replicate one (ΔVA = -12.722, p = 2.06e-06) and weak evidence that it fell in Relish replicate two population (ΔVA = -4.42, p = 9.46e-02) but no evidence of a change in Relish replicate one. Neither RandP (double-haplotype) replicate line changed significantly in VA for total longevity.

(Note: The total longevity model for Pesto, replicate two required a by-timepoint estimate of dam to converge.)

Table 3: Female Additive Genetic Variance (VA) in Mass at Emergence

	Trait	Haplotype	Rep	Sex	TP1 VA	TP1 Error	TP2 VA	TP2 Error	ΔVA	LRT	df	P
19	Mass at Emergence	RandP	R1	F	0.406	0.105	0.488	0.14	0.083	0.158	1	6.906543e-01
20	Mass at Emergence	RandP	R2	F	0.488	0.112	0.735	0.188	0.246	0.945	1	3.311108e-01
21	Mass at Emergence	Pesto	R1	F	0.81	0.146	0.958	0.173	0.148	0.399	1	5.275799e-01
22	Mass at Emergence***	Pesto	R2	F	0.168	0.075	1.315	0.208	1.146	23.154	1	1.495089e-06
23	Mass at Emergence	Relish	R1	F	0.691	0.12	0.668	0.146	-0.022	0.012	1	9.112582e-01
24	Mass at Emergence**	Relish	R2	F	0.591	0.138	1.084	0.152	0.494	3.874	1	4.905364e-02

“TP” represents “timepoint”. “Error” represents the standard error of the additive genetic variance estimate. “df” represents the degrees of freedom in the likelihood ratio test. “LRT” represents the likelihood ratio test statistic. “P (LRT)” represents the p-value given by the likelihood ratio test comparing the two timepoint estimates of VA. * p-value<= 0.1, ** p-value<= 0.05, *** p-value<= 0.001

Table 4: Female Additive Genetic Variance (VA) in Adult Longevity

	Trait	Haplotype	Rep	Sex	TP1 VA	TP1 Error	TP2 VA	TP2 Error	ΔVA	LRT	df	P
25	Adult Longevity	RandP	R1	F	6.576	2.241	7.972	3.13	1.395	0.089	1	7.660183e-01
26	Adult Longevity	RandP	R2	F	4.578	2.179	1.459	1.617	-3.118	1.223	1	2.686764e-01
27	Adult Longevity***	Pesto	R1	F	9.101	2.542	0	NA	-9.101	20.056	1	7.519994e-06
28	Adult Longevity**	Pesto	R2	F	0.515	1.368	16.063	5.168	15.548	9.058	1	2.614945e-03
29	Adult Longevity*	Relish	R1	F	5.17	1.916	1.788	0.994	-3.383	2.903	1	8.843412e-02
30	Adult Longevity	Relish	R2	F	4.194	1.984	1.017	1.334	-3.177	1.943	1	1.632995e-01

“TP” represents “timepoint”. “Error” represents the standard error of the additive genetic variance estimate. “df” represents the degrees of freedom in the likelihood ratio test. “LRT” represents the likelihood ratio test statistic. “P (LRT)” represents the p-value given by the likelihood ratio test comparing the two timepoint estimates of VA. Note that the adult longevity model for Pesto, R1, TP2 produced a boundary fit and a zero estimate for VA and was unable to estimate its error. * p-value<= 0.1, ** p-value<= 0.05, *** p-value<= 0.001

Table 5: Female Additive Genetic Variance (VA) in Total Longevity

	Trait	Haplotype	Rep	Sex	TP1 VA	TP1 Error	TP2 VA	TP2 Error	ΔVA	LRT	df	P
31	Total Longevity	RandP	R1	F	9.412	2.965	8.683	3.44	-0.729	0.022	1	8.833891e-01
32	Total Longevity	RandP	R2	F	3.656	2.222	0.234	1.534	-3.422	1.948	1	1.628321e-01
33	Total Longevity***	Pesto	R1	F	12.772	3.331	0	NA	-12.772	22.533	1	2.065710e-06
34	Total Longevity	Pesto	R2	F	2.496	1.634	0.353	1.732	-2.144	0.856	1	3.548793e-01
35	Total Longevity	Relish	R1	F	4.594	1.854	2.538	1.481	-2.056	0.745	1	3.882015e-01
36	Total Longevity*	Relish	R2	F	5.743	2.287	1.324	1.364	-4.42	2.795	1	9.458168e-02

“TP” represents “timepoint”. “Error” represents the standard error of the additive genetic variance estimate. “df” represents the degrees of freedom in the likelihood ratio test. “LRT” represents the likelihood ratio test statistic. “P (LRT)” represents the p-value given by the likelihood ratio test comparing the two timepoint estimates of VA. * p-value<= 0.1, ** p-value<= 0.05, *** p-value<= 0.001

Males

In the double-haplotype populations (RandP), we found no evidence of changes in VA of any trait in either replicate. We found strong evidence of a rise in VA in mass at emergence among males in three of four single-haplotype replicate populations (Pesto R1 ΔVA = 0.574, p = 2.02e-04, Relish R1 ΔVA = +0.37, p = 1.45e-02, Relish R2 ΔVA = +0.222, p = 2.87e-02), but no evidence of a change in Pesto replicate two. We found weak evidence of a change in VA in adult longevity for males in the Pesto, replicate two population (Pesto R2 ΔVA = 4.699, p = 6.48e-02) but no evidence of a change in replicate one, and weak evidence of a descent in VA for Relish replicate one (Relish line R1 ΔVA = -3.6, p = 5.57e-02). We found no evidence of changes in VA to any other single-haplotype populations, and we found no evidence of changes to VA in total longevity for any population.

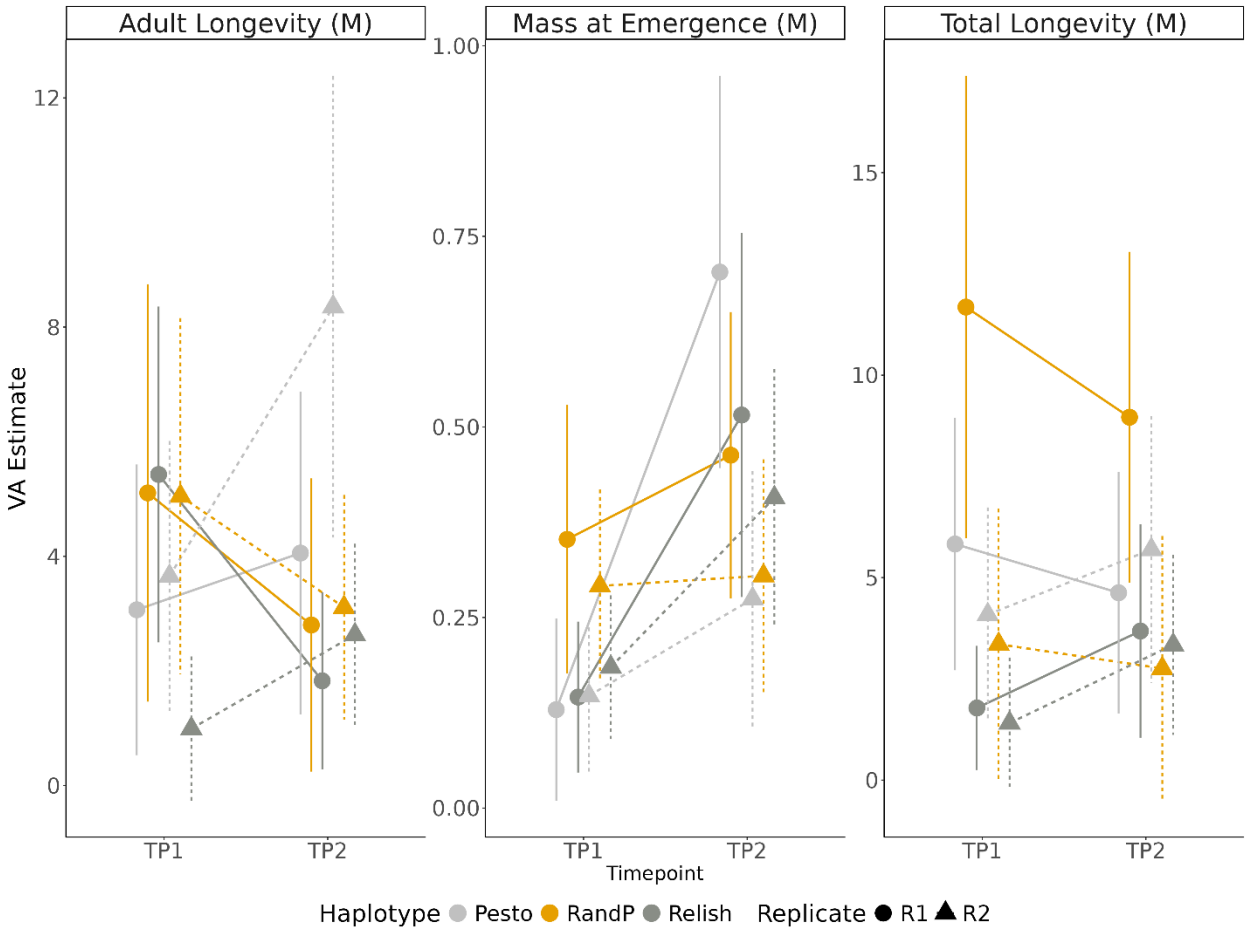


Figure 6: Male (M) Additive Genetic Variance (VA), by Timepoint (TP1 and TP2) in Adult Longevity, Mass at Emergence, and Total Longevity.

Error bars represent the estimate +/- twice the standard error. Solid lines and circles represent replicate one (R1), and dotted lines and triangles represent replicate two (R2).

Table 6: Male Additive Genetic Variance (VA) in Mass at Emergence

Trait	Haplotype	Rep	Sex	TP1 VA	TP1 Error	TP2 VA	TP2 Error	ΔVA	LRT	df	P
Mass at Emergence	RandP	R1	M	0.353	0.09	0.463	0.096	0.11	0.469	1	4.936254e-01
Mass at Emergence	RandP	R2	M	0.291	0.064	0.305	0.078	0.013	0.015	1	9.014321e-01
Mass at Emergence***	Pesto	R1	M	0.129	0.061	0.703	0.131	0.574	13.81	1	2.022776e-04
Mass at Emergence	Pesto	R2	M	0.147	0.051	0.275	0.085	0.127	1.363	1	2.429864e-01
Mass at Emergence**	Relish	R1	M	0.146	0.051	0.516	0.122	0.37	5.968	1	1.456966e-02
Mass at Emergence**	Relish	R2	M	0.185	0.048	0.407	0.086	0.222	4.783	1	2.874359e-02

“TP” represents “timepoint”. “Error” represents the standard error of the additive genetic variance estimate. “df” represents the degrees of freedom in the likelihood ratio test. “LRT” represents the likelihood ratio test statistic. “P (LRT)” represents the p-value given by the likelihood ratio test comparing the two timepoint estimates of VA. * p-value<= 0.1, ** p-value<= 0.05, *** p-value<= 0.001

Table 7: Male Additive Genetic Variance (VA) in Adult Longevity

	Trait	Haplotype	Rep	Sex	TP1 VA	TP1 Error	TP2 VA	TP2 Error	ΔVA	LRT	df	P
7	Adult Longevity	RandP	R1	M	5.108	1.858	2.803	1.305	-2.304	1.651	1	1.988403e-01
8	Adult Longevity	RandP	R2	M	5.051	1.585	3.112	1	-1.939	0.963	1	3.263169e-01
9	Adult Longevity	Pesto	R1	M	3.068	1.294	4.058	1.437	0.99	0.355	1	5.512148e-01
10	Adult Longevity*	Pesto	R2	M	3.655	1.204	8.354	2.054	4.699	3.41	1	6.481944e-02
11	Adult Longevity*	Relish	R1	M	5.429	1.494	1.829	0.789	-3.6	3.66	1	5.573870e-02
12	Adult Longevity	Relish	R2	M	0.994	0.643	2.632	0.81	1.638	2.518	1	1.125877e-01

“TP” represents “timepoint”. “Error” represents the standard error of the additive genetic variance estimate. “df” represents the degrees of freedom in the likelihood ratio test. “LRT” represents the likelihood ratio test statistic. “P (LRT)” represents the p-value given by the likelihood ratio test comparing the two timepoint estimates of VA. * p-value<= 0.1, ** p-value<= 0.05, *** p-value<= 0.001

Table 8: Male Additive Genetic Variance (VA) in Total Longevity

	Trait	Haplotype	Rep	Sex	TP1 VA	TP1 Error	TP2 VA	TP2 Error	ΔVA	LRT	df	P
13	Total Longevity	RandP	R1	M	11.68	2.91	8.96	2.084	-2.72	0.588	1	4.432250e-01
14	Total Longevity	RandP	R2	M	3.354	1.708	2.756	1.669	-0.598	0.09	1	7.646924e-01
15	Total Longevity	Pesto	R1	M	5.829	1.591	4.626	1.52	-1.204	0.289	1	5.907812e-01
16	Total Longevity	Pesto	R2	M	4.084	1.351	5.693	1.681	1.609	0.55	1	4.581451e-01
17	Total Longevity	Relish	R1	M	1.78	0.782	3.681	1.344	1.9	1.564	1	2.110403e-01
18	Total Longevity	Relish	R2	M	1.407	0.826	3.339	1.134	1.931	2.059	1	1.512642e-01

“TP” represents “timepoint”. “Error” represents the standard error of the additive genetic variance estimate. “df” represents the degrees of freedom in the likelihood ratio test. “LRT” represents the likelihood ratio test statistic. “P (LRT)” represents the p-value given by the likelihood ratio test comparing the two timepoint estimates of VA. * p-value<= 0.1, ** p-value<= 0.05, *** p-value<= 0.001

Discussion

Summary of Findings

We hypothesized that epistatic interactions between the mitochondrial and nuclear genomes influence the rate at which nuclear genetic variation is lost, because selection is less effective at eliminating nuclear alleles from a population when they can be associated with both high- and low-fitness phenotypes. We predicted that populations with two mitochondrial haplotypes would, on average, lose additive genetic variance less quickly over many generations than populations with a single mitochondrial haplotype. We found mixed support for both this hypothesis and its associated prediction. In keeping with our expectations, all double-haplotype VA values remained stable in that they did not change significantly for any trait. Single-haplotype populations, however, varied considerably in that VA estimates both rose and fell in both sexes, though the orientations of those changes were sex-specific in that a direction of change in VA in one sex did not necessarily predict the direction of change (or a change at all) in the other. Below, we outline whether we found support for our hypothesis and prediction by traits nested within sex. We categorized a trait as providing strong support for our hypothesis when all or most single-haplotype populations changed and the double-haplotype populations did not change, as providing support when the double-haplotype populations did not change but half or fewer of the single-haplotype populations changed, and no support when no populations changed. We categorized a trait as providing strong support for our prediction when both of the double-haplotype populations did not change but most or all single-haplotype

populations fell, as mixed support when the double-haplotype populations did not change but the single-haplotype populations changed in multiple directions (e.g. some rose and others fell), and no support when no changes occurred across any populations.

Table 9: Support for our hypothesis and prediction by sex and trait.

Males			
Trait	Mass at Emergence	Adult Longevity	Total Longevity
Prediction	**No Support** No replicate population lost VA in mass at emergence	**Mixed Support** No significant changes to VA in double-haplotype replicates, but both significant increases and decreases for VA in single-haplotype replicate populations	**No Support** No significant changes in VA for any replicate population
Hypothesis	**Strong Support** No change to VA in either double-haplotype replicate but significant increases in VA in three of four single-haplotype replicate populations	**Support** Both double-haplotype replicates maintained stable VA values whereas two of four single-haplotype replicate populations saw (near) significant changes in VA	**No Support** No significant changes in VA for any replicate population
Females			
Trait	Mass at Emergence	Adult Longevity	Total Longevity
Prediction	**No Support** No replicate population lost VA in mass at emergence	**Mixed Support** No significant changes to VA in double-haplotype populations but rises in VA for two single-haplotype replicates and a (nearly) significant decline in VA for one single-haplotype replicate population	**Strong Support** No significant changes in VA in double-haplotype populations but significant declines in three of four single-haplotype replicate populations
Hypothesis	**Support** No change to VA in either double-haplotype replicate but significant changes to half of the single-haplotype replicates	**Strong Support** No change to double-haplotype replicate VA but significant (or near-significant) changes to three of four single-haplotype replicates	**Support** No change to double-haplotype replicate VA but significant (or near-significant) changes to two of four single-haplotype replicates

The effects of mitochondrial diversity on additive genetic variance

A long-standing question in evolutionary biology asks what sustains the widespread genetic variation, which we measure as additive genetic variance (or VA), in nature. Implicit in this question is a puzzle: what allows the persistent, and often strong, selection that we observe operating on wild populations to continue without leading to widespread population extinction events? Aside from the production of novel genetic material through mutations, the answers proffered by existing work support, in one form or another, the preservation of existing alleles. Proposed mechanisms include negative frequency-dependent selection (Kazancioğlu and Arnqvist 2014; Brisson 2018), situational dominance reversals (Grieshop and Arnqvist 2018), or genotype-by-environment interactions, which describe a phenomenon whereby the phenotypic expression of a single genotype varies based on its backdrop. In the latter, the "environment" could represent a new geographical or ecological landscape, a different sex, a different time, or even genotype-by-genotype-by-environment interactions that involve an ecological backdrop expressed against the nuclear genome and its interactions with a secondary organellular or endosymbiotic genome (such as the genomes in mitochondria, chloroplasts, or Wolbachia) (Turelli and Barton 2004; Agashe et al. 2023).

A recent review of the evolutionary genetics of mitochondria by Dowling and Wolff (2023) explores a multitude of ways in which mitochondrial diversity influences trait values, resistance to disease, and even fitness, but acknowledges the knowledge gap with respect to how this multigenomic interplay affects evolutionary trajectories, emphasizing the

importance of evaluating how mitonuclear function affects evolutionary trajectories over time (Dowling and Wolff 2023).

Our findings that double-haplotype populations' estimates of additive genetic variance did not change significantly over time in contrast to the single-haplotype populations suggests that mitochondrial diversity may serve as a key stabilizing mechanism for the all-important genetic diversity that populations must produce or preserve to respond to selection. The stable additive genetic variance estimates of the double-haplotype populations provide evidence of balancing selection. As with intralocus sexual conflict, we find evidence that mitochondrial diversity may prevent selection from removing or fixing mitonuclear (Nmt) alleles at a given locus (Bonduriansky and Chenoweth 2009). This aligns with previous work that suggests that mitochondrial haplotypes, even introduced at different frequencies into a population (Kazancioğlu and Arnqvist 2014), often shift in frequency toward equilibrium, which could suggest roughly even fitness consequences for individuals equipped with either haplotype provided that a compatible mitonuclear complement of alleles exists in the population for both mitochondrial types (i.e. as long as fitness is not greatly reduced in carriers of one haplotype on average). Both of our double-haplotype population replicates began with exactly even frequencies of each mitochondrial haplotype, so we expect that given that mitochondrial frequencies trend toward equilibrium, two haplotypes would likely remain continue to occur at a relatively similar rate in our double-haplotype populations.

We find also that the magnitudes of changes in additive genetic variance were often substantial, frequently halving or doubling VA (in one case, VA increased by almost seven-

fold, see Table 3, Mass at Emergence for females, Pesto, R2). This is especially surprising considering the relatively small proportion of the full genome expected to be composed of Nmt genes. For instance, although the exact number of nuclear-encoded Nmt genes in *C. maculatus* is currently unknown, in *Drosophila melanogaster*, a species with a comparable number of total protein-coding genes (D. *melanogaster* ~13,500, *C. maculatus* = 14,458 Lu et al. 2024) nuclear-encoded Nmt genes compose just ~3.6% of the full genome (498 nuclear-encoded Nmt (Gallach et al. 2010) out of 13593 total, (Eslamieh et al. 2017). If mitonuclear alleles similarly constitute less than 4% of the total genome in *C. maculatus* then this small fraction of the genome alone accounts for severalfold increases in additive genetic variance for some key traits such as mass at emergence.

Sex-Specific Differences

In their 2023 review of the genetics of mitochondria, Dowling and Wolff make a case for the importance of investigating sex-specific differences in mitonuclear effects on populations over multiple generations both to further test the validity of the Mother's Curse hypothesis (Keaney et al. 2020; Leeflang et al. 2021) and to determine whether or not sexually antagonistic selection constrains traits heavily influenced by mitochondrial function (Dowling and Wolff 2023). We found major differences in the directionality and magnitude of changes to additive genetic variance in adult longevity between male and female *C. maculatus* as well as a surprisingly consistent series of increases to VA in measured male traits, especially body mass at adult emergence, a trait known to be under strong sexually antagonistic selection (Berger et al. 2014a, p. 20; Kaufmann et al. 2023). These results could be consistent with the mother's curse hypothesis (Dowling and Adrian 2019; Keaney

et al. 2020) if the male increases in additive genetic variance represent the accumulation of deleterious mutations on a single mitochondrial haplotype that universally produce a positive impact upon females and result in a wider range of potentially disadvantageous or neutral phenotypes for males, though whether this occurred is beyond the scope of our study.

Another plausible explanation is that these differences could also be an artefact of sex-specific expression of mitonuclear alleles which has been found in other taxa, for example, Gallach et al. (Gallach et al. 2010) found testis-specific expression upregulation of Nmt genes in *Drosophila* which may occur through gene duplication and subsequent relocation of the duplicated gene. The incipient shift of trait expression from autosomes to sex chromosomes could also explain sex-specific additive genetic variances, and such transitions have been shown to contribute to the resolution of strong intralocus sexual conflict in mass at emergence in *C. maculatus* (Kaufmann et al. 2021). In our study, we found not only that additive genetic variances differed by sex but also that additive genetic variance rose more often in male traits than in female traits. One explanation for this specific disparity in VA trajectory by sex may result from Nmt genes localized to (or relocated to) the X chromosome. Male X-linked VA is expected to double when dosage compensation occurs as either X-inactivation in females, halving of X-output in females, or doubling of male X-linked output (Kent et al. 2005; Parker et al. 2022, although note that Gallach et al. found that when Nmt genes were duplicated and relocated, it was overwhelmingly out of, rather onto, the X chromosome).

Limitations and Future Directions

We measured VA at a single timepoint at the beginning and a single timepoint at the end of the study, but measures of VA sampled from each population at each generation would be invaluable in identifying the trajectories of population-level additive genetic variance in future studies. These fine-scale measures could be particularly informative given that VA can fluctuate under certain circumstances, such as in environments in which an allele goes to fixation at a locus and collapses non-additive effects into additive effects (or vice versa, Kent et al. 2005; Roff 2009; Clo et al. 2020; Acker et al. 2023). We especially encourage and highlight the need for the identification of the specific nuclear loci involved in these critical mitonuclear interactions to recognize which loci, if any in specific, are instrumental in facilitating the population-level changes we see here, and we echo Dowling and Wolff (2023) in encouraging the use of modern sequencing techniques to combine comprehensive genomic analyses with the quantitative genetic methods applied here.

Chapter 2: When Mito Met Nuc: How MitoNuclear Make-Up Influences Mate Choice in a Seed Beetle

Introduction

Most eukaryotes reproduce sexually (or are at least capable of facultative sexual reproduction) (The Tree of Sex Consortium 2014), despite some major constraints imposed by it (see John Maynard Smith's "two-fold cost of sex" Maynard-Smith 1978), as well as the drawbacks of having two and not more sexes in a dyadic system given that having two additional sexes occurring at equal frequency raises the number of potential mates available to any individual in a population (Krumbeck et al. 2020). Consequently, each product of sexual coupling is a genetic blend of the participating parents, inheriting only an expected 50% of the alleles from each. This process can disrupt and separate highly functional gene complexes (Gibson et al. 2017) and, in some cases, results in offspring that are less competitive than their parents, forcing alleles present in high fitness organismal environments into new combinations every generation. This disruption, however, allows for rapid selection on superior genotypes and sometimes similarly produces novel and effective allelic combinations that can confer population-level benefits, such as the promotion of resistance against adaptively evolving pathogens (Hamilton and Zuk 1982).

The majority of all these sexually reproducing eukaryotes receive most of their energy from oxidative phosphorylation (OXPHOS), an ATP-producing process that occurs during cellular respiration in the mitochondria (Lane and Martin 2010; Lane 2011, 2015; Hill 2015). Mitochondria are vertically inherited, membrane-bound organelles that are

composed of proteins that originate from both the mitochondrial genome and the nuclear genome (the genes encoding from the latter are commonly known as N-mt genes). The compatibility between these two genomes (mitonuclear compatibility) can affect the fitness of an organism by influencing the efficiency of OXPHOS and all downstream processes that rely on its energy output such as immune function, movement, and reproduction (Barrientos et al. 2009b; Lane 2011). Although mitochondria and their DNA are vertically inherited from mothers (Lane 2011), nuclear genomes across populations of sexually reproducing organisms are subject to recombination, and thus sexual reproduction and the novel genetic pairings it produces can yield mitonuclear combinations that advantage offspring fitness by producing more efficient electron transport system complexes (Dowling et al. 2008; Hill 2015). In contrast, although asexual reproduction confers the advantage of preserving energy otherwise spent on mate-seeking and copulation, it almost inevitably dooms organisms to face Mueller's ratchet (Gordo and Charlesworth 2000). Alternately, sexually reproducing species are potentially capable of compensating against the deleterious effects of the ratchet through selecting mates whose nuclear (i.e. variable and recombining) genome possesses genetic material that can reduce or offset the dysfunction induced by mitochondrial mutations. Couplings in sexually reproducing species that maximize the best possible combinations, therefore, should be favored.

In 2013, Hill and Johnson proposed that, because the correct matching of mitochondrial and nuclear genes is so central to the fitness outcomes of sexual reproduction, females should assess potential mates for mitonuclear compatibility, and

therefore preferences should emerge among females for males that detectably signal optimally performing mitochondria. Hill and Johnson suggest that the evolution of ornaments represents exactly that: an indicator of not just phenotypic quality but of good genes, and most specifically, of good mitonuclear genes. Although their hypothesis identifies ornamentation as a specific mechanism which allows males to communicate genetic quality, their central argument is for the importance of signals of good genes even when those signals are much less obvious than ornaments. Although empirical evidence for this proposed phenomenon remains relatively scarce, Hill et al. (2019) found that redness in developing feathers of house finches (*Haemorhous mexicanus*), a condition-dependent trait under strong sexual selection by females, was correlated with the function of mitochondria in finches. On the other hand, many such signals escape easy detection by leveraging less classically obvious traits—for instance, some taxa (including *Callosobruchus maculatus*, a seed beetle and the model of this study) deploy chemosensory traits or stimuli such as pheromones to attract mates. A recent review by Buchinger and Li (2023) found that 23 of 34 known animal phyla demonstrated some indication (and, for 10 of 34, documented evidence) of chemosensory signaling alone. For instance, of the 131 surveyed orders in Arthropoda (the order to which *C. maculatus* belongs), more than half either demonstrated mate choice driven by chemosensory interaction or engaged in chemosensory sexual communication. Although chemical signals have shown strong evidence for candidacy as indicators of phenotypic quality (e.g. energetically trading off with fitness components such as longevity, varying within a population), evaluations of whether chemical signals commonly serve, in practice, as

honest signals of quality (or of mitochondrial efficiency) are empirically undertested (Johansson and Jones 2007).

Irrespective of the form a signal takes, however, whether such signals used in mate choice indicate female preference for males with compatible mitochondrial and nuclear genomes to their own remains almost entirely empirically unexplored (but see Koch and Dowling 2022; Crino et al. 2024). Although mitochondrial haplotypes (i.e. the strictly mitochondrial genotype) and the cohort of N-mt genes that partially encode the mitochondrial subunits that form the electron transport system may not necessarily always be visible through the expression of obvious traits like plumage or size, evidence of female preference for compatible males (either in terms of their mitochondrial haplotypes, nuclear genes, or both) may be clearer when females are presented with many options for potential suitors of different N-mt compositions. Moreover, evidence of mate selection favoring compatible mitochondrial or nuclear genomes could provide evidence that such signals exist even if they remain difficult to detect.

We hypothesized that mate choice is driven by the benefits of an individual procreating with a mate whose mitochondrial and N-mt OXPHOS alleles are compatible with its own. We evaluate this hypothesis using *Callosobruchus maculatus*, a seed beetle, to serve as a model organism. I describe *C. maculatus* in detail above (see Study System within Introduction), but in brief, these beetles are easy to rear, polygynandrous, and have been shown to evaluate mate quality and adjust copulation in multiple ways (e.g. males have been shown to prefer virgin mates and females actively truncate matings to avoid excessive genital damage but may kick males off early to maximize fitness by remating with

other individuals (Qi and Burkholder 1982; Crudgington and Siva-Jothy 2000; Nojima et al. 2007; Van Lieshout et al. 2014; Rousseau 2021). To test our hypothesis using *C. maculatus*, we experimentally separated naturally occurring mitochondrial and nuclear genomes by introgressing mitochondrial haplotypes onto nuclear backgrounds from geographically distinct populations more likely to have incompatible N-mt alleles. We then made four predictions: we predicted that (1) *C. maculatus* would, on average, prefer a mate whose nuclear genome is derived from the same population of origin as that of the chooser; (2) that *C. maculatus* would, on average, prefer a mate whose mitochondrial genome is derived from the same population of origin as that of the chooser; (3) that *C. maculatus* would, on average, prefer a mate whose mitochondrial and nuclear genomes are both derived from a single population of origin irrespective of the genome origins of the chooser, and (4) that *C. maculatus* would, on average, prefer a mate whose mitochondrial and nuclear genomes are both derived from the same respective populations of origin as those of the chooser (encompassing predictions one and two).

Methods

Traits Measured

We evaluated a suite of traits to quantify receptiveness of an individual to its mate. We designated traits as either behaviors that express male preference, behaviors that express female preference, or behaviors that express both. Every trait that occurred during the mating assays was measured as a duration that elapsed between two events (defined below in Table 10).

Table 10: Events during mating and how they appear.

Event	Description
Male is deposited into dish	When the male is added into the dish manually
Male begins running	The first moment at which the male departed from its initial position where it was deposited
Male begins chasing female	Usually occurs when the male first comes into contact with the female, the male begins pursuing the female actively
Male antennation begins	Males begin rapidly moving their antennae, usually over the female's abdomen
Male antennation ends	Male antennae stop moving almost entirely or entirely
Female flees male	Occurs when and if the female runs from the pursuing male
Female accepts male	When the female stops running (copulation not required to satisfy this condition)
Copulation begins	The point at which the male inserts the aedeagus into the female, typically presents as the male "leaning back" once mounted
First kick	After copulation begins, the first instance during which the female kicks the male
Copulation ends	Full separation between the two individuals

Table 11: Traits measured and their descriptions.

Trait	Latency to Chase	Chase Duration	Antennation Duration	Copulation Duration	Copulation to First Kick	Acceptance to Kick	Kick to End
Description	Male is deposited -> Male begins running	Male begins chasing female -> Female accepts male	Antennation begins -> Antennation ends	Copulation begins -> Copulation ends	Copulation begins -> First kick	Female accepts male -> First kick	First kick -> Copulation duration
Notes on Preference	Shorter latency indicates preference	Shorter chase indicates preference	Shorter antennation duration indicates preference	Longer copulation duration indicates preference	Longer copulation to first kick indicates preference	Longer acceptance to kick indicates preference	Longer kick to end indicates preference

Mating Assays

We haphazardly sampled beans from lines within each population. Each bean had at least one hatched egg laid upon it, and we kept the beans in microcentrifuge tubes to allow adult beetles to emerge singly and guarantee no unintentional matings would occur. We then paired single male and female individuals randomly from separate lines within each population (to guarantee no inbreeding) to produce offspring to be used in matings. The beans upon which their eggs were laid were again separated into microcentrifuge tubes to be reared singly. Upon a beetle's adult emergence (i.e. eclosure from a bean), we recorded the date and sexed the beetle.

We began blocks of mating assays the day that most families had produced at least two males and females each, and we selected only individuals who were two or fewer days old at the time of mating. Mating assays took place at one of two stations, or locations at which a camera and the corresponding petri dishes to be filmed were fixed, and were recorded with the main camera on either an iPad 9 Pro (station 1) or a Samsung Galaxy 9

tablet (station 2) anchored horizontally approximately 13” above the beetles’ petri dishes. Each assay within a station captured up to eighteen concurrent mating events which took place within a single, otherwise empty, 35mm petri dish. We organized eighteen dishes per round—the maximum number visible through the tablet’s aperture when reviewing footage—under the camera in a specific arrangement.

Within each assay, we chose to mate no more than one female of a given family at a time to statistically distribute families across as many rounds and stations as possible. We randomly assigned females a position between one and eighteen within the station assay and weighed each female (as part of the nuptial gift trait measurement, see Traits Measured for more details) before we placed them into the petri dishes, then giving them a five-minute acclimation period over which they were left alone (but filmed). After the acclimation period, we weighed and introduced each male to its respective female and allowed the tablet to record the pair for 45 minutes. We weighed both males and females after filming ended.

Experimental Design

We produced sixteen unique families evenly drawn from four populations: SI (southern India), BF (Burkina Faso), Candy, and Spice (see Cybrid Line Development in Introduction). From these sixteen families, we randomly selected approximately four female siblings, one female at a time, to serve as one member of a single, mated pair. We paired each female from the quartet of sisters to a male from one of the four population types (see Table 12 below). We did this to capture the average within-family preference (i.e. any preference associated with the individuals' parents' alleles). We ensured also that we evenly represented male families across all four population types of females such that we similarly captured the within-family preference of males. All populations were reared under the same conditions (see Line Cultivation and Population Rearing for more details).

Table 12: The Four Populations Used Experimentally. "N" represents the population of origin for the nuclear alleles present, and "mt" represents the population of origin for the mitochondrial haplotypes present.

	N: Burkina Faso (BF)	N: Southern Indian (SI)
Mt: Burkina Faso (BF)	BF	Candy
Mt: Southern Indian (SI)	Spice	SI

We then sampled individuals from each population and mated them to members of an unrelated line within the same population to produce offspring for mating assays. We filmed all mating assays and later visually scored the receptivity of both males and females to their respective mates (see Mating Assays for more details). We evaluated these traits using linear mixed models in R (version 4.2).

Statistical Methods

All traits were scored by two individuals reviewing video recordings and evaluated as response variables in linear mixed models in R (version 4.2, R Core Team 2025) using the glmmTMB package (Brooks et al. 2017). We included a random effect of measurer to account for any differences in measurement due to differences among observers. We visualized estimated marginal means using the emmeans (Lenth et al. 2025) package in R. We analyzed all traits separately in univariate models and modified only the relevant fixed effect to test each prediction. We analyzed mating preference data using model structures unique to each trait that we tested within the context of the four predictions described in the introduction (and recapitulated below). Each trait was analyzed across four models that corresponded to the four predictions mentioned in the introduction and included one of four unique prediction-related fixed effects: nuclear type (matched or mismatched), mitochondrial type (matched or mismatched), same overall type (same overall mitochondrial and nuclear type or different overall mitochondrial and nuclear type type), or whether the mate matched to itself (whether the mate's mitochondrial and nuclear genomes originated from the same population of origin).

Male Trait: Latency to Chase Models

For all latency to chase models, given the high frequency of zero inflation but otherwise relatively normal distribution of values, we used a tweedie distribution. The tweedie distribution uses a log link function and treats non-zero values as normally distributed but also evaluates the log odds of producing zeros given a fixed parameter on a logit scale. We ran our models in glmmTMB with fixed effects of the prediction trait measured (i.e. either a

mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not the mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male age at mating (in case females showed preference for slightly older or younger males, on average). Given our repeated measures via multiple observers, we included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and to account for any influences of time of day or administration of the assays, the round in which they were measured (out of ten possible rounds).

Male Trait: Antennation Duration Models

For all antennation duration models, we used a gamma distribution with a log link function in glmmTMB with fixed effects of the prediction trait measured (i.e. either a mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not the mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male's age at mating. We included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and the round in which they were measured (out of ten possible rounds). We added a constant of 0.01 seconds to each response value of antennation duration in to avoid zero inflation (since antennation is the means by which males detect females, it must occur but may sometimes last an almost imperceptibly short time).

Female Trait: Chase Duration Models

For all chase duration models, given the high frequency of zero inflation but otherwise relatively normal distribution of values, we used a tweedie distribution with a log link function in glmmTMB with fixed effects of the prediction trait measured (i.e. either a mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not a mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male's age at mating (in case male age influenced female responses). We included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and the round in which they were measured (out of ten possible rounds).

Female Trait: Copulation-Until-Kick Models

For all copulation-until-kick models, we used a normal distribution in glmmTMB with fixed effects of the prediction trait measured (i.e. either a mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not the mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male age at mating (to account for the possibility that males' ability to pursue a female may be affected by it). We included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and the round in which they were measured (out of ten possible rounds).

Female Trait: Acceptance-Until-Kick Models

For all acceptance-until-kick models, we used a normal distribution in glmmTMB with fixed effects of the prediction trait measured (i.e. either a mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not the mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male age at mating (to account for the possibility that males' ability to pursue a female may be affected by it). We included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and the round in which they were measured (out of ten possible rounds).

Female Trait: Kick-Until-End Models

For all kick-until-end models, we used a gamma distribution with a log link function in glmmTMB with fixed effects of the prediction trait measured (i.e. either a mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not the mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male age at mating (to account for the possibility that males' ability to pursue a female may be affected by it). We included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and the round in which they were measured (out of ten possible rounds).

Female Trait: Copulation Duration Models

For all copulation duration models, we used a normal distribution in glmmTMB with fixed effects of the prediction trait measured (i.e. either a mitochondrial match, a nuclear match, an overall combination of the mate, or whether or not the mate's mitochondrial and nuclear populations of origin matched those of the individual measured), the station at which the individual was measured, and the male age at mating (to account for the possibility that males' ability to pursue a female may be affected by it). We included random effects of the individual's mating ID (i.e. the intersection of round, the position of an individual's dish relative to the other dishes, and station) and the round in which they were measured (out of ten possible rounds).

Results

We investigated the relationship between male and female preferences for mates and nuclear type (prediction 1), mitochondrial type (prediction 2), mitonuclear match of the mate (prediction 3), and genomic familiarity (by population of origin) of the mate to the chooser (prediction 4). We found that males were significantly more likely to instantly begin chasing females who had a mismatched nuclear and mitochondrial population of origin (Table 17, $\beta = 2.244$, SE = 0.943, $p = 0.017$) or when the male's nuclear type was mismatched to the female's nuclear type (Table 13, $\beta = 2.189$, SE = 0.90, $p = 0.015$). Males were significantly less likely to immediately begin chasing females when their mitochondrial types were mismatched to their mate's (Table 14, $\beta = -1.970$, SE = 0.911, $p = 0.031$). Among those instances in which males delayed before chasing females, the delay was significantly shorter when females' mitochondrial and nuclear genomes were not both from the same population of origin as those of the male (Table 18, $\beta = 2.24$, SE = 0.94, $p = 0.017$). The time elapsed between female acceptance of a male (i.e. when she ceases running from the male) and the first time a female kicks a male was significantly shorter when the female and male mitochondrial types were mismatched (Table 15, $\beta = -60.96$, SE = 25.08, $p = 0.015$). The copulation duration was shorter ($p=0.06$) when the mitochondrial types were mismatched between the pair (Table 16, $\beta = -41.8$, SE = 22.30). We found no other significant differences (i.e. $p>0.06$) in preferences of either males or females driven by nuclear type, mitochondrial type, overall quality of mate, or the shared genomic origins of the mitochondrial and nuclear genomes.

Table 13: Latency to chase by matched or mismatched nuclear type (log scale).

Latency to Chase	Seconds	Standard Error	Z-Value	P
Nuclear Matched (Intercept)	-0.0809096	0.3646419	-0.2218878	0.8244012
Nuclear Mismatched	0.4953303	0.4359563	1.1361925	0.255876
Survey Station (Two)	0.1004778	0.4371365	0.2298545	0.8182048
Male Age at Mating (Scaled)	0.2303445	0.2162456	1.0651987	0.286786
Zero Inflation (Latency to Chase)				
Nuclear Matched (Intercept)	-4.6961366	0.6655036	-7.0565162	1.707288e-12***
Nuclear Mismatched	2.1892879	0.8978474	2.4383742	0.01475349**

Survey station represents the table upon which assays were conducted, male age at mating represents the age (in days) of the male at the time of mating, and zero inflation describes the log odds of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Table 14: Latency to chase by matched or mismatched mitochondrial type (log scale).

Latency to Chase	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	-0.2046057	0.3994922	-0.5121645	0.6085359
Mitochondria Mismatched	0.6370269	0.4387622	1.4518726	0.146537
Survey Station (Two)	0.1707422	0.4358051	0.3917856	0.6952166
Male Age at Mating (Scaled)	0.1749235	0.2178288	0.8030321	0.4219562
Zero Inflation (Latency to Chase)				
Mitochondria Matched (Intercept)	-2.6574091	0.6014736	-4.4181642	9.95428e-06***
Mitochondria Mismatched	-1.9704998	0.9109805	-2.1630537	0.03053705**

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and zero inflation describes the log odds of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Table 15: Acceptance until kick time by matched or mismatched mitochondrial type.

Acceptance to Kick	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	406.409911	30.83625	13.1796152	1.149758e-39***
Mitochondria Mismatched	-60.961500	25.08536	-2.4301628	0.01509204**
Male Mass (Scaled)	-38.366725	12.56677	-3.0530301	0.002265432**
Survey Station (Two)	-9.222985	41.95230	-0.2198445	0.8259922
Male Age at Mating (Scaled)	-28.420481	12.67210	-2.2427604	0.02491227**

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Table 16: Time elapsed from the moment the male inserts the aedeagus until full separation by matched or mismatched mitochondrial type.

Copulation Duration	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	510.520184	18.76335	27.2083722	5.170394e-163***
Mitochondria Mismatched	-41.799965	22.30320	-1.8741686	0.06090719*
Male Mass (Scaled)	-22.714059	11.16734	-2.0339727	0.04195435**
Survey Station (Two)	-12.217410	22.32716	-0.5471995	0.5842417
Male Age at Mating (Scaled)	-4.707593	11.18128	-0.4210245	0.6737372

. 'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Table 17: Latency to chase by mate quality (log scale).

Latency to Chase	Seconds	Standard Error	Z-Value	P
Mitonuclear Matched Mate (Intercept)	0.0118174	0.3671595	0.0321860	0.9743237
Mitonuclear Mismatched Mate	0.2679165	0.4402889	0.6085015	0.5428549
Survey Station (Two)	0.1073781	0.4413154	0.2433138	0.8077623
Male Age at Mating (Scaled)	0.2233974	0.2171814	1.0286212	0.3036577
Zero Inflation (Latency to Chase)				
Mitonuclear Matched Mate (Intercept)	-4.9414677	0.6899165	-7.1624140	7.926859e-13***
Mitonuclear Mismatched Mate	2.2436743	0.9430388	2.3791962	0.01735044**

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and zero inflation describes the log odds of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Table 18: Latency to chase given population-level match or mismatch of mitochondrial and nuclear genomes (log scale).

Latency to Chase	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	0.4452711	0.3372357	1.320355996	0.1867162
Different Overall Type	-1.1130186	0.4968438	-2.240178123	0.02507936**
Survey Station (Two)	0.1028193	0.4309020	0.238614176	0.8114048
Male Age at Mating (Scaled)	0.2064281	0.2145071	0.962336807	0.3358804
Zero Inflation (Latency to Chase)				
Same Overall Type (Intercept)	-2.8344341	0.5172709	-5.479593577	4.263039e-08***
Different Overall Type	-16.5251845	3447.2528490	-0.004793726	0.9961752

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and zero inflation describes the likelihood of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Discussion

We tested whether the mitochondrial and nuclear backgrounds of male and female seed beetles influenced their preferences for mates by assaying a suite of behaviors during monogamous mating trials. We hypothesized that mate choice is driven by the benefits of an individual procreating with a mate whose mitochondrial OXPHOS alleles are compatible with its own. We found support for the prediction that choosers would prefer mates whose mitochondrial types originated from the same population as their own. In both males and females, individuals expressed a preference for mates who possessed a mitochondrial type associated with their ancestral population of origin (i.e. either southern India or Burkina Faso), as measured through the latency to begin a mating chase, the duration between a female's acceptance of a male until she first kicked that male, and the full duration of copulation. Contrary to our predictions, we also found that males expressed a preference, on average, both for females whose nuclear genome was derived from an ancestral population different from that of her mitochondrial genome and for females whose nuclear genome was derived from an ancestral population that differed from that of the male's. In other words, males seemed to favor females whose nuclear genomes mismatched either the male's nuclear genome or the female's own mitochondrial genome.

The differences in mating preferences between males and females that we observed may result, in part, from the well-described intralocus sexual conflict present in *C. maculatus*, which constrains males and females, preventing them from advancing toward their respective sex-specific trait optima (Berger et al. 2014b). If mate preference, for instance, is under autosomal control and not subject to sex-specific expression,

individuals may inherit a parent's preference even when that preference might result in poorer fitness in offspring. Such sexual conflict could be further exaggerated or resolved based on the N-mt genes under selection and their expression. Given that selection for mitochondrial function only operates on females, intralocus sexual conflict could become more extreme if positive selection on autosomal N-mt genes favoring females disadvantages males. Alternately, the rewards for males who display sex-specific expression of N-mt genes could be significant—for example, Eslamieh et al. (2017) found evidence of testis-specific expression of N-mt genes in *Drosophila* that suggests that mitonuclear sexual conflict may be frequently resolved through gene duplication and relocation. However, Havird et al. (2019) found evidence that testis-specific gene duplication may be under relaxed, rather than strong selection, suggesting that sexual antagonism could be maintained rather than resolved.

Sexual selection theory has traditionally argued that detectable differences in male phenotypes, especially ornaments, are indicators of some measure of quality (Zahavi 1975; Grafen 1990), such as specific "good genes" (Hamilton and Zuk 1982) or the ability to produce costly phenotypes while remaining competitive in other ways (e.g. somatic maintenance, disease resistance, etc.). Although these ideas are widely accepted, the identities of the 'good genes' or sources of quality have been the basis for some contention (Achorn and Rosenthal 2020). Hill and colleagues (Hill and Johnson 2013; Hill 2018) have made the case that the value of mitonuclear mate choice lies in its promotion of prezygotic sorting, which facilitates compatible OXPHOS complexes and yields higher fitness offspring than might be expected by other potential drivers of mate preference. Here,

because our focus rests on the genetic background of the individuals rather than phenotypic characteristics, we lend support to Hill's argument that signals may be an indication of compatibility if not necessarily quality.

Our study focused exclusively upon the mate preferences from two ancestral populations of *C. maculatus* derived from southern India and west Africa (Burkina Faso) that represent some of the oldest mitochondrial lineages (Kebe et al. 2017; Zhang et al. 2018; Lu et al. 2024). The preferences of these populations may not be representative of other populations (or species). Moreover, the two wild populations from whom our stock cultures descend are separated by continents, and the cross-population mitochondrial differences may be greater than might be expected within a single, naturally occurring population. Although previous work has shown the mitochondrial haplotypes of these two populations to differ substantially (Kebe et al. 2017; Sayadi et al. 2017), we produced maternal, rather than strictly mitochondrial, lineages, and these preferences could be artefacts of X-linked effects specific to our given populations. Our study also highlights findings in a polygynandrous species which may show weaker preferences than those whose preferences could be more obvious (such as monogamous species). Future work should explore whether these mate preferences are commonly present in other taxa, test whether preferences persist across repeated mating events, and especially whether within-population mitochondrial diversity produces similar preferences (given that our mitochondrial haplotypes originate from populations on two different continents).

Chapter 3: What's in a "Game"? The Educational Efficacy of a Game Designed to Accurately Represent Evolution by Natural Selection to Undergraduates

Introduction

The Importance and Challenges of Teaching Evolution

Theodosius Dobzhansky famously declared that "nothing in biology makes sense except in the light of evolution" (Dobzhansky 1973). Evolution is essential in Earth's spectacular biodiversity and the fascinating processes that have produced the tree of life, and is inherent to understanding the adaptive potential of wildlife facing the threat of a rapidly changing world (Martin et al. 2023). Biology students, many of whom are pre-health professionals, benefit from learning how the most substantial challenges facing *Homo sapiens* in the modern day include the evolution of pathogens, such as the SARS-CoV-2 virus that causes COVID-19. In fact, the first core concept listed in Vision and Change: A Call to Action (Woodin et al. 2010), which formalized education priorities in U.S. biology coursework, is "Evolution—the diversity of life evolved over time by processes of mutation, selection, and genetic change."

Although the importance of evolution is rarely disputed in the biology discipline, evolution, including evolution by natural selection, is challenging to teach effectively in secondary and postsecondary academic contexts (Bishop and Anderson 1990; Abenes and Caballes 2020; Plutzer et al. 2020). Students confuse evolution and natural selection and struggle to accurately explain or identify these concepts whether on exams or in open-

ended responses (Baumgartner and Duncan 2009; White et al. 2013; Vostinar et al. 2020). Misconceptions about evolution have not only been associated with religious or contravening ideologies (Buckberry and Burke da Silva 2012) but also closely linked to popular misrepresentations of the subject through life experiences and the popular media (Abenes and Caballes 2020; Archila et al. 2024). For example, popular media may misrepresent evolution in references to it up to 96% of the time (Ferguson et al. 2022). One venue of exposure includes well-known, commercially popular video games (Kalinowski et al. 2016; Leith et al. 2016) which deploy the word 'evolution' frequently but (usually) inaccurately. Unfortunately for instructors, this suggests that the already challenging task of teaching evolution accurately may not be sufficient to ensure student comprehension. Because students enter the classroom with inaccurate ideas about the topic, instructors have to address and correct misconceptions common among students and teach the accurate content (though instructors may also be misinformed, Abenes and Caballes 2020). Taken together, long-standing challenges to understanding evolution persist (Bishop and Anderson 1990; Anderson et al. 2002; Baumgartner and Duncan 2009; Wood 2009; Abenes and Caballes 2020; Aini et al. 2024).

Evolution acceptance, defined as "agreement that evolution is valid and the best explanation from science for the unity and diversity of life on Earth, which includes speciation, the common ancestry of life and that humans evolved from non-human ancestors" (Barnes et al. 2024), also represents a barrier to learning, especially among some religious students (and especially to students of color who tend to be more religious, Barnes et al. 2020b,a). Although recent work has shown promising results when

attempting to close the acceptance gap, such as the deliberate integration of consensus messaging (Korfmacher et al. 2024), the more successful of the tested solutions sometimes involve more delicate approaches in stressing the importance of evolution in order to promote its acceptance among students for whom evolution feels in conflict with their religious views (Aini et al. 2024).

Video Games as Active Learning

Several evidence-based approaches have been established to teach evolution, including discussing misconceptions about evolution (Nelson 2008), developing cultural competency around religious beliefs when teaching (Barnes and Brownell 2017), promoting compatibility between religion and evolution (Barnes et al. 2017; Truong et al. 2018), presenting multiple perspectives on the subject (to promote acceptance), exploring the geological record (Nelson 2008), and engaging activities such as group work and other forms of active learning. Active learning received a formalized definition from Driessen et al., (2020), and a non-exhaustive list of example activities includes metacognition, discussion, group work, formative assessment, practicing core competencies, live-action visuals, conceptual class design, worksheets, and games). In the present research, we explore the potential of video games to teach evolution (Mayo 2009; Yang 2015; Ball et al. 2020). Recreational video games are so globally ubiquitous that they out-earn the film industry annually, though they have only recently been explored as a potential teaching tool in higher education (Annetta 2008). The fact that adolescents have likely encountered them prior to entering higher education (approximately 99% of boys and 94% of girls in the United States between the ages of 12 and 17 reportedly play video games) underscores their accessibility and

students' familiarity (Lenhart 2008). Lenhart (2008) demonstrated a relationship between gameplay and engagement in non-academic activities among players: playing video games with others in a shared space predicted engagement in numerous real-life activities, such as volunteering, raising money for charity, and researching politics (Lenhart 2008).

While the use of video games in academic instruction is underexplored, previous work in the context of cognition studies (Wais et al. 2021; Franceschini et al. 2022) suggests the activity can lead to boosts in short-term memory as well as the general acquisition and retention of declarative knowledge (i.e. stated information) when compared with traditional modes of instruction such as lecturing (Riopel et al. 2020). Several studies have shown that video games, especially 'serious games' (i.e. those designed for a purpose separate from pure entertainment or commercial success) may also be effective in the context of STEM disciplines (Arztmann et al. 2022). A meta-analysis evaluating the efficacy of serious games across seventy-nine studies found that serious games can significantly influence motivation, conceptual understanding, and declarative knowledge (Riopel et al. 2020), especially in natural sciences such as biology. Although video games were historically played primarily by (young) men, recent studies have also found playing video games or owning a video game console has become increasingly popular among girls and women and is associated with improvements in their academic performance (Lucas and Sherry 2004; Joiner et al. 2011; Lavalle et al. 2025) and a reduction in the effects of stereotype threat (Fordham et al. 2020). These results held even after correcting for social and economic factors. Additionally, while instructors who engage their students in active learning have reported challenges with student "buy-in" (Cavanagh et al.

2016), previous research has demonstrated that games remain effective at sustaining student motivation even when used in large (>50 students) classroom setting (Annetta 2008; Walker et al. 2008; Mayo 2009; Catton et al. 2016).

The Unexplored Potential of Video Games

Evolution is frequently referenced in popular video games and in-class simulations. Simulations have been routinely employed in general biology, genetics, and evolution classrooms for decades (Soderberg and Price 2003; Clarke-Midura et al. 2018), producing meaningful strides in academic performance when compared to traditional lecture. According to Narayanasamy et al. (2006), simulations describe state changes to some value with some relationship to time (e.g. at a single point in time, over continuous time, or without respect to time) given a ruleset. In contrast, Salen and Zimmerman define a game as "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (2003, p. 80). In other words, player agency represents the key difference between simulations and video games.

However, video games rarely represent concepts related to evolution accurately (Herrero et al. 2013; Leith et al. 2016). Despite this, engaging with video games that include an evolutionary component—even inaccurately represented—has been shown to correlate positively with the number of hours that a student spends studying evolution in a given week and has been linked to an average increase in grade and test scores in evolution coursework (Poli et al. 2012). Although video games and simulations—similar to but distinct from video games—have been linked to improved academic performance in general and improvements in student knowledge of evolution by natural selection as assessed on

quizzes and exams, and have even been shown to influence student behaviors in real life, the question of whether evolution by natural selection can be effectively and accurately taught via video games remains a major gap in the literature. Our study aims to address this gap.

Our Aim

We asked several research questions, namely, (1) What is the effect of playing a novel video game on undergraduate students' attitudes towards (a) science and (b) the video game itself?, (2) What is the effect of playing a novel video game on undergraduate biology students' self-efficacy in conceptual understanding of evolution and their actual understanding of evolution?. We then experimentally tested the effectiveness of this game as a teaching tool for use in the instruction of evolution for undergraduate students, and finally addressed the question (3) What is the impact of this novel video game on a concept inventory assessment (the Conceptual Assessment of Natural Selection)? We hypothesize that undergraduate students are often unable to fully grasp evolution by natural selection because traditional lessons teaching it can effectively communicate the results of evolution but often struggle to represent the processes shaping it. We predict that students who play a video game that allows them to directly interact with the process of evolution by natural selection in real-time for half an hour will score significantly higher on an evolution concept inventory assessment (the Conceptual Assessment of Natural Selection) than students who do not play the game for the same half-hour.

Methods

Student Learning Objectives (SLO)

We centered our video game development upon three student learning objectives (SLO) synthesized by combining principles from previous work (Lewontin 1970; Lande and Arnold 1983; Futuyma 1998), namely:

(SLO 1) …natural selection occurs within a single generation,

(SLO 2) …natural selection acts upon phenotypic variation,

(SLO 3) …evolution by natural selection occurs when heritable phenotypes cause differences in fitness.

Student Learning Objectives in Gameplay

(SLO 1) Natural Selection (single generation) and (SLO 2) selection acts upon phenotypic variation.

Students should recognize that natural selection occurs within a single generation, acting on phenotypes and not genotypes. To address the first learning objective, a key aim was for students to recognize that natural selection—distinct from evolution by natural selection—occurs not over several generations but within each (single) generation. We ensured that player choices have visible effects on the existing population of playable characters before they produce offspring.

To achieve the second learning objective, we leverage the phenotypic variation within a population of warlaks, a fictional amalgam of a diamondback terrapin and a seed

beetle. During gameplay, warlaks are capable of flying, leaping, or recoiling into a shell. All warlaks within a population share one of these player-selected traits and vary in the expressed magnitude of that trait (e.g. the distance of a given leap) and vary phenotypically in size. Students select a population size ranging from five to one-hundred individuals, each sharing the same traits (e.g. size) but differing in phenotypic values for those traits, struggling to reach a nesting habitat while predators and natural barriers obstruct their paths. These differing phenotypes frequently result in different capacities for escape and level-completion among individual warlaks, and the within-generation consequences of selection are visible to students each time a warlak is snatched up in the claws of a predator or fails to reach the breeding nest at the end of the level.

SLO 3: Evolution by Natural Selection is linked to heritable differences in fitness

Students should recognize that, when traits are heritable, trait values which promote survival and reproduction are likely to appear more frequently each generation than traits which do not promote (or which endanger) survival or reproduction. After progressing through the level, any individuals which succeed in reaching the nest (i.e. the 'end' of the level) are admitted into a breeding pool that randomly assigns individuals as mates to one another. If a single member survives, they are randomly mated with an unrelated member of the species with the same phenotype. Any organisms who are unable to reach the nesting ground or are gruesomely devoured by predators do not contribute genetically to the offspring which form the next generation. In this way, the decisions a player makes in a single generation (i.e. the current level) will directly affect the playable population in the next generation (i.e. the next level).

To demonstrate how heritability can vary across populations, we gave players the ability to select from among three discrete heritability levels at the start of each 'run', where heritability, here, represents narrow-sense heritability, or the percent of total phenotypic variation attributable to additive genetic variation (Falconer 1981). Players can adjust trait heritability before gameplay and observe responses to selection at high (0.9, or 90%), medium (0.5, or 50%), or low (0.1, or 10%) levels. For example, if a player sets trait heritability to 'high' and chooses to play as a population of flying warlaks, the mean flight strength (i.e. lift generated by a single flap of wings/press of the spacebar) of surviving individuals in one generation will closely resemble the average flight strength of their offspring in the next generation/level—there will be a correlation of approximately 0.9 (90%) between the parents' and offsprings' values for that trait. Alternately, if a player selects a 'low' heritability and once again chooses the 'flight' trait, responses to selection will likely be small each generation, and players may notice that the average flight strength of parents will have a low correlation (~0.1) with the average flight strength of their offspring. In both examples, the (natural) selection applied to each population could be identical and is observable within the generation in which it occurs; it is the player's choice of heritability, which offers us information about the influence of the genetics underlying the trait in the population and not only the parents' phenotypes, which determines the response to selection. We reinforce the different magnitudes of response to selection, via differences in heritability, through visual cues: trait heritability remains on-screen during gameplay, and changes in average trait values across all played generations are visible once a population goes extinct or a player quits.

When players select a trait and choose a 'high' heritability value, they may discover that certain trait values become more common within just a few short generations. For instance, if players choose to play as a population with the 'recoil' trait (which resembles a box turtle withdrawing fully into its shell), the heritable trait in focus will be each warlak's latency to retreat into its shell. The individuals which are most likely to survive could be those who recoil quickly rather than those who recoil slowly, as slow-to-hide warlaks may be more likely to be captured by predators. At a high heritability, within a few generations (levels), the population of warlaks will likely see a sharp decrease in latency to retreat.

Tool Development

We developed a video game with the intention of accurately representing evolution by natural selection in gameplay and testing changes in undergraduate understanding of concepts related to evolution. We built the game with the Unity Game Engine, a free software tool available for use in both educational and commercial production of many popular video games (e.g. Cuphead, Pokemon Go, and Hearthstone).

The game we developed is a two-dimensional (2D) platformer, a style common among classic console titles (e.g. Super Mario World, Donkey Kong Country, the Megaman series, etc.). Platformer (2D) gameplay generally allows the user to control a character that moves in two dimensions (i.e., left, right, up, or down), is subject to the rules of gravity, and proceeds from the 'beginning' (usually the leftmost boundary) to the 'end' (usually the rightmost boundary) of a level. We selected this style of gameplay with the aim of avoiding confusion about how the game should be played: platformer-style games have existed for

decades and are still produced in 2025 and completion of the level is usually achieved by moving from left to the right, demanding little of players in terms of navigational skills.

We published the game in WebGL and hosted it on simmer.io, a free WebGL hosting domain, which allowed the game to be played within most web browser environments such as Chrome, Safari, Firefox, and Microsoft Edge. This game was available to students using Mac OS, Windows, or Linux operating systems, though we did not ask for or require students to disclose any details about their personal laptops or operating system of choice. We conceived, drew, and animated all the visual assets used in this game with the sole exception of the warlak in the opening sequence, which was illustrated by Leah Tennery.

Gameplay

We developed a fictional character, the noble "warlak", to serve as the species-as-protagonist in the game. Centering instruction around fictional characters may reduce bias based on pre-existing knowledge of real organisms and encourages critical examination of how phenotypes behave in particular environments rather than drawing on students' reflexive assumptions about organisms they know (Beason-Abmayr 2023). At the start of a run, players select a population size of warlaks ranging from five to one hundred individuals and then control every member of that population at once. For instance, pressing the right arrow key moves every warlak to the right. Players must then traverse a harrowing and sometimes unforgiving environment (depending on the player-selected difficulty) replete with predators and obstacles barring every warlak's ability to proceed to the breeding site

at the end of each level. The game ends either with an extinction event (i.e. all warlaks die) or when they reach an adaptation threshold that varies based on difficulty—for instance, on easy mode, if 80% of warlaks survive, the player wins.

Phenotypically Variable Traits

At the start of a new run, each warlak is phenotypically unique in two ways. First, they differ in size, which has implications for warlak survival because many predators have prey size thresholds over or under, which will impact their behavior toward warlaks (i.e. either attack or ignore). Warlaks also differ in phenotypic values for one of three possible special abilities, or traits, which players may choose at the start of a run for all members of a population: either 'flight', in which a press of the spacebar represents a 'flap' of a warlak's wings, 'leap', in which a press of the spacebar causes a warlak to leap forward a certain distance (i.e. like a bullfrog), or 'recoil', in which a press of the 'down' arrow causes the warlak to recoil at a certain speed into an invulnerable state (i.e. in a manner similar to a box turtle). Players have the option to advance to the next level and generation the moment a single warlak reaches the breeding site, but they may continue to attempt to advance more members of their population into the breeding pool. By giving players the ability to decide when the selective 'cut-off' occurs, even if certain individuals might potentially have reached the breeding pool, we offer a proxy for sexual selection as individuals in real, wild populations who may survive long enough to reproduce may not necessarily get the opportunity to do so. The level ends either when the player chooses to advance to the next level or when all playable warlaks have either reached the breeding site (or it ends when they have all been devoured).

At the end of each level, a small graphic appears to inform the player of the average change in trait values (e.g. leap strength) between the starting population and the survivors. The graphic also reveals the new, expected average trait values, given the player-selected heritability, for the new and upcoming generation. Each subsequent level represents the genetic offspring of the warlaks played in the level before, which means that players observe population-level evolution of warlaks in each playthrough.

Simulated Genetics

Each warlak within a population possesses two traits controlled by genetic architecture: size and special ability (described above). We set up control of these traits by using a simulation which we developed that implemented an individual locus model (Roff 2009, p. 233). Each warlak possesses a diploid genome, each trait is controlled by ten independent loci, and at each locus two alleles are possible at the outset of a new game. Each pair of alleles generated at a given locus has starting genetic values of -1 and +1 respectively, and we randomly assign alleles at each possible locus to founder warlaks. We produce trait phenotypes for founders by implementing the following: (1) each trait begins as the assigned population mean for that trait; (2) we then randomly assign two alleles to each locus (through the process we described above) and sum the genetic values of those alleles across all loci to produce a complete genetic value for the trait controlled by those loci; (3) we divide that genetic value by the number of loci controlling that trait to standardize the trait (ten); (4) to each warlak, we add an environmental deviation, sampled from a normal distribution of values centered around zero. The variance of this distribution equals the estimated phenotypic variance minus the additive genetic variance for that trait.

At the end of each level, each surviving warlak is paired randomly with another surviving warlak (with replacement, where a given warlak's partner may have already mated with another warlak, and all warlaks are treated as hermaphroditic), and each pairing of warlaks produces a number of offspring equal to the population size (e.g. 100) divided by the number of surviving warlaks (e.g. 100 divided by 30 surviving warlaks). When the quotient of that equation does not produce a whole number, one pair produces a single additional random offspring. In the exceptional circumstance that only a single warlak survives, a warlak can mate with itself (or simulate undergoing recombination with its own gametes). This process repeats each generation until extinction or until the warlaks reach the adaptation threshold.

The 'Simple' Game

Plainly, the simple game is an additional control treatment. To test the effectiveness of simply playing any video game on improving outcomes on student assessments of evolution knowledge, we developed a version of the game meant to present a classical gaming experience without the accurate representation of evolution by natural selection that Warlak offers. To do this, we removed a player's ability to choose a population size and automatically set the population size to a single individual. Players may still choose from among the same three traits made available in the full game, and they may also choose their preferred environment. Unlike the full game, however, players can be struck repeatedly by enemies in the simple game, and we gave the warlak a health bar that falls each time a predator harms the character. If the health bar reaches zero, the game ends. If the character survives to the end of the level, rather than entering a breeding pool, the

game simply proceeds to another level without displaying a response to selection, reminding players of the heritability, or any of the 'level-cleared' canvas elements present in the full game. This process repeats until the player eventually dies, at which point the game ends without presenting a phylogeny.

Data Collection

We collected survey data to evaluate any differences before and after gameplay in student knowledge of evolution, attitudes toward science, and confidence in answering questions related to evolution. Each student completed a survey before gameplay (a "pre-game" survey) which included questions which: (1) assessed knowledge of evolution by natural selection, (2) and evaluated attitudes toward science (3) collected demographic data (though we did not include demographic descriptors in the analyses). The demographic categories included were race, ethnicity[1], gender[2], age, year of college, major (write-in), whether that major was a STEM major or not, whether the student was an international student or not, the political affiliation of the student[3], and the student's religion[4]. All submissions of demographic data by students were optional. Students completed another survey after gameplay (a "post-game" survey) which also assessed knowledge of evolution by natural selection and attitudes, but included additional categories of questions, namely:

[1] Ethnicity options offered included Caucasian, African-American, Latino or Hispanic, Asian, Native American, Native Hawaiian or Pacific Islander, Two or More, Other/Unknown, Prefer not to say
[2] Gender identity options offered included male, female, non-binary, intersex, transgender, genderqueer, or my gender is: (write-in)—we included this expanded set of options in an effort to be inclusive of queer students whose identities are often unrepresented on similar surveys (Coburn et al. 2025)
[3] Political Affiliation options offered included republican, democrat, independent, other (write-in)
[4] Religion options offered included Christian (Protestant), Christian (Catholic), Judaism, Hindu, Islam, Buddhism, Sikhism, Other (write-in)

(4) a metacognitive question evaluating students' confidence in their performance with respect to the questions about evolution, and (5) open-ended questions asking students about their experience playing the game. Finally, (6) we administered one additional survey approximately two weeks later by email which included questions belonging to the same categories as the post-game survey, but presented different Conceptual Assessment of Natural Selection questions (see Survey Development for details) than those which appeared in the post-game survey. Students were randomly assigned to one of three treatment levels: a "full game" group, a "simple game" group, and a "control" group (treatment groups). All students were shown a brief video explaining the game controls and instructions.

Treatment Groups

We randomly assigned students to one of three treatment groups that included the control group, the simple game treatment, and the full game treatment. The purpose of comparing the control group to the other categories was to test whether students' answers or attitudes changed just by the passing of time. Players in the control group did not play either game; rather, we asked that they sit patiently for thirty minutes. They were allowed to read, do homework, or otherwise entertain themselves for the balance of that time. We showed control students the same instructional video as the simple and video game treatment students. The purpose of comparing the simple game treatment to the other categories was to test the effect of playing a video game against the effect of playing a video game that displayed evolutionary processes. Students randomly assigned to the simple game group played a reduced version of the game that removed as many elements of evolution by

natural selection as possible without making the game unplayable. In specific, players no longer played as a population but rather as a single individual each generation. This individual did not change from level to level and had an HP ("hit point") bar that allowed it to receive damage repeatedly from predators before expiring unlike all playable "full game" characters which died after any contact with a predator. Additionally, the "end of level" debriefing screen that shows players the average change in phenotypes after selection was removed from the simple game. In the full game, students played as an entire population of between five and one-hundred characters, all controlled at the same time by the same suite of controls. Players selected a heritability before gameplay, and at the end of each level, surviving warlaks were included in a breeding pool to produce the next generation of playable characters.

Administration

We assessed more than 900 students at Auburn University in the spring of 2023 who were enrolled in at least one of the following courses: Principles of Biology (first-year level), Organismal Biology and Honors Organismal Biology (first- and second-year level), Anatomy and Physiology (second-year level), Evolution (second-year level), Microbiology (second-year level), Ornithology (third- and fourth-year level), Ecology (second- and third-year level), and Biomedical Physiology (fourth-year level). We randomly assigned arriving students to one of the three treatment groups described above. To statistically test concept comprehension rather than familiarity with questions, we randomly assigned students to one of four similar surveys which contained unique evolution questions but covered the same concepts (A, B, C, or D; see Survey Development for more details). All assessments

were administered in large, auditorium-style rooms on the Auburn University campus across eight dates and no students participated twice. We worked with instructors to offer all participants extra credit (not to exceed 2% of their final grade) and all participants opted in to the study. Dates and hours took place outside normal classroom hours for each of the classes listed.

Survey Development

Table 19: Conceptual Assessment of Natural Selection Survey

Question	Evolutionary Concept (Conceptual Assessment of Natural Selection)
1	evolution
2 (excluded)	selection
3	inheritance
4 (excluded)	variation
5	mutation
6 (excluded)	selection
7	evolution
8	inheritance
9	evolution
10	evolution
11	mutation
12 (excluded)	selection
13	variation
14 (excluded)	inheritance
15	evolution
16	selection
17	evolution
18	variation
19	mutation
20	selection
21	evolution
22	inheritance
23	mutation
24	evolution

To address our first research question, what is the effect of playing a novel video game on undergraduate students' attitudes towards science and the video game itself, we deployed the Scientific Attitude Inventory II, a validated concept inventory used to assess students' attitudes toward science (Moore and Foy 1997). To address our second research question, what is the effect of playing a novel video game on undergraduate biology students' self-efficacy in conceptual understanding of evolution and their actual understanding of evolution, we employed the Conceptual Assessment of Natural Selection (Kalinowski et al. 2016), a validated concept inventory that assesses students' knowledge of several components of evolution by natural selection, namely variation, evolution, selection, mutation, and inheritance. To assess students' understanding of concepts in evolution rather than their familiarity with questions, students were randomly assigned to one of four possible survey groups (A, B, C, or D) at each stage (i.e. pre-game, post-game, and final) of the survey process. Each survey asked the same demographic questions of each student but sampled an equal number of questions from a subset (19/24) of the Conceptual Assessment of Natural Selection. Of the twenty-four original questions in the Conceptual Assessment of Natural Selection, we chose to exclude five questions that concerned concepts not explored in the video game (e.g. resource allocation, seen in Table 20).

Table 20: Excluded Conceptual Assessment of Natural Selection Questions.

Excluded Question	Q2	Q4	Q6	Q12	Q14
Reason for Exclusion	Focuses on resource use (resource use absent from the game)	Focuses on how all traits vary; the game focuses on one or two traits (ability and size) at a time	Focuses on resource use (resources absent from the game)	Focuses on population growth, but game populations are fixed	Focuses on ontogeny; no ontogenetic component to the game

We preserved the original authors' categorizations of the Conceptual Assessment of Natural Selection, tagging each question as corresponding to one of five evolutionary concepts: evolution, selection, inheritance, variation, or mutation. When creating question pools to randomly select questions for each student, we ensured that every student would be evaluated using the same number of questions from each category (e.g. if there were only two "inheritance" questions present in the question pool, those questions were evenly distributed across each survey). We assessed each student's attitude toward science in both pre-treatment and post-treatment surveys, presenting a subset of eight pairs of questions from the "Scientific Attitude Inventory II" (SAI II). The authors of the concept inventory designated each statement as expressing either a positive sentiment or a negative sentiment, and each positive statement has a corresponding negative (and vice versa). To better understand students' feelings about the game, specifically, we also presented students with six original statements and asked them to score their agreement on a Likert-type scale from 0-10 (where zero represents complete disagreement and ten

represents complete agreement). In statements one through four, we directly reference gameplay, so we excluded responses from students in the control group who did not play any game. In questions five and six, which do not reference a game, we included participants at all three treatment levels.

Table 21: Original Questions Posed to Students.

Original Questions					
I understand evolution better after having played this game.	I enjoyed this game.	Evolution is easier to understand after having played this game.	I learned something about evolution by playing this game.	Learning through video games makes me more enthusiastic about pursuing a STEM career.	I would prefer to learn through video games than through my current class activities.

Data Preparation

We removed any survey entries submitted by individuals associated either directly or indirectly with the investigation (i.e. those used for piloting purposes). We anonymized all data before analysis. Although identities were not included in the analyzed dataset, we excluded any lines of data initially submitted without a student ID (in order to avoid any potential duplications or test submissions by authors).

Quantitative Statistical Analysis

We evaluated student performance on the evolution concept inventory portion of their surveys using a generalized linear model with a binomial error distribution and a logit link

function that evaluated the probability of a student getting a given number of questions correct out of a possible six; this model included the number of questions correct out of six on the post-game survey as a response variable, a fixed categorical effect of treatment level, a continuous fixed effect of pre-game survey score, a fixed categorical effect of survey (e.g. “A”, “B”, etc.), and a fixed term interacting pre-game survey score and treatment.

We analyzed student confidence in the accuracy of their own post-game survey evolution concept inventory answers using a linear mixed model that included confidence on a continuous Likert-type scale (0-10) as a response variable, fixed categorical effects of treatment, survey, pre-game survey confidence in answers, and the interaction between pre-game survey confidence and treatment level, and random effects of date taken.

We evaluated attitudes toward science using a 0-10 Likert-type scale for each Scientific Attitude Inventory statement, and each statement was designated either positive or negative and had its score adjusted to indicate either preference for or against an idea. For example, a score of 10 on a positive question expressed strong positive support for an idea whereas a score of 10 on a negative question represented strong opposition; we recalculated negative scores so that a “10” became a “0”, a “9” became a “1”, and so on. We evaluated post-game survey attitude scores using a linear model which included post-game survey attitude scores as a continuous response variable, a categorical fixed effect of treatment, a continuous fixed effect of pre-game attitude score, and a term interacting treatment and pre-game attitude score. We analyzed these data using the lme4 package in R (Version 4.3.2, R Core Team 2025).

We visualized and estimated marginal means using ggplot2 (Wickham et al. 2025) and emmeans (Lenth et al. 2025) respectively in R. In all generalized linear mixed models, we reported the maximum likelihood estimates and estimated standard errors on the latent scale. When visualizing estimated marginal means, we presented the 95% confidence interval unless otherwise stated (details presented in figure captions).

Concept Inventory Responses

We evaluated the Conceptual Assessment of Natural Selection responses of 902 student participants. We then scored respondents' pre-treatment and post-treatment surveys each on a scale from zero to six (out of six) based on the number of questions that they answered correctly. We grouped these student scores by percentage of questions answered correctly (0% - 100% of six questions asked), into a low-scoring group which included those who answered 0-2 questions correctly, a medium-scoring group who answered 3-4 questions correctly, and a high-scoring group included those who answered 5-6 questions correctly.

Qualitative Statistical Analysis

We asked students four open-ended questions: (1) "What did you think of the game?", (2) "What, if anything, about this game helped you to better understand evolution?", (3) "Would you want this game to be part of a class that taught evolution by natural selection? Why or why not?", and (4) "How could this game be improved?". Three authors of the paper coded and binned student responses into discrete categories that were developed systematically after initial review. All questions included an "NA" category and all

questions except “Would you want this game to be part of a class that taught evolution by natural selection? Why or why not?” included an “Uninformative” category.

Results

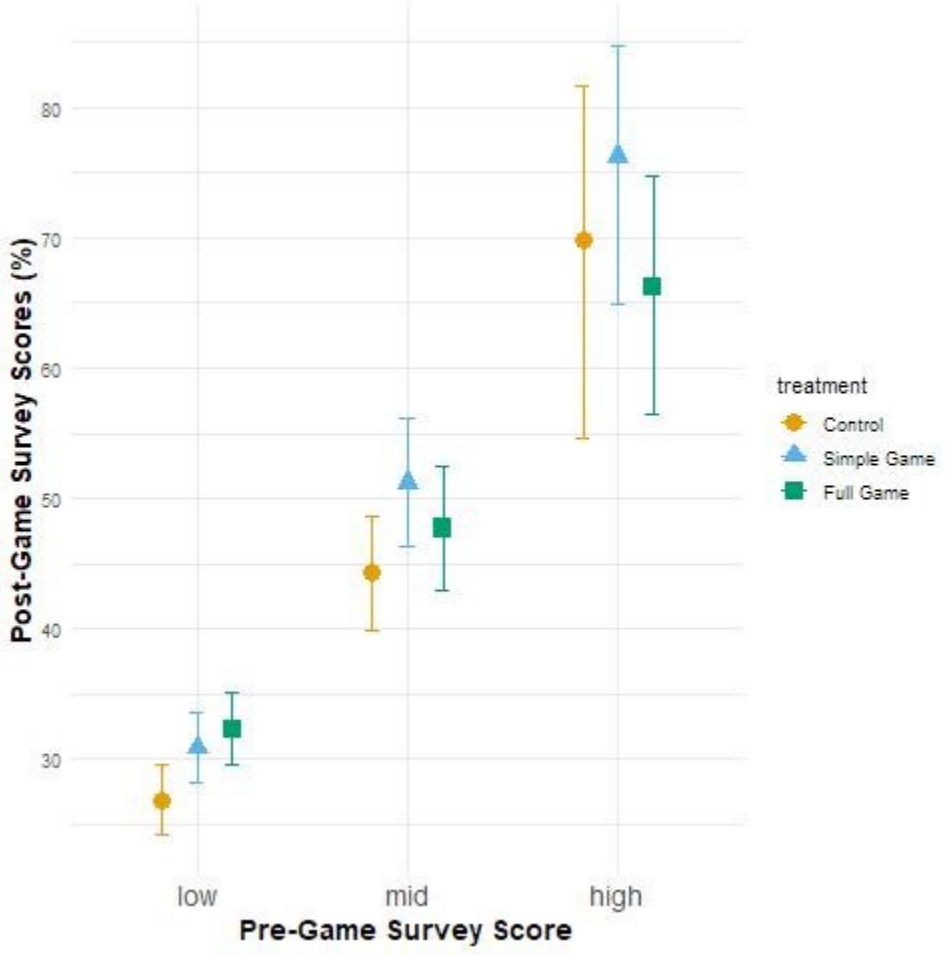


Figure 7: Post-Game Survey Scores (estimated marginal means) by Treatment and Pre-Game Survey Scores.

Error bars represent 95% confidence limits. “Low” indicates a Pre-Game Survey score of 0-2, “Mid” indicates a score of 3-4, and “High” indicates a score of 5-6 (out of a possible six).

Within the mid- and high-range pre-game score groups, we found no significant differences between the post-game scores of students who played either version of the game and those in the corresponding control group (see *Chapter 3 Methods* for details). In the low-scoring group (more than two thirds of all participants, n=619, Table 22), students who played the full and simple game scored significantly higher on average than students in the control group. The mean pre-treatment concept inventory score was 1.987, the median was 2, and the mode was 1. The post-treatment mean was 2.029, the median held at 2, but the mode rose to 2. More than two-thirds (n = 619) of participants scored a 2 or lower on the pre-game concept inventory. However, and among those, though the mode remained stable at 1, the mean rose from 1.275 to 1.691 and the median rose from 1 to 2.

Table 22: Student Performance on The Conceptual Assessment of Natural Selection. P-values at 0.000 are rounded down from $2e^{-16}$.

	Estimate	Std. Error	z value	Pr(>|z|)
Control Group/No Game (Intercept, Low Score on Pre-Game Survey)	-0.680	0.087	-7.818	0.000
Full Game	0.261	0.094	2.795	0.005
Simple Game	0.195	0.094	2.082	0.037
Scored High on Pre-Game Survey	1.842	0.342	5.394	0.000
Scored Medium on Pre-Game Survey	0.773	0.114	6.755	0.000
Took Survey B	-0.513	0.085	-6.069	0.000
Took Survey C	-0.641	0.093	-6.900	0.000
Took Survey D	-0.138	0.084	-1.641	0.101
Full Game x High Pre-Game Score	-0.429	0.405	-1.060	0.289
Simple Game x High Pre-Game Score	0.133	0.444	0.299	0.765
Full Game x Medium Pre-Game Score	-0.124	0.163	-0.759	0.448
Simple Game x Medium Pre-Game Score	0.085	0.164	0.519	0.604

Student Confidence in Answers

Student confidence from the post-game survey was strongly predicted by student pre-game survey confidence values ($\beta = 0.70$, df = 886.09, $p = 2e^{-16}$). Confidence rose significantly between the pre-game survey and post-game survey at all treatment levels on average with the largest increase in confidence among students in the full game treatment ($\beta = +1.32$, df = 887.86, p = 0.00058), followed by the simple game treatment ($\beta = +0.97$, df = 885.96, p = 0.019), compared with the "no game" control group (intercept = 1.37, df = 258.65, $p = 3.83e^{-05}$). We found a small but significant negative interaction between pre-treatment survey confidence and full-game treatment level ($\beta = -0.15$, df = 885.80, p = 0.02), though students' average confidence levels remain higher in the post-game score

until a pre-game score of approximately 8.5/10 (see Table 23) at or above which point students' post-game confidence on average was approximately equal to or lower to that of the pre-game survey. Most participants' (n = 525) confidence scores in the post-game survey fell below 7, and nearly 80% (n = 718) fell below 8.

Table 23: Student Confidence in Evolution Concept Inventory Answers.

Estimates are on the latent scale (log odds). P-values at 0.000 are rounded down from $2e^{-16}$.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Control Group/No Game (Intercept, Survey A)	1.374	0.328	258.650	4.190	0.000
Full Game	1.315	0.382	887.144	3.444	0.001
Simple Game	0.965	0.398	885.254	2.428	0.015
Confidence in Answers on Pre-Game Survey	0.703	0.050	886.095	14.164	0.000
Took Survey B	0.363	0.162	885.687	2.248	0.025
Took Survey C	-0.073	0.167	888.277	-0.436	0.663
Took Survey D	0.224	0.166	887.859	1.346	0.179
Pre-Game Evolution Score	0.098	0.049	888.482	2.005	0.045
Full Game x Pre-Game Confidence	-0.152	0.067	885.804	-2.270	0.023
Simple Game x Pre-Game Confidence	-0.112	0.069	884.913	-1.616	0.106

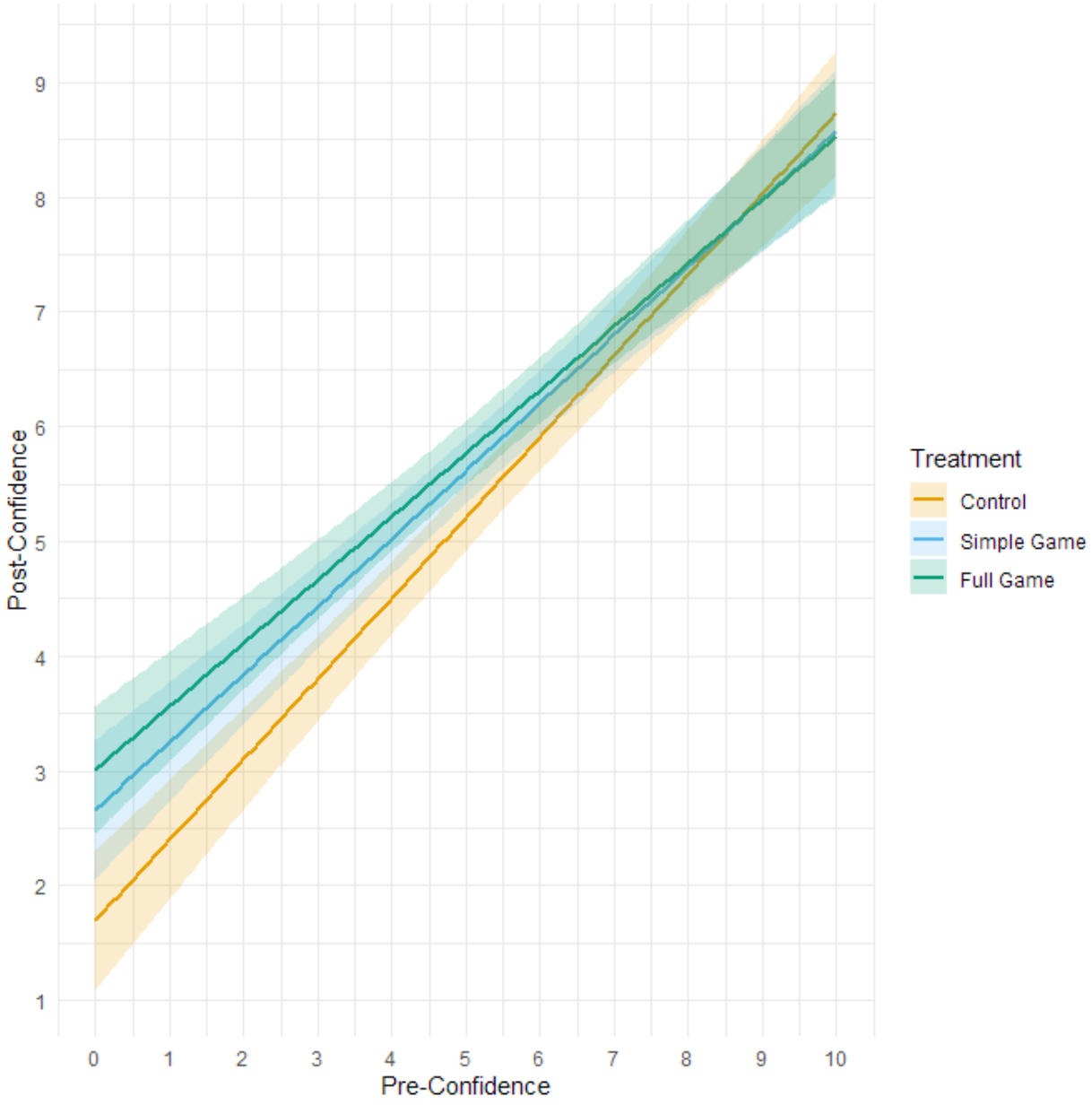


Figure 8: Participants' Confidence (estimated marginal means) on a Likert-type scale of 0-10 in the accuracy of their Conceptual Assessment of Natural Selection Answers By Treatment in pre-treatment surveys (X) and post-treatment surveys (Y).

Ribbons represent twice the standard error.

Scientific Attitude Inventory II (SAI II) Responses

We evaluated participant attitudes towards science using their responses to a subset of eight paired positive and negative prompts from the Scientific Attitude Inventory II (those from position statements 5A/5B and 6A/6B) to identify any changes in attitude that occur during gameplay. We presented students with a modified ten-point Likert-type response scale on which they could choose values on a scale from '0, strongly disagree' to '10, strongly agree'. After adjusting negative values to correspond to positive values on the scale (e.g., a response of "10" to a negative question was changed to a score of "0"), lower scores represent negative attitudes (minimum total score of 0), whereas higher scores represent positive attitudes (maximum total score of 140).

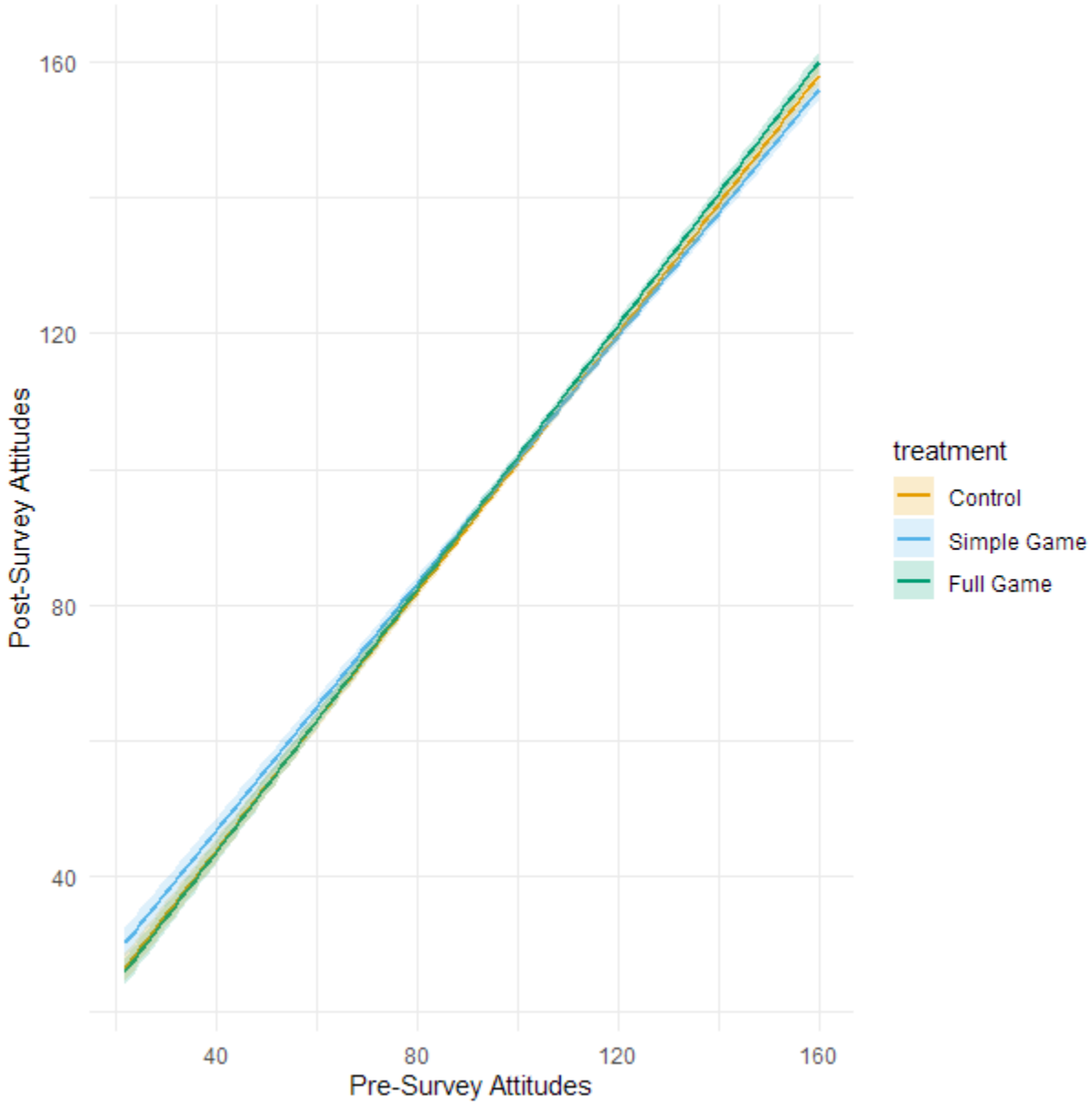


Figure 9: Overall Student Attitudes Toward Science (estimated marginal trends) Before (X) and After (Y) Treatment.

Ribbons represent twice the standard error. Responses from each question summed on a standardized scale.

Both pre-treatment and post-treatment attitudes were generally positive across all three treatments, but treatment did not explain any significant changes in attitude between pre- and post-treatment surveys (Table 24).

Table 24: Student Attitudes Toward Science After Gameplay.

P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Control Group/No Game	5.613	2.516	594.934	2.231	0.026
Full Game	-1.059	3.457	782.642	-0.306	0.759
Simple Game	4.521	3.602	781.567	1.255	0.210
Pre-Game Attitude Scores	0.953	0.023	782.998	41.526	0.000
Full Game x Pre-Game Attitude Scores	0.019	0.032	782.333	0.591	0.555
Simple Game x Pre-Game Attitude Scores	-0.041	0.033	781.612	-1.237	0.217

Student Responses to Gameplay

Table 25: Statement Prompts Provided to Students

Statement 1	Statement 2	Statement 3	Statement 4	Statement 5	Statement 6
"I understand evolution better after having played this game."	"I enjoyed this game."	"Evolution is easier to understand after having played this game."	"I learned something about evolution by playing this game."	"Learning through video games makes me more enthusiastic about pursuing a STEM career."	"I would prefer to learn through video games than through my current class activities."

Statement 1: "I understand evolution better after having played this game"

Students' agreement with the statement "I understand evolution better after having played this game" was significantly higher among those who played the full game (β +0.45, df = 602.24, p = 0.046) than those who played only the simple game (intercept = 4.947, df = 14.74, p = $1.39e^{-13}$, Table 26).

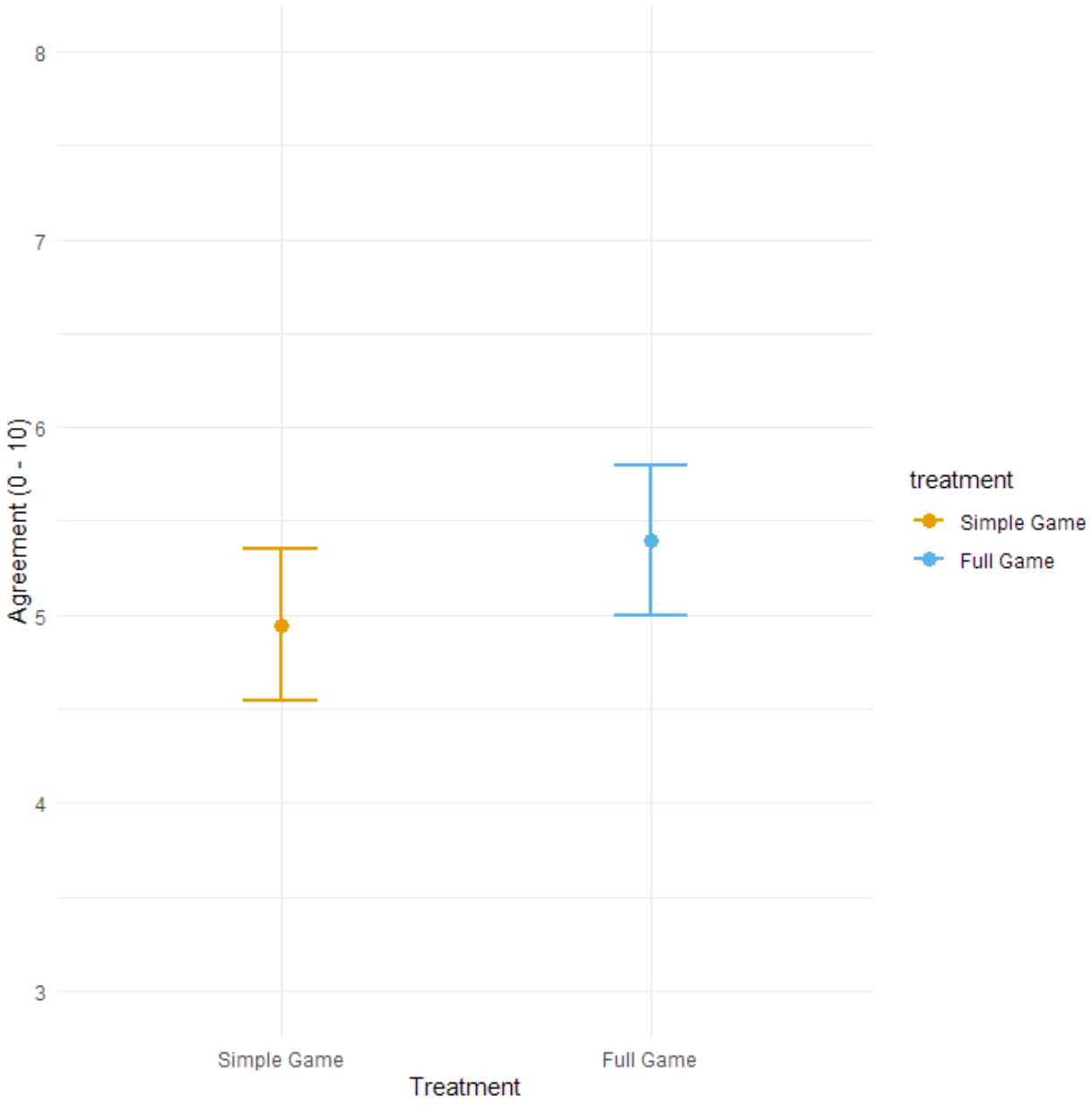


Figure 10: Student responses (estimated marginal means) to the prompt "I understand evolution better after having played this game." by Treatment.

Error bars represent twice the standard error. Responses were scored on a Likert-type scale from 0-10 where zero represents complete disagreement and ten represents complete agreement. Control excluded due to question directly identifying gameplay.

Table 26: Student responses to the statement above by treatment.

P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Simple Game	4.947	0.195	14.74	25.422	0.000
Full Game	0.450	0.226	602.24	1.996	0.046

Statement 2: “I enjoyed this game.”

Students’ agreement with the statement “I enjoyed this game” did not differ significantly between students who played the simple and full versions of the game (Table 27).

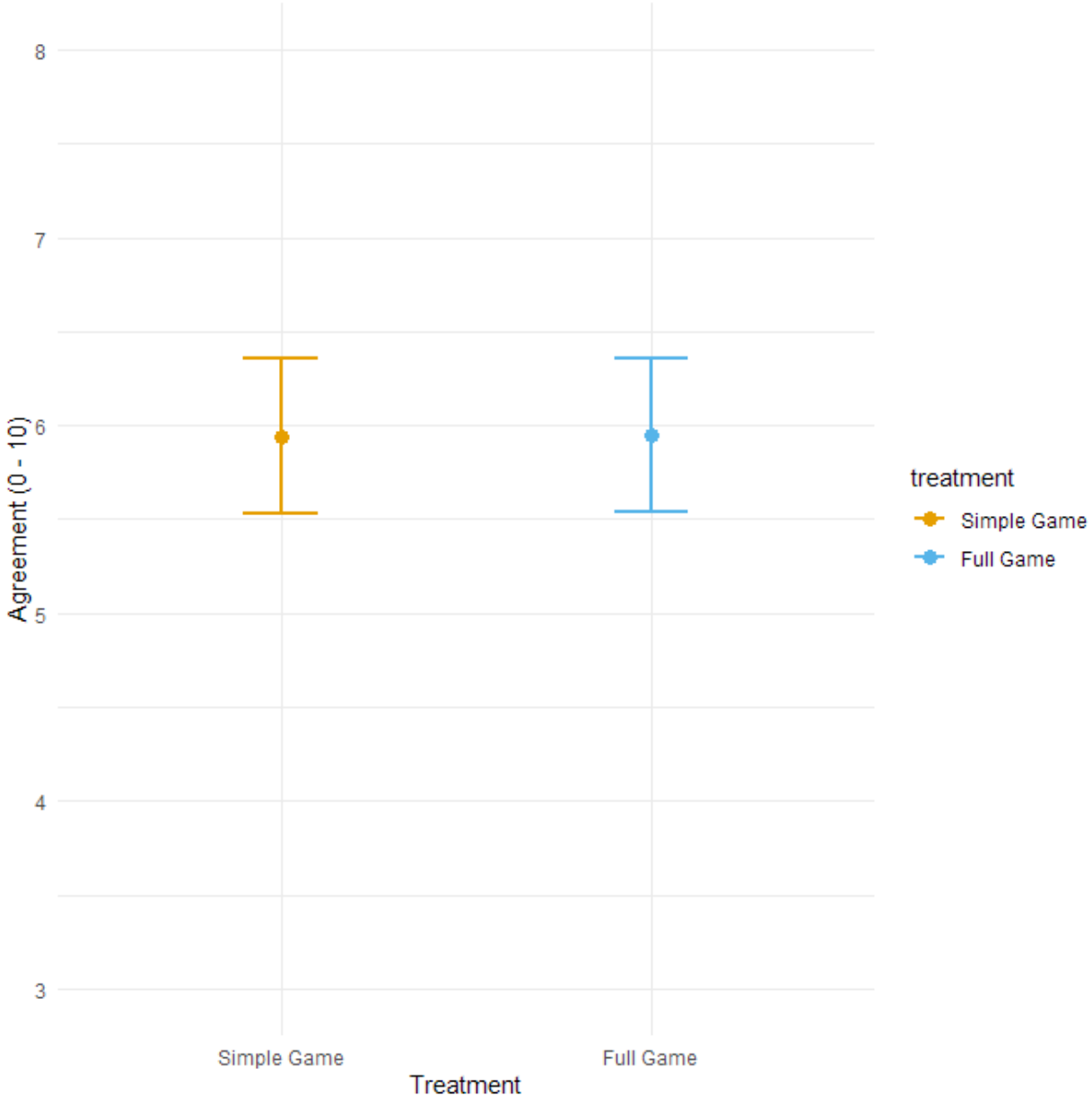


Figure 11: Student responses (estimated marginal means) to the prompt “I enjoyed this game” by Treatment.

Error bars represent twice the standard error. Responses were scored on a Likert-type scale from 0-10 where zero represents complete disagreement and ten represents complete agreement. Control excluded due to question directly identifying gameplay.

Table 27: Student responses to the statement “I enjoyed this game” by treatment. P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Simple Game	5.941	0.201	13.436	29.532	0.000
Full Game	0.007	0.223	600.026	0.031	0.975

Statement 3: "Evolution is easier to understand after playing this game."

Students' agreement with the statement "evolution is easier to understand after playing this game" did not differ significantly between students who played the simple and full versions of the game (Table 28).

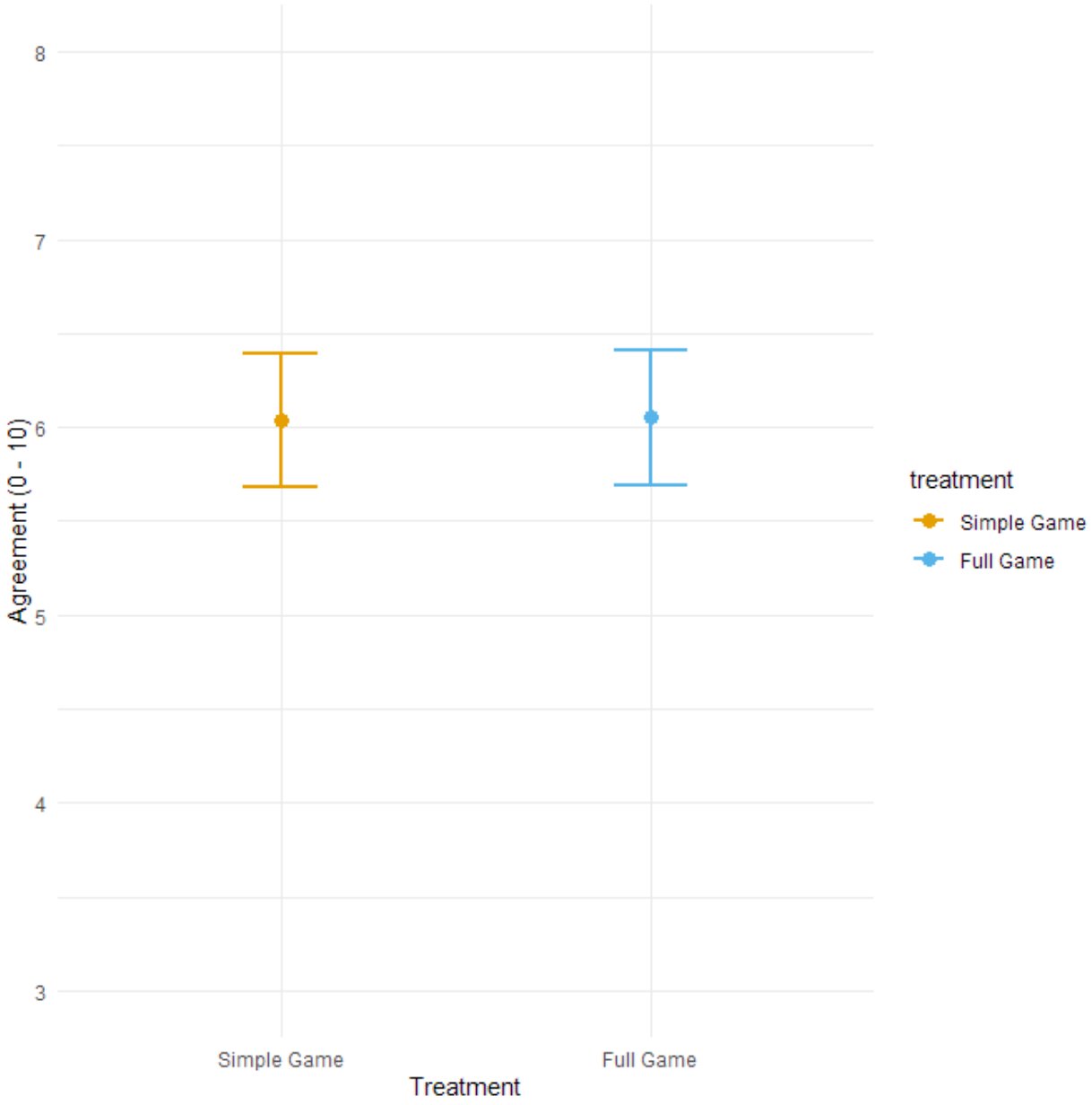


Figure 12: Student responses (estimated marginal means) to the prompt “evolution is easier to understand after playing this game” by Treatment.

Error bars represent twice the standard error. Responses were scored on a Likert-type scale from 0-10 where zero represents complete disagreement and ten represents complete agreement. Control excluded due to question directly identifying gameplay.

Table 28: Student responses to the statement “evolution is easier to understand after playing this game” by treatment.

P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Simple Game	6.033	0.171	19.804	35.340	0.000
Full Game	0.018	0.213	602.221	0.083	0.934

Statement 4: “I learned something about evolution by playing this game.”

Students’ agreement with the statement “I learned something about evolution by playing this game” was significantly higher among those who played the full game (β +0.483, df = 605.96, p = $9.23e^{-05}$) compared with those who played the simple game (Intercept = 4.836, df = 11.341, p = $2.09e^{-10}$, Table 29).

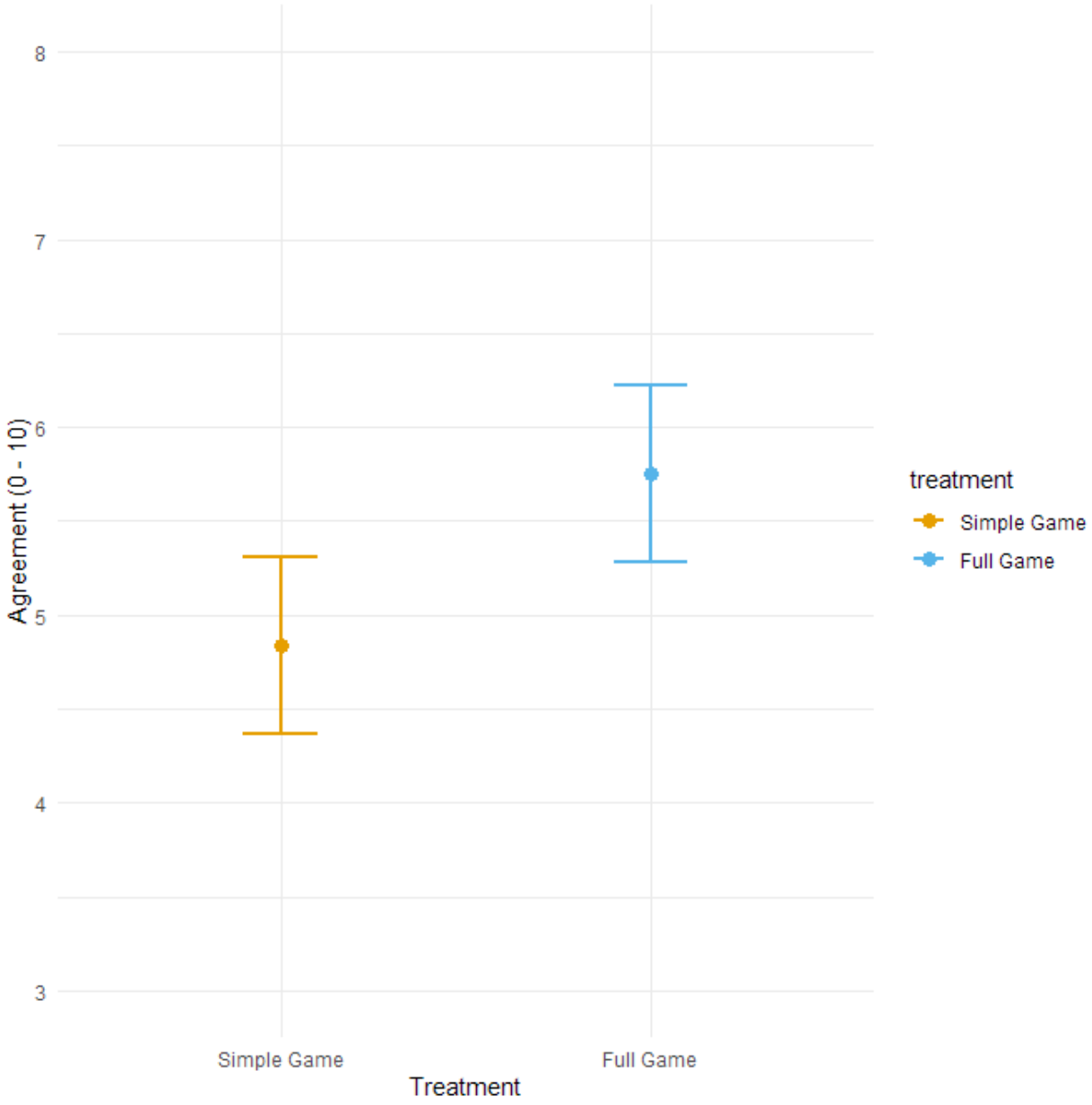


Figure 13: Student responses (estimated marginal means) to the prompt "I learned something about evolution by playing this game" by Treatment.

Error bars represent twice the standard error. Responses were scored on a Likert-type scale from 0-10 where zero represents complete disagreement and ten represents complete agreement. Control excluded due to question directly identifying gameplay.

Table 29: Student responses to the statement “I learned something about evolution by playing this game” by treatment. P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Simple Game	4.836	0.232	11.341	20.875	0
Full Game	0.913	0.232	605.963	3.935	0

Statement 5: “Learning through video games makes me more enthusiastic about pursuing a STEM career.”

Students’ agreement with the statement “learning through video games makes me more enthusiastic about pursuing a STEM career” did not differ significantly between students by treatment (Table 30).

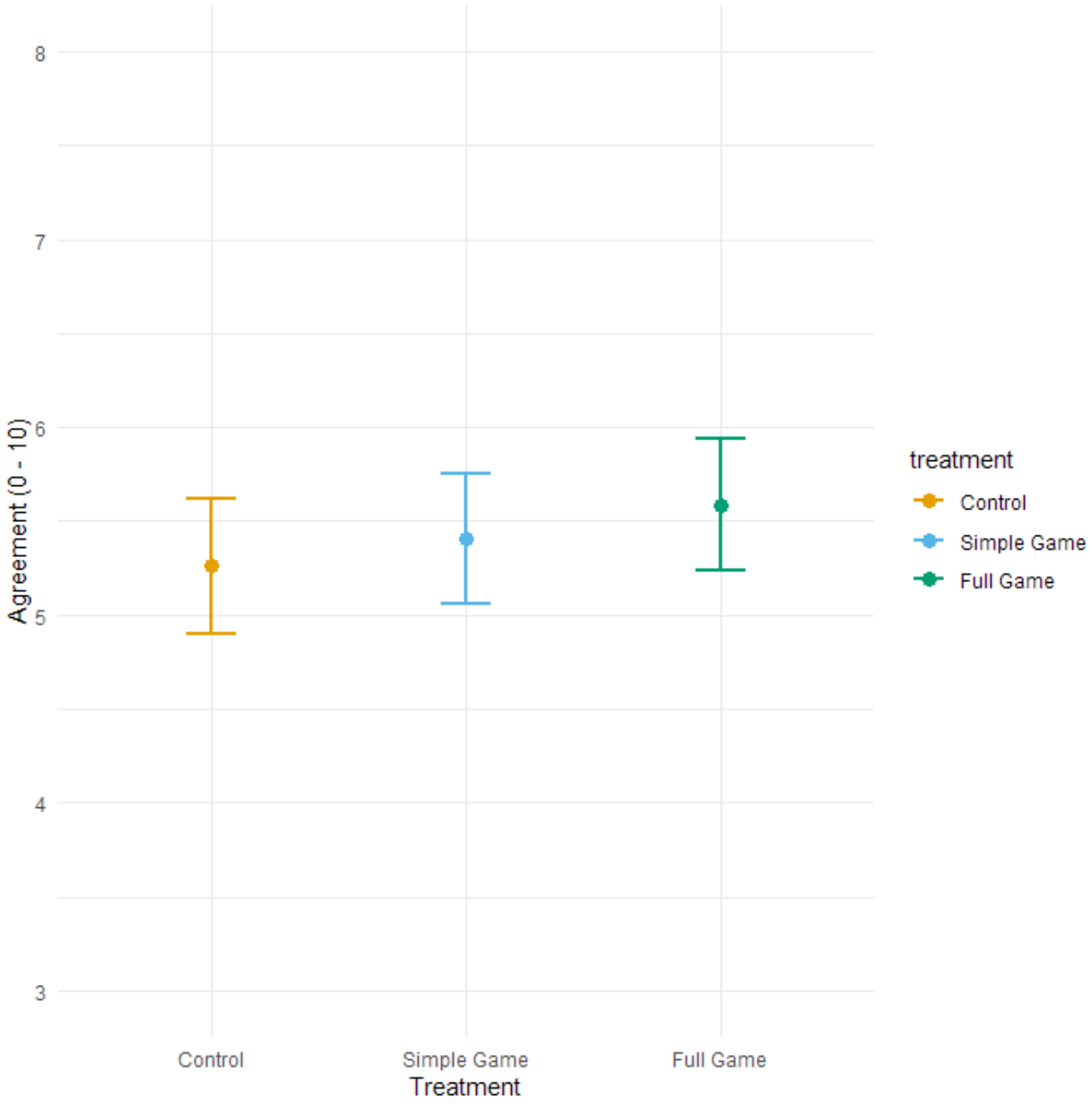


Figure 14: Student responses (estimated marginal means) to the prompt "learning through video games makes me more enthusiastic about pursuing a STEM career" by Treatment.

Error bars represent twice the standard error. Responses were scored on a Likert-type scale from 0-10 where zero represents complete disagreement and ten represents complete agreement.

Table 30: Student responses to the statement "learning through video games makes me more enthusiastic about pursuing a STEM career" by treatment.

P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Control (No Game)	5.262	0.173	20.340	30.466	0.000
Simple Game	0.142	0.218	894.728	0.654	0.513
Full Game	0.325	0.216	894.852	1.503	0.133

Statement 6: "I would prefer to learn through video games than through my current class activities."

Students' agreement with the statement "I would prefer to learn through video games than through my current class activities" did not differ significantly between students who played the simple and full versions of the game (Table 31).

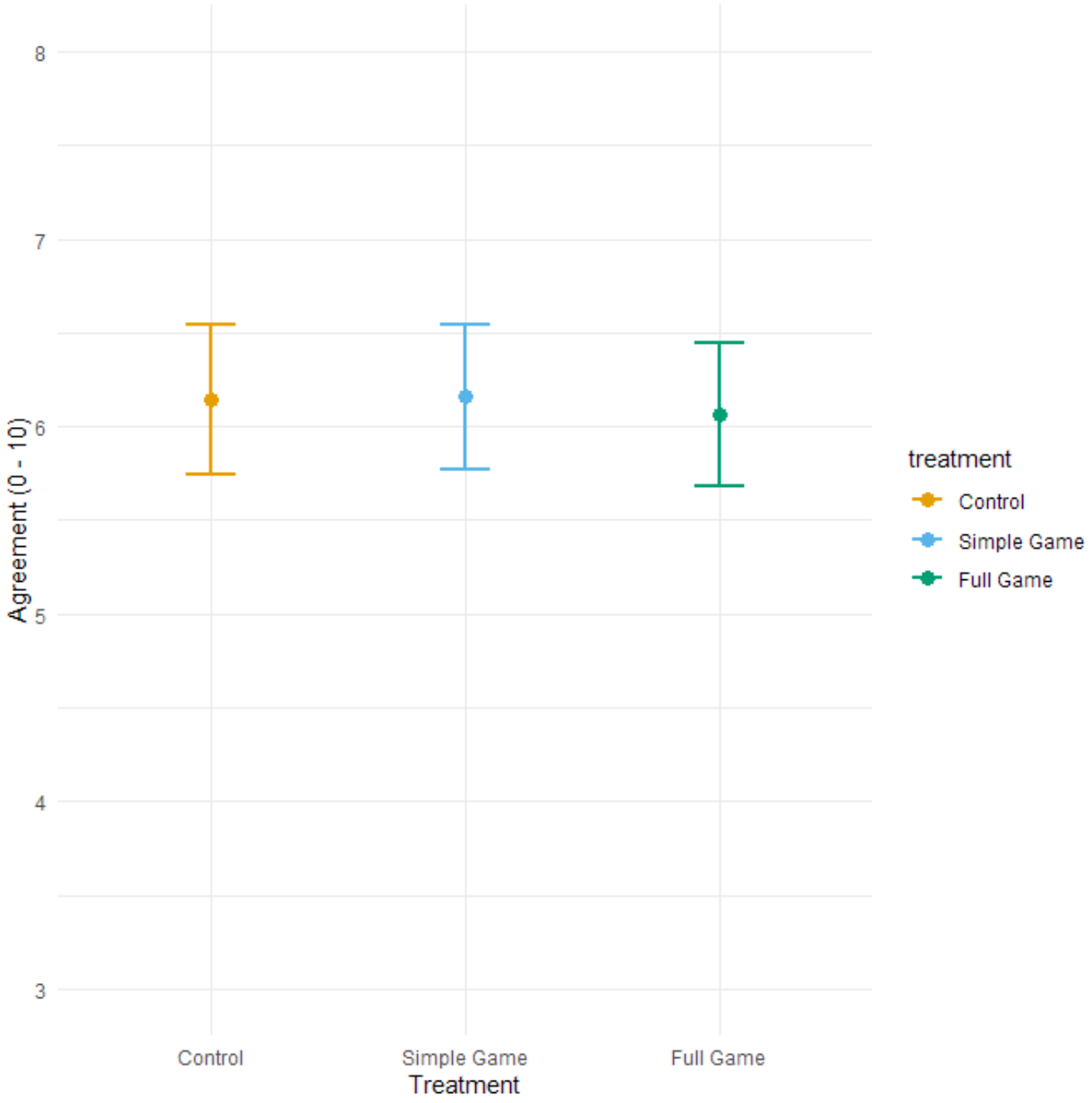


Figure 15: Student responses (estimated marginal means) to the prompt "I would prefer to learn through video games than through my current class activities" by Treatment.

Error bars represent twice the standard error. Responses were scored on a Likert-type scale from 0-10 where zero represents complete disagreement and ten represents complete agreement.

Table 31: Student responses to the statement “I would prefer to learn through video games than through my current class activities” by treatment.

P-values at 0.000 are rounded down from 2e-16.

	Estimate	Std. Error	df	t value	Pr(>|t|)
Control (No Game)	6.145	0.192	26.809	31.980	0.000
Simple Game	0.015	0.234	897.926	0.063	0.950
Full Game	-0.082	0.233	897.693	-0.351	0.725

Qualitative Feedback from Students

We asked students four open-ended questions concerning their perspectives on the game.

Open-Ended Question 1	Open-Ended Question 2	Open-Ended Question 3	Open-Ended Question 4
“What did you think of the game?”	"Would you want this game to be part of a class that taught evolution by natural selection? Why or why not?"	"How could the game be improved?"	"What about this game (if anything) helped you better understand evolution by natural selection?"

Figure 16: Open-Ended Questions Posed to Students.

Open-Ended Question 1: “What did you think of the game?”

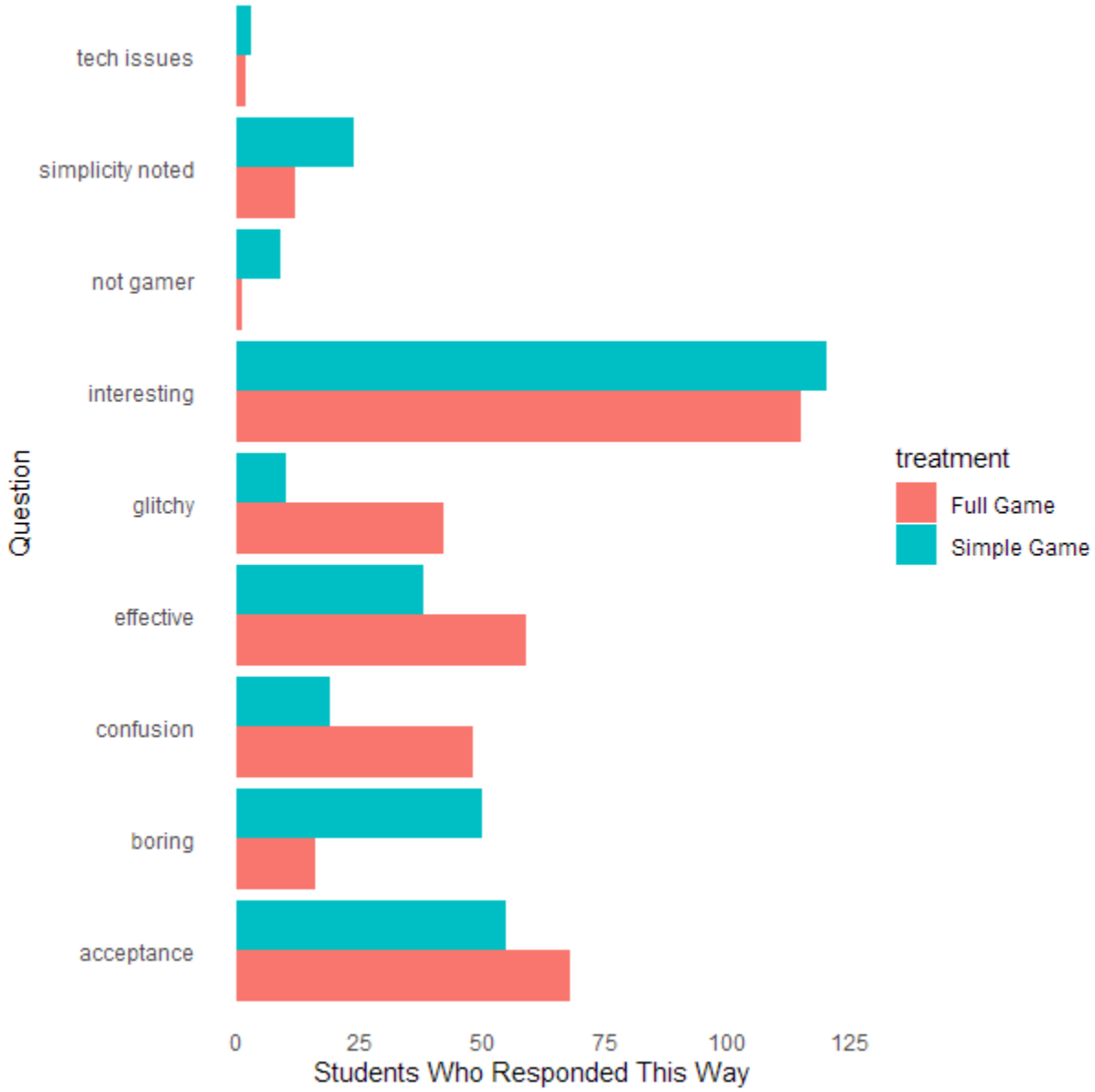


Figure 17: Student thoughts on the game among those who played either the simple game or full game.

Authors coded responses, created categories, and then binned responses into those categories. Control group excluded due to question asking about the game specifically.

We asked students who played either the full or simple game to provide an open-ended response relating their reaction to the game in broad terms. Three authors coded student responses and created eleven categories (Table 32) and binned responses appropriately. Responses binned into the uninformative category did not answer the question clearly or directly, and responses binned into NA did not provide an answer.

Table 32: Categories of Student Reactions to the Game with definitions and examples. (Students quoted directly from surveys except where bracketed)

Category	Definition	Example 1	Example 2
Glitchy	Experienced failures in software **specific to video game**	"it lagged a lot but was cool"	"I liked it, needs some of the kinks worked out though."
Effective	Helped student to learn	"The visual component of seeing what happens combined with doing something made it easier to stay engaged and learn."	"I thought it was a pretty cool design and helps understand evolution in populations better"
Interesting/Engaging/Fun	Innovative; Enjoyable; Novel	"It was a fun and creative way to learn."	"I really enjoyed the game, it was entertaining, while also teaching"
Tech Issues	Experienced limitations due to **outside** technological factors (internet, their computer, etc.)	"I like the idea of the game a lot but I don't know if it was just my computer but it seemed a	"It was hard to play on my out-dated computer"

		little glitchy and slow.“	
Acceptance	Positive responses that were not specific enough to be further coded	“I think it is so cool what you all were able to do through this game.”	“it was cash money.”
Boring	Expressed boredom or disinterest or described it as “repetitive”	“i thought it was super fun but could have more aspects to make it harder like longer levels. “	I thought the game was a little repetitive, but I did feel like I learned a little more about evolution by playing.
Simplicity Noted	Easy to understand; simple	“It was very simple, nothing that was invigorating to continue playing very long”	“kind [of] simple but fun”
Confusion	Expressed confusion	“It was alright. Was a little tough to understand the overall purpose of the game.”	“it was a little confusing at first but once I got the hang of it, it was fun”
Not a Gamer	Rejection due to a lack of enjoyment or experience with videogames	“It was good, I just haven't been a gamer.”	“I personally just do not like computer games in general”

Open-Ended Question 2: "What, if anything, about this game helped you to better understand evolution?"

We asked students who played either the full or simple game: "What, if anything, about this game helped you to better understand evolution?". The vast majority of students answered this question by explaining what concepts related to evolution they better understood after playing the game, so we categorized and binned responses according to whether those students mentioned Lewontin's components of evolution: (1) differential fitness, (2) phenotypic variation, (3) heritable fitness components, or if they instead mentioned better understanding evolution without further qualifying their response (Table 33). Responses were binned into multiple categories if their responses satisfied multiple criteria. Uninformative responses did not provide enough information to be categorized, and no response entries did not respond at all. Since the simple game was not designed to be instructive, all figures and examples reflect full game participants only.

Table 33: Student responses to the question "what, if anything, about this game helped you to better understand evolution?" were binned in the categories below.
(Students quoted directly from surveys except where bracketed)

Category	Example 1	Example 2
Differential Fitness	"Seeing different characteristics that allow for better survival"	"The ones with favorable traits survived and passed the traits down to their children"
Phenotypic Variation	"The ability of the physical appearances and sizes of the player was easier to	"the size changes, and abilities of some of the organisms to jump or fly,

	understand as to how they were able to adapt in the certain environments."	seeing them die but not others"
Heritable Fitness Components (or heritability more generally)	"the traits that individuals pass down, are only passed down to their offspring because they survive"	"Heritable traits of the warlaks that made it through the level getting passed on to their offspring"
Evolution without additional context	"I better understood that basically for evolution is just survival of the fittest and trying not to die"	"The stats at the end were probably the best part for gaining a better understanding of evolution!"

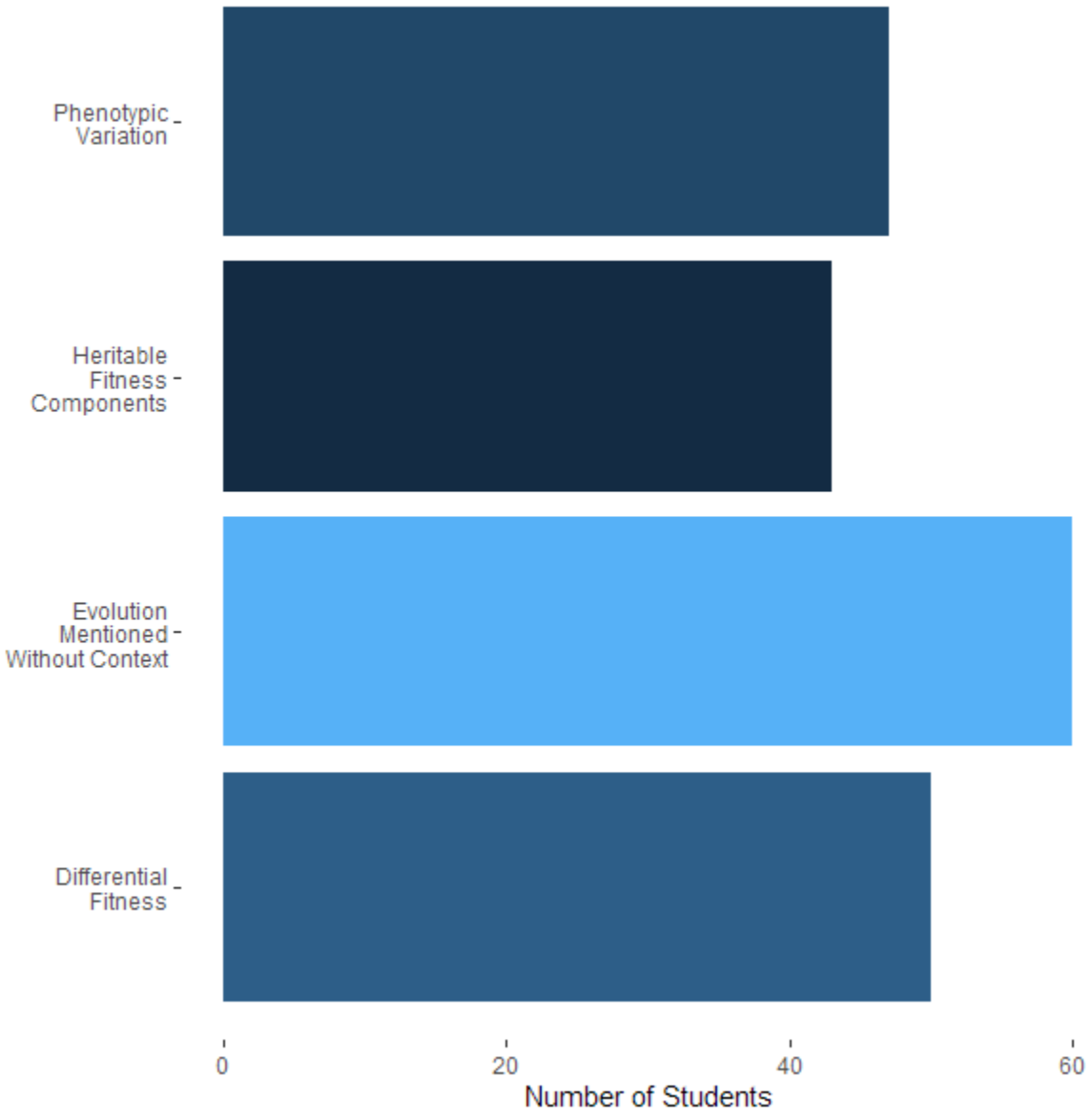


Figure 18: Student perceptions of how gameplay helped them better understand in the full game only.

Open-Ended Question 3: “Would you want this game to be part of a class that taught evolution by natural selection? Why or why not?"

We evaluated students who played either the full or simple game to answer the question “Would you want this game to be part of a class that taught evolution by natural selection? Why or why not?” and binned responses into either positive, positive-conditional, neutral, or negative sentiments (see Table 34 for details). For both the simple and full game treatments, the largest number of respondents fell into the “positive” category, and the overwhelming majority (>75% of all respondents) offered responses that were either positive or positive conditional.

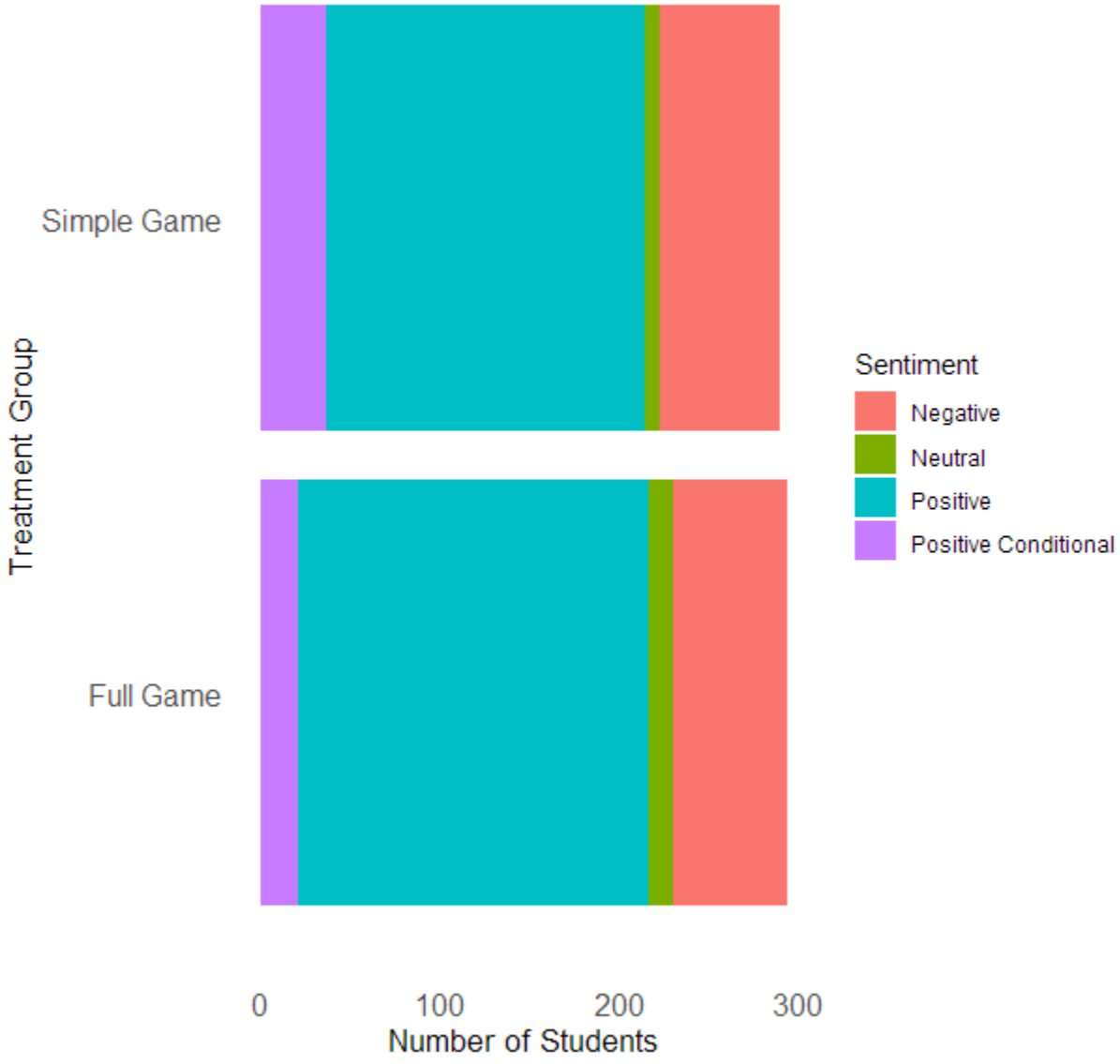


Figure 19: Student sentiments toward the game.

Table 34: Sentiments toward the game with examples

(Students quoted directly from surveys except where bracketed)

Sentiment	Positive	Positive Conditional	Neutral	Negative
Simple Game Example 1	“Yes, this game could be an interactive and fun way to learn more about evolution. Students might remember evolution and understand it better when playing a game.”	“I think this would be a nice activity, however, I don't know how it would necessarily illustrate natural selection since I could only play one turtle at a time. If you were maybe in charge of a small population all at once, that may be more informative? I'm not really sure.”	“Maybe, I don't know if I played it right. I barely got past the first few levels”	“no. It did not really show evolution.”
Simple Game Example 2	“I would want this game to be a part of a class that taught evolution by natural selection because I think it is a fun and entertaining way to learn.”	“I think this would be a nice activity, however, I don't know how it would necessarily illustrate natural selection since I could only play one turtle at a time. If you were maybe in charge of a small population all at once, that may be more informative? I'm not really sure. “	“Maybe”	“No, I am not interested in learning more about evolution, however some others might be.”
Full Game Example 1	“Yes because it is a good demonstration of natural selection, it is hard to visualize when you have only text or pictures. This allows you to	“I learn best by reading material so I think if the game included some sort of text I might like It”	“Maybe! I think it depends on the specific class.”	“No I do not think I learned anything about natural selection that I did not already know from this game”

	connect the dots and be actively involved and engaged in learning”			
Full Game Example 2	“Yes. I think it's a great idea. It did help me see in a short period of time how evolutionary concepts work.”	“i absolutely would, as long as a couple of bugs were fixed!”	“Maybe, I think it would be beneficial for visual learners to see exactly how evolution is impactful.”	“No, it doesnt improve my knowledge enough for me to care about doing it, but it is an interesting concept.”

Open-Ended Question 4: “How could the game be improved?"

We prompted students who played the full or simple game with the question “How could this game be improved?”. We created categories based on their responses and binned them. We designated any answers that did not contain enough information to be coded to be uninformative and those responses with contained no information or response as NA. As the simple game is not intended to be instructive, all values and figures reflect answers given from students who played the full game. The informative categories which received the most responses signaled a need for more bug fixes or more clear signposting (in terms of game design), and the next largest group comprised those students who felt the game needed no changes (Table 35).

Table 35: Student responses to the question ‘how could the game be improved’?

(Students quoted directly from surveys except where bracketed)

Category	Definition	Example 1	Example 2
Difficulty of Concepts	game itself was unbeatable; eagles too fast, etc.; too many things going on at one time, more challenging	“I'm not sure I just didn't understand what I was doing”	“I just didnt really understand what was going on “
Difficulty of Gameplay	topic was difficult; evolution was hard; didn’t understand what was happening; didn’t see the point; didn’t see link to evolution; bring more attention to evolution components	“Make it harder”	“I think the game could provide more information on the types of animals being involved in the game. It was difficult for me to figure out a way to survive from the large birds so I think that is helpful when trying to explains that animals must evolve in order to stay alive.”
Fix Broken Features	Requested bug fixes; noticed glitches; something didn't work; broken game; flow of game; hard to see; better graphics; make them faster; easier controls	“I think it could be programmed a little better so that it does not lag. “	“The warlacks were difficult to see for me.”
More of Existing Features	requesting more of anything the game already delivers; includes more of or changes to features already in game; anything concerning features already in the game; more depth to game	“more animal options”	“Switch it up and have different stages each time. “
More New (Non-	anything concerning features not already in the game	“Yes, if it was made easier to play and had	“obviously you could make the game extremely complex taking in all the factors that play into an environment and its ecosystem, but that would be beyond the scope of the game trying to

Existing) Features		more realistic characters and a home base.“	simply communicate the basic concepts of natural selection”
More Signposting	Request for more instructions, signposting, direction in game.	“The game could have words or instructions in it.”	“maybe provide more detail of what is going on in the game”
No Changes Needed	Students had no suggestions and liked the game ‘as-is’.	“I thought it was really good!!”	“no change needed”

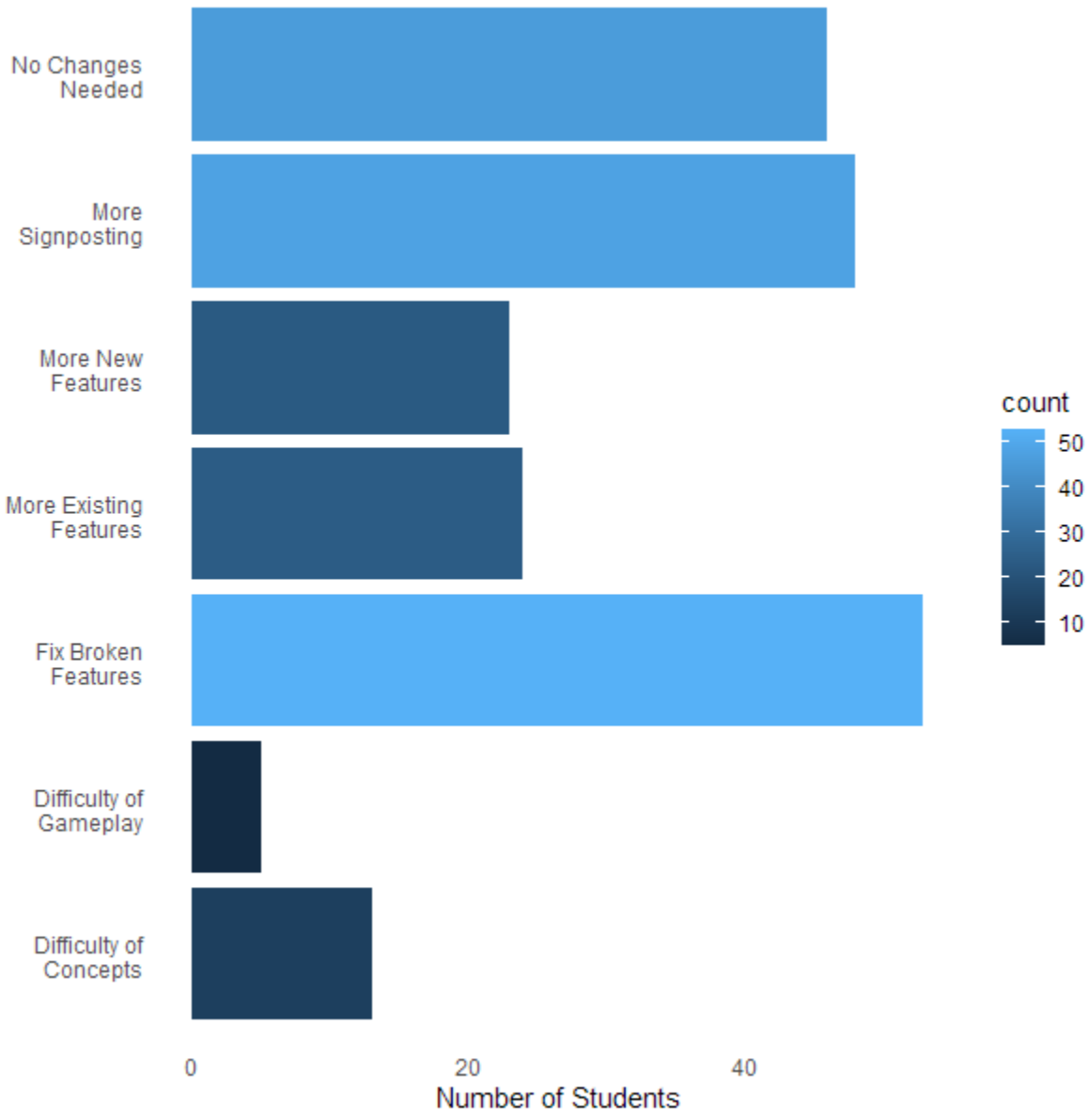


Figure 20: Student suggestions for how the game could be improved binned by category.

Discussion

Summary of Findings

We developed a game designed to effectively demonstrate the process of evolution by natural selection at work in a population of fictional characters in a simulated ecological landscape. We argued that undergraduate students struggle to understand evolution by natural selection because traditional lessons teaching it can effectively communicate the results of evolution but often struggle to represent the processes shaping it. We predicted that students who played a video game that allowed them to directly interact with the process of evolution by natural selection in real-time for half an hour would score significantly higher on an evolution concept inventory assessment (the Conceptual Assessment of Natural Selection) than students who did not play the game for the same half-hour.

Both our hypothesis and prediction held among those students in the low-scoring pre-game group in that students who played the full game were significantly more likely to get a given question correct on their post-game concept inventory assessment than students who did not play for the same half-hour (Table 22), though the differences between the full and simple games in improving post-game student concept inventory scores were small. Interestingly, students who played a simple version of the game that withheld the evolutionary components also performed significantly better than the control group (though not by as much as those who played the full game), and this may be due to games' general tendency to increase student engagement (Leith et al. 2016), perhaps

relaxing anxieties that might have impaired their performance on a pre-treatment survey. Students expressed much more frustration with glitches in the full game than in the simple game, and this may have impacted their ability to experience the full, intended effect of the game. We found the game to be effective at improving academic outcomes for less-prepared students, even without the context of a lesson or curriculum. The formation of supporting materials could further extend the effectiveness of our game.

Potential Uses

Warlak is immediately available for use and can be accessed through a link online by anyone with access to the internet. We designed the game to be run on all browsers commonly used in 2025, including Microsoft Edge, Chrome, Safari, and Mozilla Firefox. This game can be deployed in university classrooms at levels of instruction ranging from introductory biology to courses centered explicitly around evolution and can be available to students both during and outside of class hours, requiring little instruction to get students started.

In terms of addressing common academic challenges related to evolution, educators can use this digital experience to highlight aspects of evolution that are often overlooked, such as how selection on continuous traits often yields a spectrum of phenotypes rather than a handful of simple 'Mendelian' outcomes (Haskel-Ittah et al. 2020; Guevara et al. 2023) or how trait optimization is environment-specific and may vary meaningfully even across similar populations. Knowledge gaps aside, one of the most compelling reasons to adopt video games like Warlak into a classroom environment is the

expected student "buy-in" (Wang et al. 2022; Gui et al. 2023); more than 70% of the students surveyed in the full game treatment responded positively when asked whether or not they would want this game to be part of their course curriculum.

Finally, this game frees students from high-pressure social interactions that can be challenging for those who identify as introverted (Smiderle et al. 2020; Thongthip et al. 2024) or struggle with social anxiety (de Rutte et al. 2025), both traits that have been associated with student struggle in environments that promote or require direct social interaction.

Support from Literature

Although tests of the efficacy of video games as teaching tools for evolution are scarce, our findings strongly align with recent literature supporting the use of video games in education; specifically, video games have been linked to increase retention, heightened engagement, and more fun while learning (Annetta 2008; Ball et al. 2020; Franceschini et al. 2022; Lavalle et al. 2025) by work done since the turn of the millennium. Although historically, video games were dominated by male players, female players now constitute almost half of all active gamers (Fordham et al. 2020). Video games show promise also as learning tools that can reduce stereotype threat (Fordham et al. 2020), or the perception of oneself as lesser when exposed to stereotype-reinforcing biases (Steele 1997; Osborne 2007), when students are not primed with male-biased gaming stereotypes prior to gameplay. Our research did not find that half an hour of gameplay had any significant effect on interest in pursuing a STEM career, but previous research has demonstrated a positive relationship between gaming and an expressed desire to pursue work in STEM

(Clark and Ernst 2009; Mayo 2009), and it may be that longer-scale exposure to gaming promotes a sustained interest in these pursuits.

Limitations

Although our results strongly support the use of Warlak in classrooms as a teaching tool, we should note that some students expressed aversion to gaming in general. Moreover, our work evaluates the efficacy of the game at improving students' knowledge of evolution in the short term (i.e. over just a half an hour), but previous work shows that retention rates for evolution tend to be poorer in the long-term (Anderson et al. 2002; Buckberry and Burke da Silva 2012), and we suggest research that explores the effectiveness of video games over semesters, years, or entire college careers. We note also that at the time of the study, the game did retain some bugs from development (noted by several students) that may have interfered with some participants' experience.

The Future of Warlak

One of the many benefits of developing video games as course material is that they can be adaptively updated, improved, and even reimagined (often with minimal effort once initially produced) based on the changing needs and interests of students. Below, we offer some potential additions to the game based on student feedback and our reflection.

Table 36: New Ideas for 'Warlak'

New/Expanded Concept	Possibilities	Academic Benefit
Additional Environments	Pelagic marine environment, estuary, prairie, beach, arboreal environments, a petri	Expand students' perceptions of how organisms interact in different ecosystems and

	dish filled with diverse microflora	how the intersection of environments can shape evolutionary trajectories.
Weather and Climate Variation	Seasonal progression across a single generation; snow, rain, increases or decreases in predator density; season-specific obstacles.	Encourage students to consider temporal variation when contemplating evolution.
Immigration	Allow new warlaks to enter the population, bringing genetic diversity (and potentially phenotypic hallmarks like new 'colors' so that immigrant descendants can be traced)	Demonstrate the importance of gene flow across populations

Take-Home

We find that Warlak improves students' knowledge of evolution after playing for just half an hour, adding broad support to a phenomenon already widely observed in literature. We call for greater inclusion of video games in classroom environments and strongly encourage the development of new games, especially those that target subjects traditionally challenging to teach.

References

Abenes, F. M. D., and D. G. Caballes. 2020. Misconceptions of science teachers in evolution. Biometrics and Bioinformatics 12:31–38.

Achorn, A. M., and G. G. Rosenthal. 2020. It's Not about Him: Mismeasuring 'Good Genes' in Sexual Selection. Trends in Ecology & Evolution 35:206–219. Elsevier Current Trends.

Acker, P., F. Daunt, S. Wanless, S. J. Burthe, M. A. Newell, M. P. Harris, R. L. Swann, C. Gunn, T. I. Morley, and J. M. Reid. 2023. Additive genetic and environmental variation interact to shape the dynamics of seasonal migration in a wild bird population. Evolution 77:2128–2143.

Agashe, D., M. Sane, and S. Singhal. 2023. Revisiting the Role of Genetic Variation in Adaptation. The American Naturalist 202:486–502. The University of Chicago Press.

Aini, R. Q., K. Supriya, H. Dunlop, B. Edwards, S. Maas, J. Roberts, A. Summersill, Y. Zheng, S. Brownell, and M. E. Barnes. 2024. Conflict reducing practices in evolution education are associated with increases in evolution acceptance in a large naturalistic study. PLOS ONE 19:e0313490. Public Library of Science.

Akcan, A. T., B. Yıldırım, A. R. Karataş, and M. Yılmaz. 2023. Teachers' views on the effect of STEM education on the labor market. Front Psychol 14:1184730.

Altmann, R. 1894. Die Elementarorganismen und ihre Beziehungen zu den Zellen.

Anderson, D. L., K. M. Fisher, and G. J. Norman. 2002. Development and Evaluation of the Conceptual Inventory of Natural Selection. Journal of Research in Science Teaching 39:952–978.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden, and I. G. Young. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–465. Nature Publishing Group.

Annetta, L. A. 2008. Video Games in Education: Why They Should Be Used and How They Are Being Used. Theory Into Practice 47:229–239. Routledge.

Annetta, L. A., J. Minogue, S. Y. Holmes, and M.-T. Cheng. 2009. Investigating the impact of video games on high school students' engagement and learning about genetics. Computers & Education 53:74–85.

Archila, P. A., S. Restrepo, A. Truscott de Mejía, and J. Molina. 2024. STEM and Non-STEM Misconceptions About Evolution: Findings from 5 Years of Data. Sci & Educ 33:1211–1229.

Ariss, S. S., and S. A. Timmins. 1989. Employee Education and Job Performance: Does Education Matter?: http://dx.doi.org/10.1177/009102608901800101 18:1–9. SAGE PublicationsSage CA: Los Angeles, CA.

Arnqvist, G., D. K. Dowling, P. Eady, L. Gay, T. Tregenza, M. Tuda, and D. J. Hosken. 2010. GENETIC ARCHITECTURE OF METABOLIC RATE: ENVIRONMENT SPECIFIC EPISTASIS BETWEEN MITOCHONDRIAL AND NUCLEAR GENES IN AN INSECT. Evolution 64:3354–3363.

Arnqvist, G., and L. Rowe. 2005. Sexual Conflict.

Arnqvist, G., N. Vellnow, and L. Rowe. 2014. The effect of Epistasis on sexually antagonistic genetic variation. Proceedings of the Royal Society B: Biological Sciences 281. Royal Society.

Arztmann, M., L. Hornstra, J. Jeuring, and L. Kester. 2022. Effects of games in STEM education: a meta-analysis on the moderating role of student background characteristics. , doi: 10.1080/03057267.2022.2057732.

Atkinson, R. D., and M. J. Mayo. 2010. Refueling the U.S. Innovation Economy: Fresh Approaches to Science, Technology,Engineering and Mathematics (STEM) Education.

Ball, C., K.-T. Huang, S. R. Cotten, and R. V. Rikard. 2020. Gaming the SySTEM: The Relationship Between Video Games and the Digital and STEM Divides. Games and Culture 15:501–528. SAGE Publications Inc.

Ballard, J. W. O., and M. Kreitman. 1995. Is mitochondrial DNA a strictly neutral marker? Trends in Ecology & Evolution 10:485–488. Elsevier Current Trends.

Barnes, M. E., R. Q. Aini, J. P. Collins, R. D. P. Dunk, E. A. Holt, J. Jensen, J. R. Klein, T. Misheva, L. S. Nadelson, M. J. Reiss, W. L. Romine, A. Shtulman, A. L. Townley, J. R. Wiles, Y. Zheng, and S. E. Brownell. 2024. Evaluating the current state of evolution acceptance instruments: a research coordination network meeting report. Evo Edu Outreach 17:1.

Barnes, M. E., and S. E. Brownell. 2017. A Call to Use Cultural Competence When Teaching Evolution to Religious College Students: Introducing Religious Cultural Competence in Evolution Education (ReCCEE). CBE Life Sciences Education 16. American Society for Cell Biology.

Barnes, M. E., H. M. Dunlop, G. M. Sinatra, T. M. Hendrix, Y. Zheng, and S. E. Brownell. 2020a. "Accepting Evolution Means You Can't Believe in God": Atheistic Perceptions of Evolution among College Biology Students. LSE 19:ar21. American Society for Cell Biology (lse).

Barnes, M. E., J. Elser, and S. E. Brownell. 2017. Impact of a Short Evolution Module on Students' Perceived Conflict between Religion and Evolution. The American Biology Teacher.

Barnes, M. E., K. Supriya, H. M. Dunlop, T. M. Hendrix, G. M. Sinatra, and S. E. Brownell. 2020b. Relationships between the Religious Backgrounds and Evolution Acceptance of Black and Hispanic Biology Students. LSE 19:ar59. American Society for Cell Biology (lse).

Barnes, M. E., K. Supriya, Y. Zheng, J. A. Roberts, and S. E. Brownell. 2021. A New Measure of Students' Perceived Conflict between Evolution and Religion (PCoRE) Is a Stronger Predictor of Evolution Acceptance than Understanding or Religiosity. LSE 20:ar42. American Society for Cell Biology (lse).

Barreto, F. S., E. T. Watson, T. G. Lima, C. S. Willett, S. Edmands, W. Li, and R. S. Burton. 2018. Genomic signatures of mitonuclear coevolution across populations of Tigriopus californicus. Nat Ecol Evol 2:1250–1257. Nature Publishing Group.

Barrientos, A., F. Fontanesi, and F. Díaz. 2009a. Evaluation of the Mitochondrial Respiratory Chain and Oxidative Phosphorylation System Using Polarography and Spectrophotometric Enzyme Assays. Current Protocols in Human Genetics 63:19.3.1-19.3.14. John Wiley & Sons, Ltd.

Barrientos, A., F. Fontanesi, and F. Díaz. 2009b. Evaluation of the Mitochondrial Respiratory Chain and Oxidative Phosphorylation System Using Polarography and Spectrophotometric Enzyme Assays. Current Protocols in Human Genetics 63:19.3.1-19.3.14. John Wiley & Sons, Ltd.

Barton, N., A. Etheridge, and A. Véber. 2016. The infinitesimal model. bioRxiv 039768. Cold Spring Harbor Laboratory.

Barton, N. H., and P. D. Keightley. 2002. Understanding quantitative genetic variation. Nature Reviews Genetics 2001 3:1 3:11–21. Nature Publishing Group.

Barton, N. H., and M. Turelli. 1989. Evolutionary Quantitative Genetics: How Little Do We Know? Annual Review of Genetics 23:337–370.

Baumgartner, E., and K. Duncan. 2009. Evolution of Students' Ideas about Natural Selection through a Constructivist Framework. The American Biology Teacher 71:218–227. University of California Press.

Beason-Abmayr, B. 2023. The fictional animal project: A flexible tool for helping students learn physiology. Physiology 38:5732003. American Physiological Society.

Berger, D., E. C. Berg, W. Widegren, G. Arnqvist, and A. A. Maklakov. 2014a. Multivariate intralocus sexual conflict in seed beetles. Evolution 68:3457–3469.

Berger, D., E. C. Berg, W. Widegren, G. Arnqvist, and A. A. Maklakov. 2014b. Multivariate intralocus sexual conflict in seed beetles. Evolution 68:3457–3469. John Wiley & Sons, Ltd (10.1111).

Biot-Pelletier, D., S. Bettinazzi, I. Gagnon-Arsenault, A. K. Dubé, C. Bédard, T. H. M. Nguyen, H. L. Fiumera, S. Breton, and C. R. Landry. 2023. Evolutionary Trajectories are Contingent on Mitonuclear Interactions. Molecular Biology and Evolution 40:msad061.

Bishop, B. A., and C. W. Anderson. 1990. Student conceptions of natural selection and its role in evolution. Journal of Research in Science Teaching 27:415–427. John Wiley & Sons, Ltd.

Blier, P., M. Passamonti, N. Pichaud, A. E. Harada, T. M. Healy, and R. S. Burton. 2019. Variation in Thermal Tolerance and Its Relationship to Mitochondrial Function Across Populations of Tigriopus californicus. Frontiers in Physiology | www.frontiersin.org 1:213.

Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends in Ecology & Evolution 24:280–288.

Brisson, D. 2018. Negative Frequency-Dependent Selection Is Frequently Confounding. Front. Ecol. Evol. 6. Frontiers.

Brooks, M., K. Kristensen, K. van Benthem, A. Magnusson, C. Berg, A. Nielsen, H. Skaug, M. Mächler, and B. Bolker. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal.

Buchinger, T. J., and W. Li. 2023. Chemical communication and its role in sexual selection across Animalia. Communications Biology 6:1178. Nature Publishing Group UK London.

Buckberry, S., and K. Burke da Silva. 2012. Evolution: Improving the Understanding of Undergraduate Biology Students with an Active Pedagogical Approach. Evo Edu Outreach 5:266–273. BioMed Central.

Bürger, R. 2005. A multilocus analysis of intraspecific competition and stabilizing selection on a quantitative trait. Journal of Mathematical Biology 50:355–396. Springer.

Burton, R. S. 1990a. Hybrid breakdown in developmental time in the copepod Tigriopus californicus. Evolution 44:1814–1822. Evolution.

Burton, R. S. 1990b. Hybrid breakdown in developmental time in the copepod Tigriopus californicus. Evolution 44:1814–1822. Evolution.

Butler, D., B. Cullis, and A. Gilmour. 2007. Asreml-R: an R package for mixed models using residual maximum likelihood.

Camfield, E. K., L. Beaster-Jones, A. D. Miller, and K. M. Land. 2020. Using Writing in Science Class to Understand and Activate Student Engagement and Self-Efficacy. Pp. 89–105 *in* Active Learning in College Science. Springer International Publishing.

Catton, K. B., A. Galang, and A. T. Bulk. 2016. Disruption in Large Classes during Active Learning Sessions.

Cavanagh, A. J., O. R. Aragón, X. Chen, B. A. Couch, M. F. Durham, A. Bobrownicki, D. I. Hanauer, and M. J. Graham. 2016. Student Buy-In to Active Learning in a College Science Course. CBE—Life Sciences Education 15:ar76. American Society for Cell Biology.

Checa-Romero, M., and J. M. Gimenez-Lozano. 2025. Video games and metacognition in the classroom for the development of 21st century skills: a systematic review. Front. Educ. 9:1485098. Frontiers.

Clark, A. C., and J. V. Ernst. 2009. Gaming in technology education: the study of gaming can teach life skills for the twenty-first century that employers want … these include analytical thinking, team building, multitasking, and problem solving under duress. The Technology Teacher 68:21–27. International Technology Education Association.

Clarke-Midura, J., D. S. Pope, S. Maruca, J. K. Abraham, and E. Meir. 2018. Iterative design of a simulation-based module for teaching evolution by natural selection. Evo Edu Outreach 11:4.

Clo, J., J. Ronfort, and D. Abu Awad. 2020. Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. Journal of Evolutionary Biology 33:1203–1215.

Coburn, K., K. Troy, C. A. Busch, N. Barber-Choi, K. M. Bonney, B. Couch, M. E. García-Ojeda, R. Hutto, L. Famble, M. Flagg, T. Gladding, A. Kowalkowski, C. Landaverde, S. M. Lo, K. MacLeod, B. Mbogo, T. Misheva, A. Trinh, R. Vides, E. Wieboldt, C. Gormally, and J. Maloy. 2025. Cisnormative Language and Erasure of Trans* and Genderqueer Student Representation in Biology Education Research. LSE 24:ar3. American Society for Cell Biology (lse).

Connallon, T., and S. F. Chenoweth. 2019. Dominance reversals and the maintenance of genetic variation for fitness. PLoS Biol 17:e3000118.

Craig Stillwell, R., and C. W. Fox. 2009. Geographic variation in body size, sexual size dimorphism and fitness components of a seed beetle: local adaptation versus phenotypic plasticity. Oikos 118:703–712.

Crick, F. 1970. Central Dogma of Molecular Biology. Nature 227:561–563. Nature Publishing Group.

Crudgington, H. S., and M. T. Siva-Jothy. 2000. Genital damage, kicking and early death. Nature 407:855–856. Nature Publishing Group.

Darwin, C. 1871. The Descent of Man, and Selection in Relation to Sex.

Davidesco, I., and C. Milne. 2019. Implementing Cognitive Science and Discipline-Based Education Research in the Undergraduate Science Classroom. CBE—Life Sciences Education 18:es4. NLM (Medline).

de Rutte, J., S. Myruski, E. Davis, A. Findley, and T. A. Dennis-Tiwary. 2025. A randomized clinical trial investigating the clinical impact of a game-based digital therapeutic for social anxiety disorder. J Anxiety Disord 111:103000.

Deslauriers, L., L. S. McCarty, K. Miller, K. Callaghan, and G. Kestin. 2019. Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. Proceedings of the National Academy of Sciences 116:19251–19257. National Academy of Sciences.

Dobzhansky, T. 1937. Genetic Nature of Species Differences. The American Naturalist 71:404–420.

Dobzhansky, T. 1973. Nothing in biology makes sense except in the light of evolution. American Biology Teacher 35:125–129.

Dowling, D. K., and R. E. Adrian. 2019. Challenges and Prospects for Testing the Mother's Curse Hypothesis. Integrative and Comparative Biology 59:875–889.

Dowling, D. K., U. Friberg, and J. Lindell. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. Trends in Ecology & Evolution 23:546–554. Elsevier Current Trends.

Dowling, D. K., A. L. Nowostawski, and G. Arnqvist. 2007. Effects of cytoplasmic genes on sperm viability and sperm morphology in a seed beetle: Implications for sperm competition theory? Journal of Evolutionary Biology 20:358–368.

Dowling, D. K., and J. N. Wolff. 2023. Evolutionary genetics of the mitochondrial genome: insights from Drosophila. Genetics 224:iyad036.

Driessen, E. P., J. K. Knight, M. K. Smith, and C. J. Ballen. 2020. Demystifying the Meaning of Active Learning in Postsecondary Biology Education. CBE—Life Sciences Education 19:ar52. American Society for Cell Biology.

Edmands, S., and R. S. Burton. 1999a. CYTOCHROME C OXIDASE ACTIVITY IN INTERPOPULATION HYBRIDS OF A MARINE COPEPOD: A TEST FOR NUCLEAR-NUCLEAR OR NUCLEAR-CYTOPLASMIC COADAPTATION. Evolution 53:1972–1978. John Wiley & Sons, Ltd.

Edmands, S., and R. S. Burton. 1999b. CYTOCHROME C OXIDASE ACTIVITY IN INTERPOPULATION HYBRIDS OF A MARINE COPEPOD: A TEST FOR NUCLEAR-NUCLEAR OR NUCLEAR-CYTOPLASMIC COADAPTATION. Evolution 53:1972–1978. John Wiley & Sons, Ltd.

Eslamieh, M., A. Williford, and E. Betrán. 2017. Few Nuclear-Encoded Mitochondrial Gene Duplicates Contribute to Male Germline-Specific Functions in Humans. Genome Biol Evol 9:2782–2790.

Fairbairn, D. J., and D. A. Roff. 2006. The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. Heredity 97:319–328. Nature Publishing Group.

Falconer, D. S. 1981. Introduction to quantitative genetics.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford.

Fordham, J., R. Ratan, K.-T. Huang, and K. Silva. 2020. Stereotype Threat in a Video Game Context and Its Influence on Perceptions of Science, Technology, Engineering, and Mathematics (STEM): Avatar-Induced Active Self-Concept as a Possible Mitigator. American Behavioral Scientist 64:900–926. SAGE Publications Inc.

Fox, C. W. 1993a. Multiple Mating, Lifetime Fecundity and Female Mortality of the Bruchid Beetle, Callosobruchus maculatus (Coleoptera: Bruchidae). Functional Ecology 7:203. JSTOR.

Fox, C. W. 1993b. The influence of maternal age and mating frequency on egg size and offspring performance in Callosobruchus maculatus (Coleoptera: Bruchidae). Oecologia 96:139–146.

Fox, C. W., M. L. Bush, D. A. Roff, and W. G. Wallin. 2004. Evolutionary genetics of lifespan and mortality rates in two populations of the seed beetle, Callosobruchus maculatus. Heredity 92:170–181.

Franceschini, S., S. Bertoni, M. Lulli, T. Pievani, and A. Facoetti. 2022. Short-Term Effects of Video-Games on Cognitive Enhancement: the Role of Positive Emotions. J Cogn Enhanc 6:29–46.

Futuyma, D. J. 1998. Wherefore and whither the naturalist? American Naturalist, doi: 10.2307/2463288.

Gallach, M., C. Chandrasekaran, and E. Betrán. 2010. Analyses of Nuclearly Encoded Mitochondrial Genes Suggest Gene Duplication as a Mechanism for Resolving Intralocus Sexually Antagonistic Conflict in Drosophila. Genome Biol Evol 2:835–850.

Gibson, A. K., L. F. Delph, and C. M. Lively. 2017. The two-fold cost of sex: Experimental evidence from a natural system. Evolution Letters 1:6. Wiley-Blackwell.

Gomulkiewicz, R., and R. D. Holt. 1995. When does Evolution by Natural Selection Prevent Extinction? Evolution 49:201–207. Oxford University Press.

Gordo, I., and B. Charlesworth. 2000. The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics 154:1379–1387. Oxford Academic.

Grafen, A. 1990. Biological signals as handicaps. Journal of Theoretical Biology 144:517–546.

Grafen, A. 2018. The left hand side of the Fundamental Theorem of Natural Selection. Journal of Theoretical Biology 456:175–189. Academic Press.

Gregory, T. R. 2009. Understanding Natural Selection: Essential Concepts and Common Misconceptions. Evolution: Education and Outreach 2:156–175. BioMed Central Ltd.

Grieshop, K., and G. Arnqvist. 2018. Sex-specific dominance reversal of genetic variation for fitness. PLOS Biology 16:e2006810. Public Library of Science.

Guevara, E., S. Gopalan, D. J. Massey, M. Adegboyega, W. Zhou, A. Solis, A. D. Anaya, S. E. Churchill, J. Feldblum, and R. R. Lawler. 2023. Getting it right: Teaching undergraduate biology to undermine racial essentialism. Biology Methods & Protocols 8:bpad032.

Gui, Y., Z. Cai, Y. Yang, L. Kong, X. Fan, and R. H. Tai. 2023. Effectiveness of digital educational game and game design in STEM learning: a meta-analytic review. IJ STEM Ed 10:36.

Hamilton, W. D., and M. Zuk. 1982. Heritable True Fitness and Bright Birds: A Role for Parasites? Science 218:384–387. American Association for the Advancement of Science.

Haskel-Ittah, M., R. G. Duncan, and A. Yarden. 2020. Students' Understanding of the Dynamic Nature of Genetics: Characterizing Undergraduates' Explanations for Interaction between Genetics and Environment. LSE 19:ar37. American Society for Cell Biology (lse).

Havird, J. C., E. S. Forsythe, A. M. Williams, J. H. Werren, D. K. Dowling, and D. B. Sloan. 2019. Selfish Mitonuclear Conflict. Cell Press.

Havird, J. C., and H. J. McConie. 2019. Sexually Antagonistic Mitonuclear Coevolution in Duplicate Oxidative Phosphorylation Genes. Integrative and Comparative Biology 59:864–874.

Henderson, C. R. 1976. A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. Biometrics 32:69. JSTOR.

Henderson, C. R. 1950. Estimation of Genetic Parameters (Abstract). Ann. Math. Slat. 21.

Henderson, C. R. 1986. Estimation of Variances in Animal Model and Reduced Animal Model for Single Traits and Single Records. Journal of Dairy Science 69:1394–1402. Elsevier.

Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid. 2018. The Contemporary Evolution of Fitness. Annual Review of Ecology, Evolution, and Systematics, doi: 10.1146/annurev-ecolsys-110617-062358.

Herrero, D., H. Del-Castillo, D. Herrero, and D. Castillo. 2013. Can a failed simulation game be useful for teaching about reality that it aims to simulate? Learning about Evolution and Natural Selection through Spore video game CAN A FAILED SIMULATION GAME BE USEFUL FOR TEACHING ABOUT REALITY THAT IT AIMS TO SIMULATE? LEARNING EVOLUTION AND NATURAL SELECTION THROUGH SPORE VIDEO GAME IN A BIOLOGY CLASS.

Hill, G. E. 2020. Mitonuclear Compensatory Coevolution. Elsevier Ltd.

Hill, G. E. 2015. Mitonuclear ecology. Molecular Biology and Evolution, doi: 10.1093/molbev/msv104.

Hill, G. E. 2018. Mitonuclear Mate Choice: A Missing Component of Sexual Selection Theory? BioEssays 40:1700191. John Wiley & Sons, Ltd.

Hill, G. E. 2017. The mitonuclear compatibility species concept. American Ornithological Society.

Hill, G. E., J. C. Havird, D. B. Sloan, R. S. Burton, C. Greening, and D. K. Dowling. 2019. Assessing the fitness consequences of mitonuclear interactions in natural populations. Biological Reviews, doi: 10.1111/brv.12493.

Hill, G. E., and J. D. Johnson. 2013. The mitonuclear compatibility hypothesis of sexual selection. Proceedings of the Royal Society B: Biological Sciences 280. The Royal Society.

Hughes, K. A., and M. H. Burleson. 2000. Evolutionary Causes of Genetic Variation in Fertility and other Fitness Components. Pp. 7–33 *in* J. L. Rodgers, D. C. Rowe, and W. B. Miller, eds. Genetic Influences on Human Fertility and Sexuality: Theoretical and Empirical Contributions from the Biological and Behavioral Sciences. Springer US, Boston, MA.

Ifa, A., and I. Guetat. 2018. Does public expenditure on education promote Tunisian and Moroccan GDP per capita? ARDL approach. Journal of Finance and Data Science 4:234–246. KeAi Communications Co.

Johansson, B. G., and T. M. Jones. 2007. The role of chemical communication in mate choice. Biol Rev Camb Philos Soc 82:265–289.

Joiner, R., J. Iacovides, M. Owen, C. Gavin, S. Clibbery, J. Darling, and B. Drew. 2011. Digital Games, Gender and Learning in Engineering: Do Females Benefit as Much as Males? Journal of Science Education and Technology 20:178–185. Springer.

Kalinowski, S. T., M. J. Leonard, and M. L. Taper. 2016. Development and Validation of the Conceptual Assessment of Natural Selection (CANS). CBE Life Sciences Education 15. American Society for Cell Biology.

Kaufman, H. G. 1978. Continuing education and job performance: A longitudinal study. Journal of Applied Psychology 63:248–251.

Kaufmann, P., J. M. Howie, and E. Immonen. 2023. Sexually antagonistic selection maintains genetic variance when sexual dimorphism evolves. Proceedings of the Royal Society B: Biological Sciences 290:20222484. Royal Society.

Kaufmann, P., M. E. Wolak, A. Husby, and E. Immonen. 2021. Rapid evolution of sexual size dimorphism facilitated by Y-linked genetic variance. Nat Ecol Evol 5:1394–1402. Nature Publishing Group.

Kawecki, T. J. 1995. Expression of genetic and environmental variation for life history characters on the usual and novel hosts in Callosobruchus maculatus (Coleoptera:Bruchidae). Heredity 75:70–76. Nature Publishing Group.

Kazancioğlu, E., and G. Arnqvist. 2014. The maintenance of mitochondrial genetic variation by negative frequency-dependent selection. Ecology Letters 17:22–27. John Wiley & Sons, Ltd.

Keaney, T. A., H. W. S. Wong, D. K. Dowling, T. M. Jones, and L. Holman. 2020. Mother's curse and indirect genetic effects: Do males matter to mitochondrial genome evolution? Journal of Evolutionary Biology 33:189–201. Blackwell Publishing Ltd.

Kebe, K. K., N. Alvarez, M. Tuda, O. Arnqvist, C. W. Fox, M. Semb ene, and A. ı Esp ındola. 2017. Global phylogeography of the insect pest Callosobruchus maculatus (Coleoptera: Bruchinae) relates to the history of its main host, Vigna unguiculata. Journal of Biogeography 44:2515–2526.

Kent, J. W., T. D. Dyer, and J. Blangero. 2005. Estimating the additive genetic effect of the X chromosome. Genet Epidemiol 29:377–388.

Koch, E. L., S. H. Sbilordo, and F. Guillaume. 2020. Genetic variance in fitness and its cross-sex covariance predict adaptation during experimental evolution. Evolution 74:2725–2740.

Korfmacher, I. M., M. Hammann, and C. Konnemann. 2024. The role of consensus messaging in teaching evolution. International Journal of Science Education 0:1–20. Routledge.

Krumbeck, Y., G. W. A. Constable, and T. Rogers. 2020. Fitness differences suppress the number of mating types in evolving isogamous species. R Soc Open Sci 7:192126.

Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the "animal model". Philos Trans R Soc Lond B Biol Sci 359:873–890.

Lande, R., and S. J. Arnold. 1983. The Measurement of Selection on Correlated Characters. Evolution 37:1210–1226. [Society for the Study of Evolution, Wiley].

Lane, N. 2015. Life force: why energy shapes evolution. The Biochemist 37:6–11.

Lane, N. 2011. The costs of breathing. Science 334:184–185. American Association for the Advancement of Science.

Lane, N., and W. Martin. 2010. The energetics of genome complexity. Nature 467:929–934. Nature Publishing Group.

Lane, N., and W. F. Martin. 2015. Eukaryotes really are special, and mitochondria are why. Proceedings of the National Academy of Sciences 112:E4823–E4823. Proceedings of the National Academy of Sciences.

Lavalle, A., M. A. Teruel, A. Maté, and J. Trujillo. 2025. Study of gender perspective in STEM degrees and its relationship with video games. Entertainment Computing 52:100889.

Leeflang, H. L., S. V. Dongen, and P. Helsen. 2021. Mother's curse on conservation: assessing the role of mtDNA in sex-specific survival differences in ex-situ breeding programs. Animal Conservation, doi: 10.1111/ACV.12740. John Wiley & Sons, Ltd.

Leith, A. P., R. A. Ratan, and D. Y. Wohn. 2016. The (De-)evolution of Evolution Games: A Content Analysis of the Representation of Evolution Through Natural Selection in Digital Games. Journal of Science Education and Technology 25:655–664.

Lenhart, A. J. E. A. R. C. J. 2008. Teens, Video Games, and Civics: Teens' Gaming Experiences Are Diverse and Include Significant Social Interaction and Civic Engagement. Pew Internet & American Life Project. Pew Internet & American Life Project. 1615 L Street NW Suite 700, Washington, DC 20036. Tel: 202-419-4500; Fax: 202-419-4505; Web site: http://pewinternet.org.

Lenth, R. V., B. Banfai, B. Bolker, P. Buerkner, I. Giné-Vázquez, M. Herve, M. Jung, J. Love, F. Miguez, J. Piaskowski, H. Riebl, and H. Singmann. 2025. emmeans: Estimated Marginal Means, aka Least-Squares Means.

Lewontin, R. C. 1970. JSTOR: Annual Review of Ecology and Systematics, Vol. 1 (1970), pp. 1-18. Annual Review of Ecology and Systematics.

Lu, H.-R., C.-Y. Mao, L.-J. Zhang, J.-W. He, X.-S. Wang, X.-Y. Zhang, W.-L. Fan, Z.-Z. Huang, L. Zong, C.-H. Cui, F.-M. Wu, X.-L. Wang, Z. Zou, X.-Y. Li, and S.-Q. Ge. 2024. High-quality reference genome of cowpea beetle Callosobruchus maculatus. Sci Data 11:799.

Lucas, K., and J. L. Sherry. 2004. Sex Differences in Video Game Play: Communication Research 31:499–523. Sage PublicationsSage CA: Thousand Oaks, CA.

Lynch, M. 1996. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. Molecular Biology and Evolution 13:209–220.

Martin, R. A., C. R. B. da Silva, M. P. Moore, and S. E. Diamond. 2023. When will a changing climate outpace adaptive evolution? WIREs Climate Change 14:e852.

Maynard-Smith, J. 1978. The evolution of sex. Cambridge University Press Cambridge.

Mayo, M. J. 2009. Video games: A route to large-scale STEM education? National Academies Press.

Messina, F. J. 1990. Alternative life-histories in Callosobruchus maculatus: environmental and genetic bases. Bruchids and legumes. Proc. 2nd international symposium, Okayama, 1989 303–315. Kluwer; Series Entomologica, 46.

Meyer, J. W. 1977. The Effects of Education as an Institution. American Journal of Sociology 83:55–77. University of Chicago Press.

Meyer, K. 1991. Estimating variances and covariances for multivariate animal models by restricted maximum likelihood. Genetics, Selection, Evolution 23:67–83.

Meyer, K. 1989. Estimation of genetic parameters. Pp. 161–167 *in* W. G. Hill and T. F. M. McKay, eds. Evolution and Animal Breeding. CAB International.

Meyer, K. 2008. Likelihood calculations to evaluate experimental designs to estimate genetic variances. Heredity 101:212–221. Nature Publishing Group.

Moore, R. W., and R. L. H. Foy. 1997. The scientific attitude inventory: A revision (SAI II). Journal of Research in Science Teaching 34:327–336.

Morrissey, M. B., D. J. Parker, P. Korsten, J. M. Pemberton, L. E. B. Kruuk, and A. J. Wilson. 2012. The prediction of adaptive evolution: Empirical application of the secondary theorem of selection and comparison to the breeder's equation. Evolution, doi: 10.1111/j.1558-5646.2012.01632.x.

Narayanasamy, V., K. W. Wong, C. C. Fung, and S. Rai. 2006. Distinguishing games and simulation games from simulators. Comput. Entertain. 4:9.

Nelson, C. E. 2008. Teaching evolution (and all of biology) more effectively: Strategies for engagement, critical reasoning, and confronting misconceptions. Integrative and Comparative Biology 48:213–225.

Newall, E., and M. J. Reiss. 2023. Evolution hesitancy: challenges and a way forward for teachers and teacher educators. Evo Edu Outreach 16:5.

Ng, T. W. H., and D. C. Feldman. 2009. How broadly does education contribute to job performance? Personnel Psychology 62:89–134. John Wiley & Sons, Ltd.

Nojima, S., K. Shimomura, H. Honda, I. Yamamoto, and K. Ohsawa. 2007. Contact sex pheromone components of the cowpea weevil, Callosobruchus maculatus. J Chem Ecol 33:923–933.

Osborne, J. W. 2007. Linking Stereotype Threat and Anxiety. https://doi.org/10.1080/01443410601069929 27:135–154. Routledge.

Parker, D. J., K. S. Jaron, Z. Dumas, M. Robinson-Rechavi, and T. Schwander. 2022. X chromosomes show relaxed selection and complete somatic dosage compensation across Timema stick insect species. Journal of Evolutionary Biology 35:1734–1750.

Pennisi, E. 2016. EVOLUTIONARY BIOLOGY. Do genomic conflicts drive evolution? Science (New York, N.Y.).

Petrina, S. 2022. Status and Trends of STEM Education in Canada. Status and Trends of STEM Education in Highly Competitive Countries: Country Reports and International Comparison. ERIC.

Plutzer, E., G. Branch, and A. Reid. 2020. Teaching evolution in U.S. public schools: a continuing challenge. Evolution: Education and Outreach 2020 13:1 13:1–15. BioMed Central.

Poli, D., C. Berenotto, S. Blankenship, B. Piatkowski, G. A. Bader, and M. Poore. 2012. Bringing evolution to a technological generation: A case study with the video game SPORE. University of California Press.

Qi, Y. T., and W. E. Burkholder. 1982. Sex pheromone biology and behavior of the cowpea weevilCallosobruchus maculatus (Coleoptera: Bruchidae). J Chem Ecol 8:527–534.

Queller, D. C. 2017. Fundamental Theorems of Evolution. The American naturalist 189:345–353. University of Chicago PressChicago, IL.

R Core Team. 2025. R: A Language and Environment for Statistical Computing.

Radwan, J. 2008. Maintenance of genetic variation in sexual ornaments: A review of the mechanisms. Genetica.

Rand, D. M., R. A. Haney, and A. J. Fry. 2004. Cytonuclear coevolution: the genomics of cooperation. Trends in Ecology & Evolution 19:645–653. Elsevier.

Rank, N. E., P. Mardulyn, S. J. Heidl, K. T. Roberts, N. A. Zavala, J. T. Smiley, and E. P. Dahlhoff. 2020. Mitonuclear mismatch alters performance and reproductive success in naturally introgressed populations of a montane leaf beetle*. Evolution 74:1724–1740. Society for the Study of Evolution.

Riopel, M., L. Nenciovici, P. Potvin, P. Chastenay, P. Charland, J. B. Sarrasin, and S. Masson. 2020. Impact of serious games on science learning achievement compared with more conventional instruction: an overview and a meta-analysis. https://doi.org/10.1080/03057267.2019.1722420 55:169–214. Routledge.

Roff, D. A. 2009. Modeling Evolution: an introduction to numerical methods. OUP Oxford.

Rousseau, S. E. J. 2021. Male mate choice and selectivity in relation to female mating status in the seed beetle Callosobruchus maculatus. The University of Bergen.

Sayadi, A., E. Immonen, C. Tellgren-Roth, and G. Arnqvist. 2017. The Evolution of Dark Matter in the Mitogenome of Seed Beetles. Genome Biology and Evolution 9:2697–2706.

Shaw, R. G. 2018. From the Past to the Future: Considering the Value and Limits of Evolutionary Prediction. The American Naturalist 193:1–10.

Shaw, R. G. 1991. The comparison of quantitative genetic parameters between populations. Evolution 45:143–151.

Shaw, R. G., and F. H. Shaw. 2013. Quantitative genetic study of the adaptive process. Heredity 112:13–20.

Smiderle, R., S. J. Rigo, L. B. Marques, J. A. Peçanha de Miranda Coelho, and P. A. Jaques. 2020. The impact of gamification on students' learning, engagement and behavior based on their personality traits. Smart Learning Environments 7:3.

Soderberg, P., and F. Price. 2003. An examination of problem-based teaching and learning in population genetics and evolution using EVOLVE, a computer simulation. International Journal of Science Education 25:35–55. Routledge.

Stearns, S. C. 1989. Trade-Offs in Life-History Evolution. Functional Ecology 3:259. JSTOR.

Steele, C. M. 1997. A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist 52:613–629. American Psychological Association, US.

Tekinbas, K. S., and E. Zimmerman. 2003. Rules of Play: Game Design Fundamentals. MIT Press.

The Tree of Sex Consortium. 2014. Tree of Sex: A database of sexual systems. Sci Data 1:140015.

Thongthip, P., K. Intawong, S. Chernbumroong, P. Worragin, Y. Tongpaeng, and K. Puritat. 2024. Design and analysis of personalized serious games for information literacy: catering to introverted and extraverted individuals through game elements. Humanit Soc Sci Commun 11:725. Palgrave.

Truong, J. M., M. E. Barnes, and S. E. Brownell. 2018. Can Six Minutes of Culturally Competent Evolution Education Reduce Students' Level of Perceived Conflict Between Evolution and Religion?

Turelli, M., and N. H. Barton. 2004. Polygenic Variation Maintained by Balancing Selection: Pleiotropy, Sex-Dependent Allelic Effects and G E Interactions. Oxford Academic.

Van Der Giezen, M. 2011. Mitochondria and the Rise of Eukaryotes. BioScience 61:594–601. Oxford Academic.

Van Lieshout, E., K. B. McNamara, and L. W. Simmons. 2014. Why Do Female Callosobruchus maculatus Kick Their Mates? PLOS ONE 9:e95747. Public Library of Science.

Vostinar, A. E., B. Z. Johnson, and K. Connors. 2020. Artificial Life in Game Mods for Intuitive Evolution Education. arXiv. arXiv.

Wais, P. E., M. Arioli, R. Anguera-Singla, and A. Gazzaley. 2021. Virtual reality video game improves high-fidelity memory in older adults. Sci Rep 11:2552. Nature Publishing Group.

Walker, J. D., S. H. Cotner, P. M. Baepler, and M. D. Decker. 2008. A Delicate Balance: Integrating Active Learning into a Large Lecture Course. CBE Life Sci Educ 7:361–367.

Wang, L.-H., B. Chen, G.-J. Hwang, J.-Q. Guan, and Y.-Q. Wang. 2022. Effects of digital game-based STEM education on students' learning achievement: a meta-analysis. International Journal of STEM Education 9:26.

Weaver, R. J., S. Rabinowitz, K. Thueson, and J. C. Havird. 2022. Genomic Signatures of Mitonuclear Coevolution in Mammals. Mol Biol Evol 39:msac233.

White, P. J. T., M. K. Heidemann, and J. J. Smith. 2013. A New Integrative Approach to Evolution Education. BioScience 63:586–594. Oxford Academic.

Wickham, H., W. Chang, L. Henry, T. L. Pedersen, K. Takahashi, C. Wilke, K. Woo, H. Yutani, D. Dunnington, T. van den Brand, Posit, and PBC. 2025. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics.

Wilson, A. J., D. Réale, M. N. Clements, M. M. Morrissey, E. Postma, C. A. Walling, L. E. B. Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. Journal of Animal Ecology 79:13–26.

Wise, D. A. 1975a. Academic achievement and job performance. American Economic Review 65:350–366.

Wise, D. A. 1975b. Academic achievement and job performance. American Economic Review 65:350–366.

Wolak, M. E. 2012. nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. Methods in Ecology and Evolution 3:792–796.

Wolff, J. N., D. M. Tompkins, N. J. Gemmell, and D. K. Dowling. 2016. Mitonuclear interactions, mtDNA-mediated thermal plasticity, and implications for the Trojan Female Technique for pest control. Scientific Reports, doi: 10.1038/srep30016.

Wood, W. B. 2009. Innovations in Teaching Undergraduate Biology and Why We Need Them. Annual Review of Cell and Developmental Biology 25:93–112. Annual Reviews .

Woodin, T., V. C. Carter, and L. Fletcher. 2010. Vision and Change in Biology Undergraduate Education, A Call for Action—Initial Responses. LSE 9:71–73. American Society for Cell Biology (lse).

Xie, Y., M. Fang, and K. Shauman. 2015. STEM Education. Annual Review of Sociology 41:331–357. Annual Reviews Inc.

Yang, J. 2015. Video Games for STEM Learning: How Does It Work?

Yasri, P., and R. Mancy. 2014. Understanding Student Approaches to Learning Evolution in the Context of their Perceptions of the Relationship between Science and Religion. International Journal of Science Education 36:24–45.

Zahavi, A. 1975. Mate selection—A selection for a handicap. Journal of Theoretical Biology 53:205–214.

Zhang, L.-J., L. Wu, C.-Y. Wei, X.-L. Liu, H.-J. Xue, X.-K. Yang, and R.-E. Nie. 2018. The complete mitochondrial genome of the cowpea weevil, Callosobruchus maculates (Coleoptera: Chrysomelidae: Bruchinae) and a related phylogenetic analysis of Chrysomelidae. Taylor & Francis 3:645–647. Taylor & Francis.

Appendices

Chapter 2: Additional Tables

Table 37: Antennation duration by matched or mismatched nuclear type.

'Survey station' represents the table upon which assays were conducted, and 'male age at mating' represents the age (in days) of the male at the time of mating. * <0.1, **<0.05, ***<0.001.

Antennation Duration	Seconds	Standard Error	Z-Value	P
Nuclear Matched (Intercept)	3.9448011	0.09446360	41.7600120	0***
Nuclear Mismatched	0.1097694	0.11845661	0.9266634	0.3541013
Survey Station (Two)	-0.1102527	0.11868686	-0.9289373	0.3529216
Male Age at Mating (Scaled)	-0.0489182	0.05827015	-0.8395071	0.4011848

Table 38: Chase Duration by matched or mismatched nuclear type.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male, and zero

inflation describes the likelihood of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Chase Duration	Seconds	Standard Error	Z-Value	P
Nuclear Matched (Intercept)	0.93841275	0.3887438	2.41396171	0.01578012**
Nuclear Mismatched	0.22193655	0.4562562	0.48642967	0.6266626
Male Mass (Scaled)	-0.01153137	0.2249001	-0.05127331	0.9591077
Survey Station (Two)	0.51191845	0.4580722	1.11754980	0.2637593
Male Age at Mating (Scaled)	-0.22855951	0.2274396	-1.00492409	0.3149334
Zero Inflation (Chase Duration)				
Nuclear Matched (Intercept)	-3.49471298	0.7359261	-4.74872807	2.047e-06***
Nuclear Mismatched	0.66783602	0.9681159	0.68983065	0.4903007

Table 39: Female acceptance until first kick (from the female) by matched or mismatched nuclear type.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Acceptance to Kick	Seconds	Standard Error	Z-Value	P
Nuclear Matched (Intercept)	375.519242	30.12194	12.46663336	1.135217e-35***
Nuclear Mismatched	-1.648112	25.40308	-0.06487843	0.9482708
Male Mass (Scaled)	-35.946714	12.81335	-2.80541143	0.005025239**
Survey Station (Two)	-7.344787	42.69435	-0.17203184	0.8634125
Male Age at Mating (Scaled)	-32.265216	12.82619	-2.51557244	0.01188393**

Table 40: From the start of copulation until first kick (from the female) by matched or mismatched nuclear type.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Copulation to First Kick	Seconds	Standard Error	Z-Value	P
Nuclear Matched (Intercept)	313.37774	27.85026	11.25223703	2.257916e-29***
Nuclear Mismatched	-14.89881	22.08389	-0.67464615	0.4999006
Male Mass (Scaled)	-21.29798	11.15009	-1.91011755	0.05611808*
Survey Station (Two)	-1.11629	39.90909	-0.02797083	0.9776854
Male Age at Mating (Scaled)	-19.38285	11.16279	-1.73637982	0.0824967*

First Kick Until Copulation Ends	Seconds	Standard Error	Z-Value	P
Nuclear Matched (Intercept)	4.85819251	0.16411244	29.6028295	1.374078e-192***
Nuclear Mismatched	0.12076071	0.13752600	0.8780936	0.3798929
Male Mass (Scaled)	0.02259664	0.06935765	0.3257988	0.7445766
Survey Station (Two)	-0.03164401	0.23286683	-0.1358889	0.8919091
Male Age at Mating (Scaled)	0.13191102	0.06929726	1.9035531	0.05696842*

Table 41: From the start of copulation until full separation by matched or mismatched nuclear type.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Table 42: Antennation duration by matched or mismatched mitochondrial type.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Antennation Duration	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	4.02311015	0.10073784	39.9364342	0***
Mitochondria Mismatched	-0.06323780	0.11912048	-0.5308726	0.595507
Survey Station (Two)	-0.09912227	0.11808514	-0.8394136	0.4012373
Male Age at Mating (Scaled)	-0.04511582	0.05905069	-0.7640185	0.4448562

Table 43: Chase duration by matched or mismatched mitochondrial type.

‘Survey station’ represents the table upon which assays were conducted, ‘male age at mating’ represents the age (in days) of the male at the time of mating, male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male, and zero inflation describes the likelihood of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Chase Duration	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	1.077157381	0.3970517	2.712889468	0.006669935**
Mitochondria Mismatched	-0.074834755	0.4541230	-0.164789616	0.8691096
Male Mass (Scaled)	-0.004755252	0.2237725	-0.021250389	0.9830459
Survey Station (Two)	0.535406743	0.4544392	1.178170256	0.2387287
Male Age at Mating (Scaled)	-0.219568281	0.2284601	-0.961079251	0.3365123
Zero Inflation (Chase Duration)				
Mitochondria Matched (Intercept)	-3.163819575	0.6347350	-4.984473546	6.213075e-07***
Mitochondria Mismatched	0.008027357	0.9297771	0.008633636	0.9931114

Table 44: Time elapsed from the start of copulation until first kick by matched or mismatched mitochondrial type.

‘Survey station’ represents the table upon which assays were conducted, ‘male age at mating’ represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Copulation to First Kick	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	313.37774	27.85026	11.25223703	2.257916e-29***
Mitochondria Mismatched	-14.89881	22.08389	-0.67464615	0.4999006
Male Mass (Scaled)	-21.29798	11.15009	-1.91011755	0.05611808*
Survey Station (Two)	-1.11629	39.90909	-0.02797083	0.9776854
Male Age at Mating (Scaled)	-19.38285	11.16279	-1.73637982	0.0824967*

Table 45: Time elapsed from the first time the female kicks the male (post-copulation) until full separation by matched or mismatched mitochondrial type.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

First Kick Until Copulation Ends	Seconds	Standard Error	Z-Value	P
Mitochondria Matched (Intercept)	4.85819251	0.16411244	29.6028295	1.374078e-192***
Mitochondria Mismatched	0.12076071	0.13752600	0.8780936	0.3798929
Male Mass (Scaled)	0.02259664	0.06935765	0.3257988	0.7445766
Survey Station (Two)	-0.03164401	0.23286683	-0.1358889	0.8919091
Male Age at Mating (Scaled)	0.13191102	0.06929726	1.9035531	0.05696842*

Table 46: Antennation duration by mate quality.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating. * <0.1, **<0.05, ***<0.001.

Antennation Duration	Seconds	Standard Error	Z-Value	P
Mitonuclear Matched Mate (Intercept)	3.92933232	0.09538037	41.1964486	0***
Mitonuclear Mismatched Mate	0.14110276	0.11844500	1.1912935	0.2335384
Survey Station (Two)	-0.11588535	0.11876928	-0.9757182	0.3292041
Male Age at Mating (Scaled)	-0.04766169	0.05823674	-0.8184127	0.4131215

Table 47: Chase duration by mate quality.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male, and zero inflation describes the likelihood of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Chase Duration	Seconds	Standard Error	Z-Value	P
Mitonuclear Matched Mate (Intercept)	1.009588418	0.3813409	2.647469411	0.00810967**
Mitonuclear Mismatched Mate	0.076444085	0.4563503	0.167511840	0.8669673
Male Mass (Scaled)	0.006572335	0.2239172	0.029351625	0.9765842
Survey Station (Two)	0.511583118	0.4614288	1.108693560	0.2675624
Male Age at Mating (Scaled)	-0.222056079	0.2264227	-0.980714568	0.3267335
Zero Inflation (Chase Duration)				
Mitonuclear Matched Mate (Intercept)	-19.823795535	3884.8235164	-0.005102882	0.9959285
Mitonuclear Mismatched Mate	17.165033803	3884.8235415	0.004418485	0.9964746

Table 48: Time elapsed from the moment the female stops fleeing the male until the female first kicks the male by mate quality.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Acceptance to Kick	Seconds	Standard Error	Z-Value	P
Mitonuclear Matched Mate (Intercept)	364.30313	29.64865	12.2873447	1.059286e-34***
Mitonuclear Mismatched Mate	25.07331	25.80305	0.9717192	0.3311903
Male Mass (Scaled)	-34.21152	12.90686	-2.6506466	0.008033785**
Survey Station (Two)	-11.36851	41.47908	-0.2740781	0.7840246
Male Age at Mating (Scaled)	-31.82584	12.80460	-2.4855005	0.01293694**

Table 49: Time elapsed from the start of copulation until first kick by mate quality.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Copulation to First Kick	Seconds	Standard Error	Z-Value	P
Mitonuclear Matched Mate (Intercept)	302.674637	27.62440	10.9567850	6.165076e-28***
Mitonuclear Mismatched Mate	11.193495	22.47044	0.4981432	0.6183831
Male Mass (Scaled)	-20.887983	11.26073	-1.8549400	0.06360479*
Survey Station (Two)	-4.808362	39.10725	-0.1229532	0.9021441
Male Age at Mating (Scaled)	-19.149827	11.17884	-1.7130423	0.08670476*

Table 50: Time elapsed from the first time the female kicks the male (post-copulation) until full separation by mitonuclear quality of the mate.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

First Kick Until Copulation Ends	Seconds	Standard Error	Z-Value	P
Mitonuclear Matched Mate (Intercept)	4.87048247	0.16428530	29.6464904	3.763908e-193***
Mitonuclear Mismatched Mate	0.08388641	0.13959327	0.6009345	0.5478836
Male Mass (Scaled)	0.03094956	0.07004868	0.4418293	0.6586127
Survey Station (Two)	-0.02858107	0.23039284	-0.1240537	0.9012728
Male Age at Mating (Scaled)	0.13218871	0.06943307	1.9038293	0.05693242*

Table 51: Time elapsed from the moment the male inserts the aedeagus until full separation by the mitonuclear quality of the mate.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Copulation Duration	Seconds	Standard Error	Z-Value	P
MitoNuclear Matched Mate (Intercept)	477.566410	18.04706	26.4622776	2.635573e-154***
MitoNuclear Mismatched Mate	27.161715	22.39051	1.2130903	0.2250953
Male Mass (Scaled)	-19.690478	11.26044	-1.7486417	0.08035298*
Survey Station (Two)	-15.230948	22.70507	-0.6708169	0.5023371
Male Age at Mating (Scaled)	-7.614182	11.15959	-0.6822994	0.4950496

Table 52: Antennation duration given population-level match or mismatch of mitochondrial and nuclear genomes.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating. * <0.1, **<0.05, ***<0.001.

Antennation Duration	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	4.0246365	0.08916673	45.1360785	0***
Different Overall Type	-0.1158813	0.13239749	-0.8752530	0.3814363
Survey Station (Two)	-0.1015510	0.11797878	-0.8607564	0.3893722
Male Age at Mating (Scaled)	-0.0525272	0.05835163	-0.9001838	0.3680224

Table 53: Chase duration given population-level match or mismatch of mitochondrial and nuclear genomes.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male, and zero inflation describes the likelihood of producing a zero value given the variable. * <0.1, **<0.05, ***<0.001.

Chase Duration	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	1.099817e+00	0.4467076	2.4620505696	0.01381452**
Different Overall Type	6.349686e-01	0.4959437	1.2803241090	0.2004312
Male Mass (Scaled)	5.597821e-05	0.2210410	0.0002532481	0.9997979
Survey Station (Two)	5.669147e-01	0.4510877	1.2567727346	0.2088359
Male Age at Mating (Scaled)	-1.691418e-01	0.1839186	-0.9196554358	0.3577528
Zero Inflation (Chase Duration)				
Same Overall Type (Intercept)	-2.954766e+00	0.5357981	-5.5147005541	3.493743e-08***
Different Overall Type	-7.825842e-01	1.1992618	-0.6525548967	0.5140433

Table 54: Time elapsed from the moment the female stops fleeing the male until the female first kicks the male given population-level matching or mismatching of both the mitochondrial and nuclear genomes.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Acceptance to Kick	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	368.110648	29.48535	12.4845262	9.068149e-36***
Different Overall Type	24.163298	28.07305	0.8607292	0.3893872
Male Mass (Scaled)	-36.167175	12.76906	-2.8324076	0.004619891**
Survey Station (Two)	-6.710255	42.65390	-0.1573187	0.8749937
Male Age at Mating (Scaled)	-31.749638	12.80621	-2.4792382	0.01316633**

Table 55: Time elapsed from the moment the male inserts the aedeagus until the female first kicks the male given population-level matching or mismatching of both the mitochondrial and nuclear genomes.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Copulation to First Kick	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	299.751843	26.92836	11.13145684	8.818236e-29***
Different Overall Type	28.033377	24.50154	1.14414759	0.2525624
Male Mass (Scaled)	-21.901790	11.10897	-1.97154015	0.04866213**
Survey Station (Two)	-2.148722	39.29709	-0.05467891	0.9563943
Male Age at Mating (Scaled)	-18.678590	11.14037	-1.67665855	0.09360927*

Table 56: Time elapsed from the first time the female kicks the male (post-copulation) until full separation of the pair given population-level match or mismatch of mitochondrial and nuclear genomes .

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

First Kick Until Copulation Ends	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	4.92493358	0.15758721	31.25211547	2.089552e-214***
Different Overall Type	-0.06526858	0.15314878	-0.42617758	0.6699785
Male Mass (Scaled)	0.02586595	0.06948536	0.37225039	0.7097064
Survey Station (Two)	-0.01862621	0.22825068	-0.08160418	0.9349615
Male Age at Mating (Scaled)	0.12990070	0.06953815	1.86804946	0.06175518*

Table 57: Time elapsed from the moment the male inserts the aedeagus until full separation by the mitonuclear quality of the mate.

'Survey station' represents the table upon which assays were conducted, 'male age at mating' represents the age (in days) of the male at the time of mating, and male mass (scaled) represents the scaled average of two mass measures (in mg) for the given male. * <0.1, **<0.05, ***<0.001.

Copulation Duration	Seconds	Standard Error	Z-Value	P
Same Overall Type (Intercept)	486.279747	16.87339	28.8193191	1.228595e-182***
Different Overall Type	10.990007	25.10824	0.4377052	0.6616
Male Mass (Scaled)	-20.995512	11.25988	-1.8646308	0.06223317*
Survey Station (Two)	-10.915123	22.61296	-0.4826933	0.6293135
Male Age at Mating (Scaled)	-7.235762	11.22964	-0.6443452	0.5193516

Evolution by Natural Selection Questions (CANS)

(Students are issued a sample of the questions in this category)

Anteaters are mammals that live in South America and eat only ants. Anteaters have several traits that help them catch and eat ants efficiently. Firstly, anteaters have remarkably large claws that allow them to easily rip open ant hills. Anteaters feed by sticking their tongue into tunnels in the ant hills. Their entire head and mouth is adapted for catching ants. Their tongues are 24 inches long and covered with sticky saliva. Anteaters cannot open their mouth, and do not have teeth. Even the stomachs of anteaters are unique: unlike most mammals, anteaters do not secrete acid in their stomachs. None is needed. Ants naturally contain formic acid; ants eaten by anteaters digest in their own acid. Biologists have concluded that anteaters evolved all of these unique traits from ancestors that looked similar to rats.

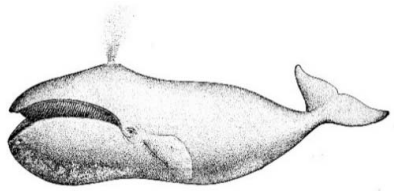
1.Which of the following is the best description of how anteaters evolved long tongues?
a. Anteaters grew long tongues because they needed to reach inside ant hills.
b. Anteaters grew long tongues because they constantly stretched their tongues.
c. Random mutations occurred because anteaters needed to change.
d. Each year, anteaters with the longest tongues were most likely to live.
e. Changes like this depend on many factors, so it is impossible to answer.

3.A female anteater gives birth to a baby. What traits of the mother will the baby inherit?
a. The mother's traits that helped her survive and reproduce.
b. The mother's traits that changed because she used or did not use them.
c. The mother's traits that were changed by the environment during her lifetime.
d. The mother's traits that were determined by genes.
e. The mother's traits determined by genes plus one or more other traits listed above.

5.Anteaters evolved long claws from ancestors that had shorter claws. Think about the first anteater to
have claws as long as modern anteaters. Why did this individual have such long claws?
a. The anteater dug up many anthills, and these efforts affected its claws.
b. The anteater was lucky a genetic accident gave it long claws.
c. The anteater needed long claws to dig up ants, so they developed.
d. The anteater needed long claws to eat, so a mutation changed its DNA.

7.Modern anteaters do not have teeth, but their ancestors did. Which of the following is the best
description of what caused anteaters to lose their teeth?
a. Anteaters did not use their teeth while feeding on ants.
b. Anteaters did not need their teeth to survive and raise their young.
c. Anteaters without teeth had more young than anteaters with teeth.
d. This happened entirely by chance.
e. Changes like this depend on many factors, so it is impossible to answer.

8. An ancestor of modern anteaters has a tongue that is only half as long as modern anteaters. Because
her tongue is relative short, she has to work hard to extend her tongue far enough inside ant hills to
catch ants. How will these efforts affect the tongues of her offspring?
a. Her efforts will probably give her offspring slightly longer tongues.
b. Her efforts will not affect the tongues of her offspring.
c. Growth is affected by many factors; the effects of her actions cannot be predicted.



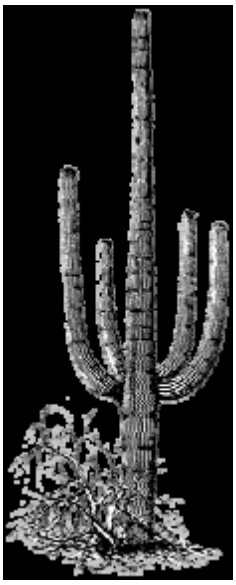
Bowhead whales are the only species of large whales that live their entire life in the icy water of the Arctic Ocean. They have a couple adaptations that help them do this. First, bowhead whales have a thick layer of fat under their skin called blubber that helps keep them warm. The blubber of bowhead whales is
18 inches thick, which is thicker than any other whale. Second,
bowhead whales have a remarkably thick skull. This allows
them to break thick ice in order to get air to breathe. No other whales have such thick skulls. Available
evidence shows that bowhead whales evolved from ancestors that lived in the warm waters of the
Pacific Ocean and did not have either thick blubber or thick skulls.

9.Which of the following is the best description of the role cold water played in the evolution of thick
blubber?
a. It caused mutations that gave whales thicker and thicker blubber.
b. It helped determine which whales each generation survived and which did not.
c. It directly influenced the growth and development of the whales.
d. It made the whales work to stay warm and these efforts caused whales to change.
e. It forced all the whales to change so they could survive and reproduce.

10. What is the best way to describe the evolutionary changes that occurred among the whales while
the species evolved thick skulls?
a. The skull of each whale got a little thicker during its lifetime.
b. Whales with thick skulls reproduced and became more common.
c. The population changed randomly each year.
d. Mutations increased the skull thickness of more and more whales each year.

11. What was most likely true regarding genetic mutations that occurred during the time bowhead
whales were evolving thick blubber?
a. Most of the mutations helped the whales survive in their new environment.
b. The number of mutations occurring in the whale population increased when the whales moved into the icy water of the Arctic Ocean and then decreased when the whales finished adapting to their environment.
c. The mutations occurred because whales needed to adapt in order to survive and reproduce in the icy water of the Arctic Ocean.
d. The mutations were not affected by water temperature or the needs of the whales.

13. Orcas (also known as “killer whales”) hunt bowhead whales. A biologist wants to know how fast
bowhead whales can swim when chased by orcas. She observes female whales being chased by orcas
and measures the maximum speed these whales can swim. What will she most likely observe? Why?
a. All of the whales swim the same speed because they are the same species.
b. There will be notable differences among the whales because each whale has had different amounts of nutrition and exercise during its life.
c. There will be notable differences among the whales because each whale has different genes.
d. Both “b” and “c”.
e. It is impossible to predict given the information provided.



Saguaro cacti live in the scorching hot desserts of Arizona and Mexico where it is common for less than eight inches of rain to fall in a year. Not surprisingly, saguaro cacti have special traits that help them live in this harsh landscape. Like many cacti, saguaro have no leaves. The leaves of the ancestors of saguaro evolved into spines. This prevents water from evaporating from leaves and deters animals from feeding on the saguaro. Photosynthesis occurs within the stem and branches of saguaros. Water loss is further minimized by a waxy covering on the "skin" of saguaro. Beneath the ground, saguaros have an extensive root system—much longer than plants living in wetter climates. These roots are also remarkably shallow. Most of them are only a few inches below the surface of the ground. This enables saguaros to absorb water from rainfall before it evaporates back into the air.

15. Which of the following is the best description of how saguaro cacti evolved to have long roots?
a. Saguaro cacti needed to develop longer roots to survive in the desert.
b. Mutations occurred because the climate of Arizona and Mexico was hot.
c. Saguaros with short roots produced fewer seeds each year than saguaros with longer roots.
d. Each generation of saguaro cacti worked hard to grow roots as long as possible.
e. Changes like this depend on many factors, so it is impossible to answer.

16. During the time period saguaro cacti were evolving to their current form, there were years with very little rain. What likely happened to the saguaro cacti during the driest years?
a. The saguaro cacti managed to obtain the water they needed.
b. Saguaro cacti with the shortest roots died.
c. The saguaro cacti survived with less water than normal.
d. The saguaro cacti grew longer roots.

17. Every individual plant and animal is affected by the environment during its lifetime. For example, a person will become tan if exposed to the sun, and a tree will grow slanted if it lives on a windy ridge.
What role did the responses of individuals to their environment (like these) play in the evolution of

waxy skin among saguaro cacti?
a. Responses like these were the sole reason saguaro cacti evolved waxy skin.
b. Responses like these contributed to saguaro cacti evolving waxy skin.
c. Responses like these played no role in the evolution of waxy skin in saguaro cacti.
d. Responses like these might have played a role in the evolution of cacti—if saguaro respond to intense sunlight or drought by growing waxier skin.

18.A large population of saguaro cacti thrives in a wide valley in Southern Arizona. The valley floor is flat
and there are no gullies, rock outcrops, or variation in the soil. Which of the following statements best
describes similarities and differences among full-grown saguaro cacti living in the valley?
a. The saguaro cacti share all the same traits and are essentially identical to each other.
b. The saguaro cacti share all of the most important traits, and the small differences between them
do not affect how long they live or how well they reproduce.
c. The saguaro cacti are all identical on the inside, but have many differences in appearance.

19. The ancestors of modern saguaro cacti did not have long and sharp spines. Consider the first ancestor of saguaro cacti to grow spines that were as long and sharp as the spines on saguaro cacti living
in Arizona today. Why did this cactus grow such sharp spines?
a. It was fortunate a genetic mistake gave it extra sharp spines.
b. The cactus needed sharper spines to stop animals from eating it.
c. Animals chewing on the cactus caused it to grow sharper spines.
d. Mutations changed the DNA of this cacti because it was injured by an animal.
e. The hot climate caused this change.

20. Saguaro cacti produce fruits that contain thousands of seeds. These seeds are often eaten by birds. When a seed is eaten, it passes through the bird's digestive system unharmed and falls to the ground. If a seed lands in suitable soil, escapes being eaten by mice, and receives enough rainfall during the first years of its life, it may eventually grow into a seedling. Which of the following is the best description of what influences whether saguaro cacti produce seedlings?
a. The production of seedlings is purely a matter of chance.
b. Chance plays a big role, but the characteristics of individual cacti are also important.
c. The production of seedlings is not influenced by chance.

Malaria is a tropical disease caused by a single-celled parasite. People become infected with malaria
when they are bitten by a mosquito that carries the single-celled parasite. Infection with the malaria
parasite causes fever, vomiting, and aches in people. Each year, over 200 million people suffer from
malaria, and one million people die.
A common approach for preventing malaria is killing the mosquitoes that spread the parasite. In the 1940s the insecticide DDT was discovered to be highly effective at killing mosquitoes. DDT was sprayed in many tropical countries and initially killed 99% percent of the mosquitoes in the areas where it was used. This dramatically reduced the number of people contracting malaria in those areas.
Soon after DDT was first used, health workers in Africa discovered that mosquito populations evolved to be resistance to DDT: each year DDT was applied, fewer and fewer of the mosquitoes exposed to DDT died. By the end of the 1940s DDT was no longer effective in some regions and had to be replaced with other insecticides. This happened everywhere DDT was used. Switching insecticides proved to be only a temporary fix; mosquitoes evolved resistance to each insecticide that has been used to combat malaria. This is likely to be a serious health problem for years to come.

21. Which of the following is the best explanation of the process that caused mosquito populations in Africa to become resistant to DDT?
a. The immune systems of mosquitoes exposed to DDT developed resistance. These mosquitoes passed some of this resistance to their offspring so that each generation of mosquitoes became more likely to survive exposure to DDT.
b. Some mosquitoes were fortunate to be naturally able to survive exposure to DDT (even though DDT had never been used in their area before), and passed this ability to their offspring.
c. The mosquitoes became resistant, because if they did not, they would die.
d. DDT caused widespread mutations in the DNA of mosquitoes.

22. Consider a female mosquito that was exposed to DDT during the years a population was evolving
resistance to DDT. She survives and later lays a cluster of eggs. How will her exposure to DDT likely
affect her offspring?
a. Her exposure to DDT will give her offspring increased resistance to DDT.
b. Her exposure to DDT will have no effect on her offspring.
c. The effect on her offspring of her DDT exposure cannot be predicted.

23. What was most likely true regarding the genetic mutations that occurred during the years mosquitoes were evolving resistance to DDT?
a. The number and effect of mutations that occurred was not influenced by DDT.
b. Most of the mutations that occurred helped the mosquitoes survive.

c. The number of mutations occurring in the population increased when DDT was first applied, and then decreased when the mosquitoes finished adapting.
d. The mutations occurred because mosquitoes needed to survive.

24. The chemical deltamethrin (DM) is another insecticide that is used to kill mosquitoes. Not surprisingly, mosquito populations sprayed with DM for several years evolve to become resistant. However, if DM spraying is stopped for a few years, the population will lose its resistance. What is the most likely description of how this occurs?
a. Mosquitoes do not need to be resistant to DM when it is not present.
b. The immune systems of mosquitoes not exposed to DM gradually lose their ability to cope with DM.
c. When DM is not present, mosquitoes that are resistant to DM do not survive as well as mosquitoes that are not resistant to DM.
d. This was purely a random event.

Metacognition

How confident do you feel about your answers to the questions about evolution?

(scale of 1-10, 1 is not confident at all, 10 is extremely confident)

Scientific Attitude Inventory II Questions (Moore and Foy 1997)

(All students are issued all questions in this category)

Likert-type Scale: Strongly Disagree, Disagree, Slightly Disagree, Neutral, Slightly Agree, Agree, Strongly Agree

Questions with a “letter” appended represent categories under which subsequent questions fall—all questions are either ‘negative’ or ‘positive’ and paired with a question designed to be their opposite.

5-A. Progress in science requires public support in this age of science; therefore, the public should be made aware of the nature of science and what it attempts to do. The public can understand science and it ultimately benefits from scientific work.

12. Most people can understand science.
23. People must understand science because it affects their lives.
29. Every citizen should understand science.

5-B. Public understanding of science would contribute nothing to the advancement of science or to human welfare; therefore, the public has no need to understand the nature of science. They cannot understand it and it does not affect them.
6. Only highly trained scientists can understand science.
8. Most people are not able to understand science.
38. Scientific work is only useful to scientists.

6-A. Being a scientist or working in a job requiring scientific knowledge and thinking would be a very interesting and rewarding life's work. I would like to do scientific work.
1. I would enjoy studying science.
27. I would like to work with other scientists to solve scientific problems.
30. I may not make great discoveries, but working in science would be fun.
36. I would like to be a scientist.
40. Working in a science laboratory would be fun.

6-B. Being a scientist or working in a job requiring scientific knowledge and thinking would be dull and uninteresting; it is only for highly intelligent people who are willing to spend most of their time at work. I would not like to do scientific work.
13. The search for scientific knowledge would be boring.
14. Scientific work would be too hard for me.
22. I do not want to be a scientist.
37. Scientists do not have enough time for their families or for fun.
39. Scientists have to study too much.

Demographic Questions

(All students are issued all questions in this category)

1. Gender Identity
    a. Male
    b. Female
    c. Non-binary

    d. Intersex
    e. Transgender
    f. Genderqueer
    g. My gender is: (write-in)
2. Year of Birth (write-in)
3. Ethnicity
    a. Caucasian
    b. African-American
    c. Latino or Hispanic
    d. Asian
    e. Native American
    f. Native Hawaiian or Pacific Islander
    g. Two or More
    h. Other/Unknown
    i. Prefer not to say
4. Year of College
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior
    e. Super Senior (5+ years)
    f. Graduate Student
5. Are you a STEM major? (STEM stands for science, engineering, technology, and mathematics)
    a. Yes
    b. No
6. Are you an international student*?
    a. *We define an international student as a student who is not a U.S. citizen, permanent resident, or refugee, and is on a temporary visa or requires a visa to study in the United States.
        i. Yes
        ii. No
7. Are you a biology major?
    a. Yes
    b. No
8. I most recently studied evolution by natural selection (as part of a course):
    a. This semester
    b. Last semester
    c. Before last semester (while enrolled at Auburn)
    d. Never (while enrolled at Auburn
9. Are you a first-generation college student*?

    a. *We define a student as a first-generation college student if neither of their parents attended college.
        i. Yes
        ii. No.
10. Do you intend to graduate with a STEM degree?
    a. Yes
    b. No
11. Major (write-in)
12. Religion
    a. Christian (Protestant)
    b. Christian (Catholic)
    c. Judaism
    d. Hindu
    e. Islam
    f. Buddhism
    g. Sikhism
    h. Other (write-in)
13. Political Affiliation
    a. Republican
    b. Democrat
    c. Independent
    d. Other (write-in)

“Fun” Questions

Likert-type Scale: Completely Disagree, Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree, Completely Agree

1. I would prefer to learn through video games than through my current class activities.
2. I learned something about evolution by playing this game.
3. I enjoyed this game.
4. Evolution is easier to understand when it’s presented in a game.
5. I understand evolution better after playing this game.
6. Learning through video games makes me more enthusiastic about pursuing a STEM career.

OPEN-ENDED QUESTIONS

1. What did you think of this game?
2. What about this game (if anything) helped you better understand evolution by natural selection?
3. Could it be improved to better teach evolution by natural selection, and if so, how?
4. Would you want this game to be part of a class that taught evolution by natural selection? Why or why not?