Received: 19 October 2021 | Accepted: 14 March 2022

DOI: 10.1002/aps3.11472

APPLICATION ARTICLE

Applications in Plant Sciences

Selection on the gametophyte: Modeling alternation of generations in plants

Elissa S. Sorojsrisom[1,2] | **Benjamin C. Haller**[3] | **Barbara A. Ambrose**[2] | **Deren A. R. Eaton**[1]

[1]Department of Ecology, Evolution, and Environmental Biology, Columbia University, 1200 Amsterdam Ave., Schermerhorn Ext., New York 10027, USA

[2]New York Botanical Garden, 2900 Southern Blvd., Bronx, New York 10458, USA

[3]Department of Computational Biology, Cornell University, Ithaca, New York 14853, USA

Correspondence

Elissa S. Sorojsrisom, New York Botanical Garden, 2900 Southern Blvd., Bronx, New York 10458, USA.
Email: ess2239@columbia.edu

This article is part of the special issue "Methodologies in Gametophyte Biology."

Abstract

Premise: The degree of gametophyte dependence on the sporophyte life stage is a major feature that differentiates the life cycles of land plants, yet the evolutionary consequences of this difference remain poorly understood. Most evolutionary models assume organisms are either haploid or diploid for their entire lifespan, which is not appropriate for simulating plant life cycles. Here, we introduce *shadie* (**S**imulating **Ha**ploid–**Di**ploid **E**volution), a new, simple Python program for implementing simulations with biphasic life cycles and analyzing their results, using SLiM 3 as a simulation back end.

Methods: We implemented evolutionary simulations under three realistic plant life cycle models supported in *shadie*, using either standardized or biologically realistic parameter settings to test how variation in plant life cycles and sexual systems affects patterns of genome diversity.

Results: The dynamics of single beneficial mutation fixation did not vary dramatically between different models, but the patterns of spatial variation did differ, demonstrating that different life histories and model parameters affect both genetic diversity and linkage disequilibrium. The rate of linkage disequilibrium decay away from selected sites varied depending on model parameters such as cloning and selfing rates, through their impact on effective population sizes.

Discussion: Evolutionary simulations are an exciting, underutilized approach in evolutionary research and education. *shadie* can aid plant researchers in developing null hypotheses, examining theory, and designing empirical studies, in order to investigate the role of the gametophyte life stage, and the effects of variation in plant life cycles, on plant genome evolution.

KEYWORDS

alternation of generations, gametophyte, haploid selection, life cycle, simulation, sporophyte

All land plants undergo an alternation of generations between multicellular haploid (gametophyte) and diploid (sporophyte) life stages. The diploid sporophyte phase is the dominant life stage for the majority of land plants, and has historically received the lion's share of attention from the scientific community; however, increasing recognition of the potential for selection acting on the gametophyte phase to drive patterns of gene and genome evolution (Beaudry et al., 2020) is bringing new attention to this understudied life history phase. The major land plant lineages exhibit substantial variation in the extent to which the gametophyte is free-living versus dependent on the sporophyte (Figure 1), as well as in numerous additional life history factors affecting population demography in either life stage. Understanding how such life history differences among lineages have affected their evolution requires examining these processes in the context of a detailed understanding of plant life cycles.

Haploid gametophytes experience selection differently than diploid sporophytes. This is due in part to ploidy differences: gametophytes lack heterozygosity, and thus do not experience dominance effects (Crow and Kimura, 1970), potentially increasing the strength of both positive and purifying selection (Immler, 2019). Indeed, gametophytes have shown the potential to exhibit rapid adaptive evolution (Lamborn et al., 2005), which can also lead to genetic conflicts between gametophyte and sporophyte genomes when the direction of selection is antagonistic (Bell, 1995; Haig and Wilczek, 2006). Despite their very different morphologies, gametophytes and sporophytes tend to express similar sets of genes, with estimates of shared expression ranging from 60% in *Arabidopsis* Heynh. to >90% in bryophytes and ferns (Honys and Twell, 2004; Szövényi et al., 2011; Sigel et al., 2018; Beaudry et al., 2020). Such extensive overlap suggests selection on one life stage may often affect the other through pleiotropy. Population census sizes and, more importantly, effective sizes (N_e), can vary dramatically between gametophyte and sporophyte life stages (Bengtsson and Cronberg, 2009), further affecting the efficacy of selection. This is likely to correlate with differences among life stages in the extent of selfing and cloning (Shaw, 2000; Eppley et al., 2007), the presence of separate sexes, and the interdependence between life stages (e.g., sporophytes that grow dependently from gametophytes). The latter can also be an important mediator of maternal effects (Jesson et al., 2012; Chettoor et al., 2016). Despite the many potential effects of the gametophyte life stage on patterns of genome evolution, we lack a model-based framework for predicting these effects, and studies exploring the impact of alternation of generations on evolutionary processes are few (Hoban et al., 2012; Bessho and Otto, 2022).

Empirical studies of gametophyte evolution have largely focused on angiosperms (Beaudry et al., 2020), which have a highly reduced gametophyte phase, making conclusions from such studies difficult to apply to other plant lineages. For example, in angiosperms the opportunity for selection on the female gametophyte (embryo sac) is predicted to be low compared to the male gametophyte (pollen), because embryos are highly protected and dependent on the sporophyte (Lenormand and Dutheil, 2005). This prediction cannot be extended to other clades, like pteridophytes or bryophytes, where gametophytes are free-living (Stoeckel et al., 2021). This demonstrates the primary challenge for studying the evolutionary consequences of selection on gametophytes: the gametophyte life stages among plant lineages vary in numerous ways that make it difficult to isolate the effects of one aspect in the absence of many others. Evolutionary simulations offer a powerful approach for investigating such complex life histories. For example, one could simulate genes that evolved with or without selection acting on the gametophyte life stage, and in an organism with a free-living versus reduced gametophyte, to investigate their combined effects on measurable population genomic statistics. One could further specify that selection acts antagonistically between the sporophyte and gametophyte generations to model genetic conflict, or test these factors in combination with numerous additional quantitative parameters, such as varying the relative sizes of the gametophyte and sporophyte populations. Through their power to examine elements of plant life cycles in isolation or combination, evolutionary simulations can aid researchers in generating null hypotheses and new predictions for empirical research.

Although it may sound simple in principle, designing evolutionary simulations to accurately model plant life cycles

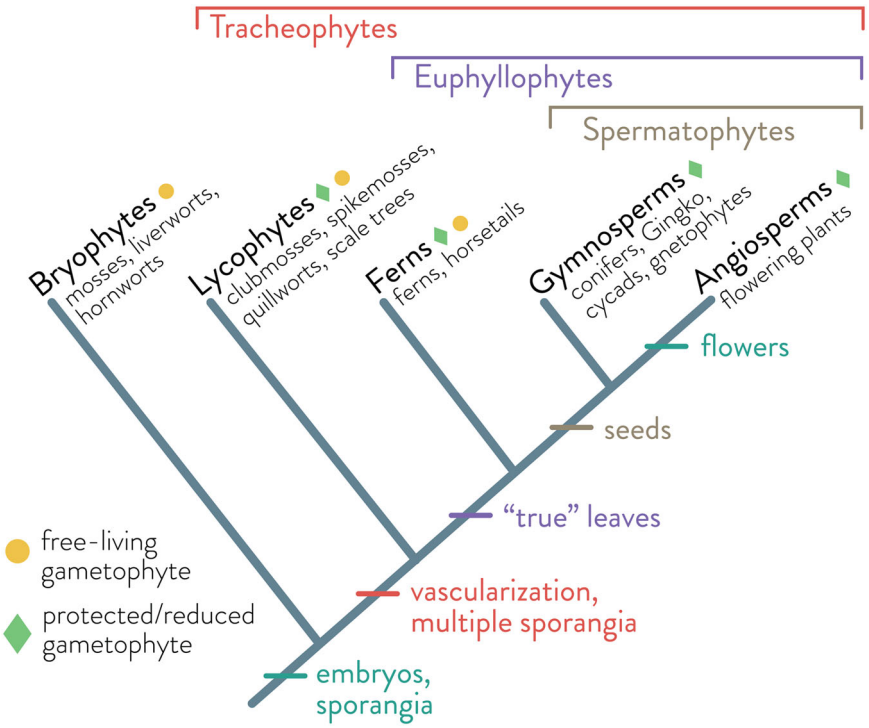

FIGURE 1 Phylogeny of land plants depicting the major plant lineages modeled in *shadie* with key traits. The presence within each lineage of species with free-living versus protected/reduced gametophytes is indicated by a yellow circle versus green diamond, respectively.

requires a nuanced understanding of plant reproduction and the alternation-of-generations process. Here we review the major plant sexual systems and life cycles with the goal of simplifying this variation into a set of codified parameters representing the most significant forms of variation while maintaining a clear link between parameters and biologically relevant structures and processes. We introduce a simple Python program, *shadie* (**S**imulating **Ha**ploid–**Di**ploid **E**volution), for implementing evolutionary simulations of complex, parameterized demographic histories under different biphasic plant life cycles. Finally, we implement a simple set of simulations to explore how variation in plant life cycles and sexual systems affects patterns of genome diversity through impacts on evolutionary dynamics.

Evolutionary simulations

The Wright–Fisher (WF) process models a population evolving in discrete time units (Ewens, 2004), and is commonly used in evolutionary simulations. In its most simple form, it can be used to predict the change in allele frequencies per generation due to drift, based only on the population size, *N*. To do this, it requires a number of simplifying assumptions to exclude the effects of all other evolutionary forces. These assumptions include: (1) no selection; (2) constant population size; (3) panmixia; (4) discrete time and non-overlapping generations; and (5) no new mutations. In practice, this model is typically extended by relaxing one or more of these assumptions to allow additional evolutionary forces to act on the population, such as mutation, selection, population size fluctuations (i.e., demography), migration, and non-random mating, and can even include complex evolutionary scenarios like epistasis, genetic conflicts, biased sex ratios, and many more.

The results of simple evolutionary models can often be solved or approximated with mathematical or statistical equations that are much faster than full evolutionary simulations, as with the use of the coalescent (Kingman, 1982) to model the ancestry of neutrally evolving populations. However, models with non-neutral dynamics (e.g., selection) violate assumptions of most simple models, or become mathematically intractable, and thus require computational simulation. Such simulation programs may still resemble the WF model (e.g., with random mating and non-overlapping generations), or they may be sufficiently different that one might instead call them a non-Wright–Fisher (nonWF) model.

SLiM 3 (Haller and Messer, 2019) is a state-of-the-art program for simulating evolution using WF and nonWF models. By default, SLiM assumes that individuals are diploid with a reduced gametophyte generation, as is typical of most animals and seed plants. It is possible to design simulations in SLiM with a haploid-dominant life cycle, or even simulations of alternation of generations (with both the haploid and diploid stages explicitly simulated with selection). However, doing so can be quite complex and requires intricate knowledge of the Eidos programming language used by SLiM, as well as an intimate understanding of the organismal life history of interest. Below, we describe the sexual systems and life cycles common to the major lineages of land plants, and discuss them in the context of representing their particularities using parameters in our *shadie* Python program, which acts as a simplified front end for SLiM.

Introducing *shadie*

shadie is a Python library designed to: (1) write SLiM scripts to model haploid–diploid plant life cycles, (2) run SLiM simulations, (3) analyze and visualize simulation results, and (4) combine these steps into simple and reproducible Python code intended to be run from interactive Jupyter notebooks (Kluyver et al., 2016). The results of SLiM simulations are stored as tree-sequence files (Kelleher et al., 2018; Haller et al., 2019) containing a record of all relevant genealogical histories and mutations. Analysis of tree-sequence files is performed in *shadie*'s *postsim* module, which uses the *pyslim* (https://github.com/tskit-dev/pyslim [accessed 13 September 2021]) and *tskit* (Kelleher et al., 2018) Python libraries to load and simplify tree sequences, and to calculate population genetic statistics. Visualization tools are integrated throughout *shadie* to aid in designing and verifying simulations. Visualizing chromosomes, genealogies, and statistics is performed using the interactive visualization libraries *toyplot* (Shead, 2014) and *toytree* (Eaton, 2020).

Beginners can use *shadie* to perform complex SLiM simulations with just a few lines of Python code, while advanced users of SLiM can generate a baseline SLiM script with *shadie* and then make their own adjustments in Eidos. Scripts generated by *shadie* can be exported and executed in SLiM's interactive modeling environment, SLiMgui, for further real-time simulation visualization.

In addition to serving research purposes, *shadie* was also designed as a didactic tool for teaching students about plant life cycles and evolution, and its documentation features high-quality scientific illustrations, by E. Sorojsrisom, that elucidate plant biology concepts. Simulations are a powerful tool for teaching students a nuanced understanding of evolution; 21st-century students find technology highly intuitive and may benefit from the similarities between simulations and sim-based computer games (Rieber, 2005). Teaching exercises can be created in Jupyter notebooks and shared through the cloud-based service Binder (Project Jupyter et al., 2018) to allow students to run *shadie* code in a web browser without the need to install any software.

Biological background: Plant mating systems

To sexually reproduce, all plants must eventually make eggs and sperm, but whether this is done by a single individual,

by separate male and female individuals, or something in between, varies among the major plant lineages. The degree of separation between the male and female portions of the life cycle is an important distinguishing feature of a plant life cycle.

Haploid gametophytes form archegonia and/or antheridia, specialized structures in which eggs and sperm, respectively, are produced. In monoicous species, the gametophyte is hermaphroditic and forms both archegonia and antheridia (Figure 2A). In dioicous species, the gametophytes are unisexual: the female gametophyte, or megagametophyte, produces only archegonia, whereas the male gametophyte, or microgametophyte, produces only antheridia (Figure 2B). Both monoicy and dioicy occur in bryophytes and pteridophytes (lycophytes and ferns), whereas all spermatophytes (seed plants) are dioicous.

The egg and sperm produced by gametophytes fuse to form new diploid sporophytes. Sporophytes produce spores and can similarly be distinguished by the type of spores they produce. Homosporous species form haploid spores that are morphologically indistinguishable (Figure 2A–C) and can be associated with monoicy or dioicy. Spores of homosporous species have the potential to develop into female, male, or hermaphroditic gametophytes, and identity may depend on chemical signals from other gametophytes in the surrounding area (Boavida and McCormick, 2010). Heterosporous species form two differentiated types of spores: larger megaspores, which will develop into female gametophytes, and smaller microspores, which will develop into male gametophytes (Figure 2D). All bryophytes are homosporous (Tanurdzic and Banks, 2004), whereas both homospory and heterospory occur in pteridophytes, and only heterospory occurs in spermatophytes.

Monoecy and dioecy (note: although they appear very similar, these terms are distinct from monoicy and dioicy) apply only to spermatophytes, which form specialized

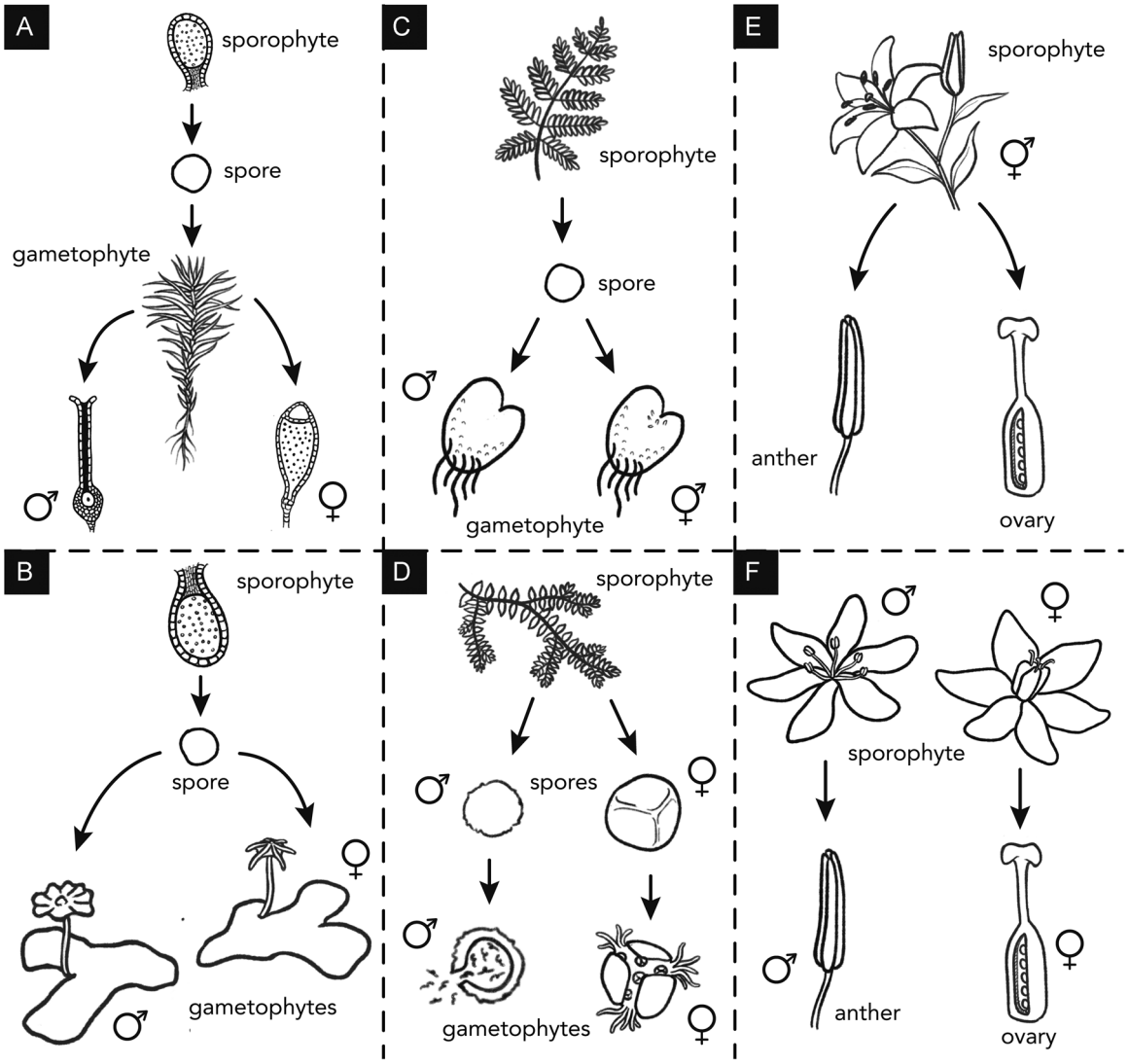

FIGURE 2 Land plant mating systems: (A) monoicy, (B) dioicy, (C) homosporous, (D) heterosporous, (E) monoecious, (F) dioecious. The mating systems of plants are affected by whether the gametophyte, sporophyte, or both life stages have distinct sexes and vary greatly between lineages.

structures such as flowers and cones. In seed plants, sporophytes produce megaspores inside an ovary, where they are retained throughout the formation of the megagametophyte and fertilization. Progeny are released from the mother sporophyte as diploid zygotes (seeds). Microspores mature into microgametophytes, or pollen grains, and are formed inside specialized male structures such as anthers or pollen cones. In monoecious species (Figure 2E), both male and female diploid reproductive structures occur on the same plant (this is distinct from monoicy, because the gametophyte life stages of spermatophytes are unisexual). In dioecious species (Figure 2F), an individual sporophyte can only form reproductive structures that produce ovaries or pollen, but not both (note: in both monoecy and dioecy, the gametophyte life stages are unisexual, so they are both forms of dioicy).

Monoicous gametophytes have the potential to perform gametophytic selfing (i.e., intragametophytic selfing; Figure 3), in which the egg and sperm come from the same haploid gametophyte. This is a severe type of inbreeding, producing a sporophyte that is homozygous at all loci (de Groot et al., 2012). Sporophytic selfing (i.e., intergametophytic selfing; Figure 3) is a less extreme type of inbreeding in which the egg and sperm that produce a zygote are from two different gametophytes formed by the same sporophyte (i.e., two different recombinant products of the same sporophyte). Monoicous species can engage in both gametophytic and sporophytic selfing. By contrast, because their male and female gametophytes are separate, dioicous species can only engage in sporophytic selfing.

The sexual system of an organism should be a core component of evolutionary simulations, because it constrains the way individuals originate and interact, and thus how genes are passed from one generation to the next. For example, a monoecious angiosperm can engage in sporophytic selfing, barring some other pre- or post-zygotic barrier mechanism, whereas a dioecious angiosperm cannot. The ability to configure the parameters of these sexual systems can be important for accurately modeling plant life histories, and also provides a basis for detailed investigations of the evolutionary effects of specific processes such as selfing (e.g., Igic and Busch, 2013; Haig, 2016). Although some sexual systems, such as that of monoecious spermatophytes, fit well with the assumptions of basic evolutionary models like the WF process, other sexual systems, such as those of monoicous bryophytes, homosporous pteridophytes, and

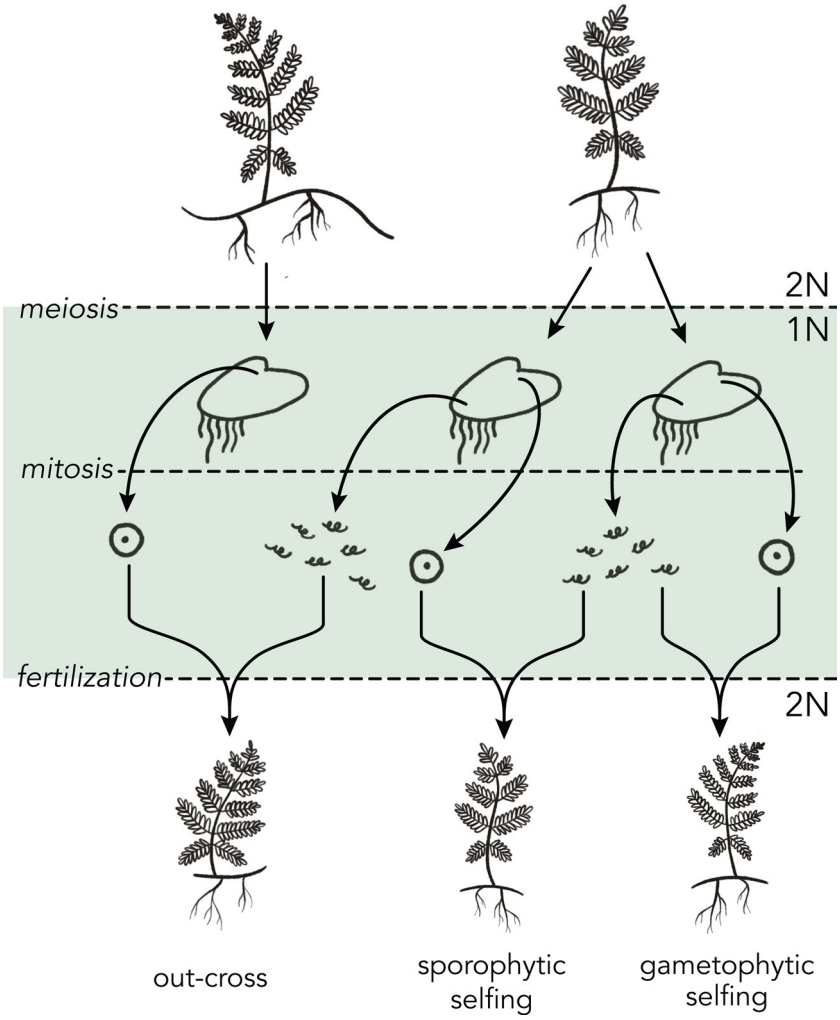

FIGURE 3 Diagrammatic representation of outcrossing versus sporophytic and gametophytic selfing. Eggs are represented by a circle with a dot in the center and sperm are represented by a curlicue. The two sporophyte parents are separate, genomically distinct individuals.

dioecious spermatophytes, deviate considerably and can require many additional parameters to accurately describe.

Biological background: Plant life cycles

This section provides a description of six generalized land plant life cycle categories. Three groups, bryophytes, pteridophytes, and spermatophytes, were chosen based on shared life cycle similarities, with two sexual system variations for each group. There are, of course, exceptions and variations for each of the life cycles presented here. In its most general form, the life cycle of a land plant proceeds as follows: (1) the haploid gametophyte life stage undergoes mitosis to form archegonia and antheridia, where eggs and sperm are produced; (2) an egg is fertilized by a single sperm to form a diploid zygote sporophyte; (3) once mature, the sporophyte creates specialized diploid structures called sporangia, which may be housed within further specialized structures; (4) diploid sporocytes, or spore "mother cells," are created by mitosis within the sporangia; (5) each sporocyte performs two meiotic divisions to create four haploid spores; (6) spores develop into new haploid gametophytes by mitosis. It is important to note that the meiotic divisions in step 5 are the only opportunity for recombination to occur throughout the life cycle. All subsequent haploid structures that arise from a spore (i.e., gametophyte, archegonia, antheridia, gametes) are genetically identical to the original spore.

The bryophyte life cycle consists of a long-lived, free-living gametophyte stage and an ephemeral diploid sporophyte that remains attached and nutritionally dependent on the gametophyte until it dies (Shaw and Beer, 1999). In monoicous bryophytes (Figure 4A) and dioicous bryophytes (Figures 4B and 5A), each archegonium gives rise to a single egg, while each antheridium produces many sperm that, when released, fertilize eggs retained in the archegonia of other gametophytes (in monoicous species, eggs may be fertilized by sperm from the same gametophyte). Fertilization produces a zygote sporophyte, which develops as a superficial structure growing from the gametophyte (in dioicous species, only from female gametophytes). Sporangia develop within the sporophyte and release spores when mature. The spores are dispersed into the surrounding environment and develop into new gametophytes. The sporophyte then dies, but the gametophyte parent continues living, and may reproduce again.

Lycophytes and ferns (known collectively as the paraphyletic group, pteridophytes) have life cycles in which

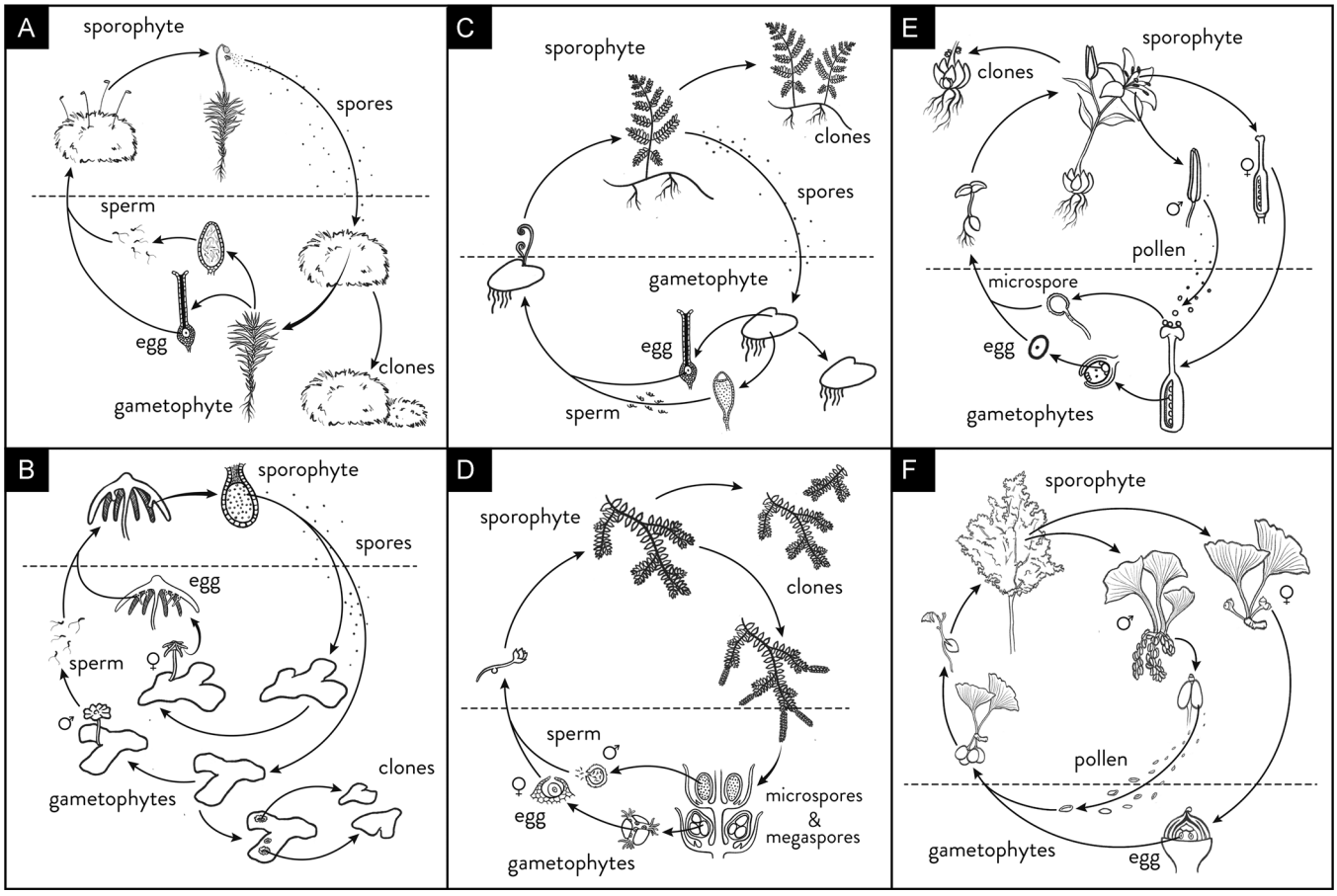

FIGURE 4 Diagrammatic, generalized life cycles of (A) monoicous bryophytes, (B) dioicous bryophytes, (C) homosporous pteridophytes, (D) heterosporous pteridophytes, (E) monoecious spermatophytes, and (F) dioecious spermatophytes.

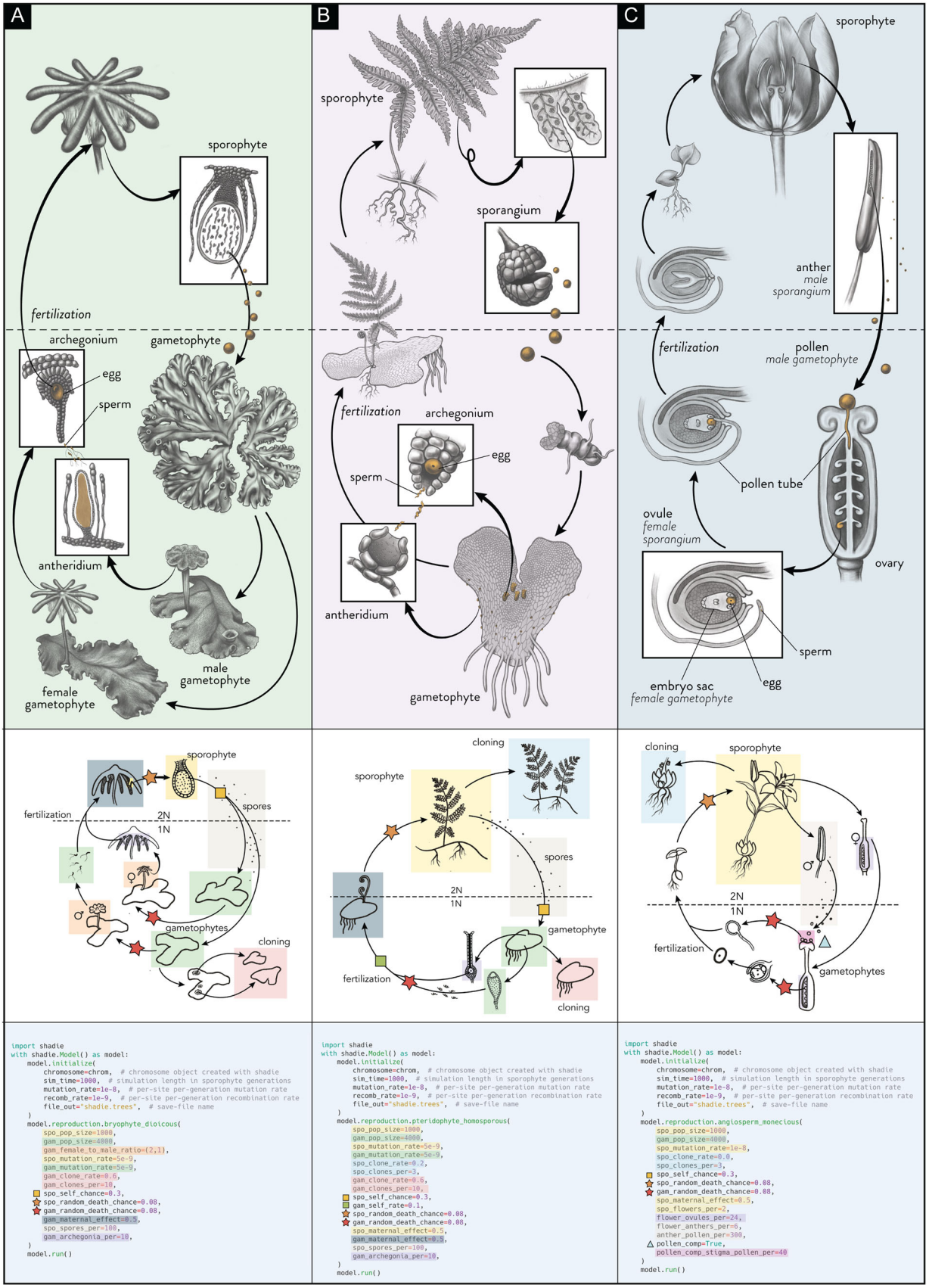

FIGURE 5 Relating model parameters to biological phenomena across different life cycles of major lineages. Realistic life cycle (top), diagrammatic simplification (middle), and user-facing code (bottom) for three of the six default life cycle models available in *shadie*. Models shown are (A) dioicous bryophyte, (B) homosporous pteridophyte, and (C) monoecious angiosperm. In the realistic diagrams (top), haploid tissue involved in reproduction is highlighted in yellow.

the dominant sporophyte life stage produces an ephemeral gametophyte. In homosporous pteridophytes (Figures 4C and 5B), each spore generated by a sporophyte germinates into a free-living gametophyte, or prothallus: a small, flat photosynthetic structure that produces archegonia and/or antheridia. When mature, flagellated sperm are released by the antheridia and require water to swim to an archegonium, where fertilization of an egg creates a new sporophyte zygote. The gametophyte supports the developing sporophyte nutritionally for a time, before it eventually dies away and the sporophyte lives independently. In heterosporous pteridophytes (Figure 4D), sporophytes retain their spores within specialized structures (e.g., strobili and cones in lycophytes, sporocarps in ferns) and gametophytes are highly reduced compared to homosporous gametophytes, consisting of only a few cells. During development and maturation, the micro- and megagametophytes are retained in the spore (called "endosporic") and are thus highly dependent on the sporophyte. When mature, the spores are released while still containing the highly reduced gametophytes, which quickly undergo fertilization to form a new zygote (Kumar, 2001).

In seed plants (gymnosperms and angiosperms), the sporophyte is the dominant life stage. For simplicity, we will use angiosperm terminology for the description of this group; however, the life cycle of gymnosperms (conifers, ginkgos, cycads) is fairly similar, and is also shown for comparison in Figure 4F. In monoecious spermatophytes (Figures 4E and 5C) and dioecious spermatophytes (Figure 4F), the sporophyte produces specialized reproductive structures (e.g., cones in gymnosperms, flowers in angiosperms). In angiosperms, the flowers have a pistil (consisting of a stigma, style, and ovary) and/or stamens (consisting of an anther and filament). Each ovule contained within an ovary produces a single megasporocyte, which will undergo meiosis to produce four recombinant megaspores. Only one megaspore will survive, while the other three degenerate (Boavida and McCormick, 2010). The ovule matures until it forms the megagametophyte, or embryo sac. Microsporocytes are generated within the anther, making it analogous to the microsporangium in other lineages. Each microsporocyte undergoes meiosis to form four recombinant microspores, each of which matures into a single pollen grain, a highly reduced microgametophyte. At maturity, a pollen grain consists of one vegetative cell and two sperm cells, each formed via mitosis of the original microspore. Mature pollen grains are released from the anthers and land on the stigma of a flower. The pollen grain then germinates, growing a pollen tube down the style and into the ovary, where it enters the micropyle and releases its two sperm. One sperm fertilizes the egg to produce a diploid zygote, while the other fertilizes the diploid central cell within the embryo sac to produce the triploid endosperm, which will nourish the developing zygote (Mascarenhas, 1989); the zygote and endosperm together compose the seed. The seed is retained on the sporophyte while the embryo develops within; once mature, the seed is released and will germinate into a new sporophyte. Unlike the life cycles described previously, the reduced gametophytes of seed plants do not give rise to multiple archegonial or antheridial structures. Rather, each gametophyte produces just a single egg or a single sperm (two identical sperm packaged together, in fact) that will engage in sexual reproduction.

Although the many complex details of these plant life cycles can seem trivial, explicit quantitative modeling of such differences among plant lineages may lead to important insights into how selection has differently shaped their evolutionary histories. For example, parameters defining the number of antheridia, archegonia, and gametes produced by each gametophyte describe the potential for a selected mutation to be passed on, in the next generation, to one versus many sporophytes. Similarly, selection can be defined to act only on the phenotype of one life stage (as opposed to both) to represent that a reduced phase, such as the sporophyte in bryophytes, does not interact strongly with the environment, or to represent different kinds of selection pressures, such as sperm competition in spermatophytes (Peters and Weis, 2018). Such settings will affect the strength of selection through their effects on dominance (affecting haploid versus diploid stages) and effective population sizes (passing genes on to many versus few offspring). The outcomes of these processes can be studied through their effects on population genetic statistics calculated on simulated genomes.

METHODS

shadie life cycle model parameters

We developed six models in *shadie* to produce scripts for performing evolutionary simulations in SLiM (using its nonWF model type); these models represent the life cycles of monoicous and dioicous bryophytes, homosporous and heterosporous pteridophytes, and monoecious and dioecious spermatophytes. There are, of course, many more variations in the life cycles of specific plant species; these can be modeled with adjustments to the default models presented here. Examples of the *shadie* Python syntax for writing scripts to implement three of the six life cycle models are shown in Figure 5, paired with illustrations to demonstrate the connections between parameter values and life cycle processes. This article is primarily focused on describing *shadie* functionality from a biological perspective; a publication describing further technical details of the life cycle models and their back-end implementation in SLiM by *shadie* is forthcoming.

To investigate the influence of lineage-specific life cycle processes on evolutionary outcomes, we first simulated genomes under parameter settings where processes shared by multiple lineage types were set to standardized values (Table 1). Examples of this include

TABLE 1 Standardized and realistic parameter settings for simulations.[a]

	Wright–Fisher	Bryophyte: monoicous		Bryophyte: dioicous		Pteridophyte: heterosporous	
Simulation type	**Standard**	**Standard**	**Realistic**	**Standard**	**Realistic**	**Standard**	**Realistic**
sim_time	10,000	10,000	10,000	10,000	10,000	10,000	10,000
recomb_rate	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}	1×10^{-7}
spo_pop_size	3000	3000	3000	3000	3000	3000	3000
gam_pop_size	3000	3000	3000	3000	3000	3000	3000
spo_mutation_rate	1×10^{-6}	5×10^{-7}	5×10^{-7}	5×10^{-7}	5×10^{-7}	5×10^{-7}	5×10^{-7}
gam_mutation_rate	N/A	5×10^{-7}	5×10^{-7}	5×10^{-7}	5×10^{-7}	5×10^{-7}	5×10^{-7}
spo_fem_ratio	N/A	N/A	N/A	N/A	N/A	N/A	N/A
gam_fem-ratio	N/A	N/A	N/A	0.5	0.66666	0.5	0.5
spo_clone_rate	N/A	N/A	N/A	N/A	N/A	0.0	0.1
spo_clones_per	N/A	N/A	N/A	N/A	N/A	0.0	3
spo_self_rate_per_egg	N/A	0.0	0.44	0.0	0.22	0.0	0.1
gam_clone_rate	N/A	0.0	0.8	0.0	0.8	N/A	N/A
gam_clones_per	N/A	0.0	10	0.0	10	N/A	N/A
gam_self_rate_per_egg	N/A	0.0	0.44	N/A	N/A	N/A	N/A
spo_maternal_effect	N/A	N/A	N/A	N/A	N/A	0.0	0.0
gam_maternal_effect	N/A	0.0	0.0	0.0	0.0	N/A	N/A
spo_random_death_chance	N/A	0.0	0.0	0.0	0.0	0.0	0.0
gam_random_death_chance	N/A	0.0	0.0	0.0	0.0	0.0	0.0
spo_spores_per	N/A	10	10	10	10	10	10
gam_archegonia_per	N/A	10	10	10	10	10	10

Note: N/A = the parameter is not applicable to the specific life cycle model.

[a]Selection coefficient, $s = 0.5$; dominance coefficient, $h = 0.5$ for all simulations. The results are reported in Figure 6.

setting identical values for the gametophytic selfing rate in monoicous species, or for the female-to-male ratio in dioicous species. These simulations are intended to highlight the effects of the presence versus absence of specific life cycle attributes in each lineage, as opposed to quantitative differences among them. A full description of each model parameter can be referenced in Appendix S1. (For this article, we excluded maternal effects because we observed nonlinear dynamics emerging that will require further study, and for simplicity, we also excluded the random chance of death parameter.)

Lineages are likely to exhibit even more dramatic evolutionary differences in simulations using realistic life cycle parameters than in our standardized parameter simulations. We thus also simulated genomes under parameter settings (Table 1) informed by empirical estimates to investigate how more realistic values may change patterns of genome evolution. To provide reasonable values for each life cycle process, we compiled a table of values drawn from observations or measurements in the literature, when available, or estimated them to the best of our ability based on natural history knowledge (Appendix S2). Some parameters exhibit substantial variation among lineages. For example, the number of archegonia per gametophyte is estimated to vary by >2× between bryophytes and pteridophytes, and the gametophytic selfing rate is estimated to vary by 88× between monoicous bryophytes and monoicous pteridophytes.

The *shadie* genome model

A SLiM genome model is used to define the number of sites over which mutations can be observed in simulations, as well as the potential for different alleles at each site to affect the fitness of individuals. *shadie* includes a *chromosome* submodule for defining genome models in Python syntax. This study presents a simple simulation to test whether signatures of selection left in the genome are detectably different for genes that affect only sporophyte fitness versus genes that affect only gametophyte fitness. To this end, we designed a chromosome with just two sites, one expressed in the sporophyte generation and one expressed in the gametophyte generation. These sites were placed sufficiently

far apart to ensure that linkage disequilibrium between sites did not interfere. We defined a linear chromosome 15 Mbp in length containing two genes (Figure 6D). The first gene is located 5 Mbp from the start, and defined such that mutations in this region yield a 10% fitness benefit only during the gametophyte phase (i.e., affect only the gametophyte phenotype), with no effect during the sporophyte phase. The second gene is located 10 Mbp from the start, and is contrastingly defined such that mutations yield a 10% fitness benefit only during the sporophyte phase, with a 0.5 dominance coefficient (i.e., fitness of a heterozygote is intermediate between that of the homozygotes), with no effect during the gametophyte phase. Mutations occurring outside of these two genic regions have no effect on fitness. Both genes are 1000 bp long and can experience repeated mutations, which occur at a rate of 5×10^{-7} per site per life stage.

The *shadie* demographic model

A demographic model is used to define the carrying capacity of populations through time and their evolutionary relationships (e.g., divergence times, migration rates). (Note that the population size in this context is not *effective population size*, which is an emergent property of the number of individuals and the many processes of a simulation affecting the probability that a sample in one generation is an ancestor of samples in the next, but rather is *census population size*.) *shadie* can be used to define a demographic model involving a single population, or in more complex scenarios where two or more populations diverge from a common ancestor. For our full simulations, we defined a single population evolving for 10,000 generations with a carrying capacity of 3000 individuals for both the gametophyte and sporophyte phases.

shadie simulation and post-simulation analysis

Here we performed simulations for only the monoicous bryophyte, dioicous bryophyte, and heterosporous pteridophyte models, as these were the only models that had been optimized for speed at the time of writing. Each was run under both "standardized" and "realistic" parameter settings (Table 1). We also simulated data under an approximation of the WF process (implemented as a SLiM nonWF model type) using the same chromosome and demography for comparison. For each model, we performed 20 replicate simulations from different random seeds, using a mutation rate of 5×10^{-7} per site per life stage generation and a recombination rate of 1×10^{-7} per site per diploid life stage generation (there is no recombination in the haploid life stage).

The resulting tree-sequence files were loaded and analyzed in *shadie's postsim* analysis module. This module includes functions to update the *tskit* tree-sequence class objects to correct the ages of nodes and mutations to compensate for the way in which the alternation-of-generations life cycles are coded in *SLiM/shadie* (one haploid generation + one diploid generation represents one standard SLiM generation), and to assign all extant genomes to a single population (haploids and diploids are assigned to different population IDs during *SLiM/shadie* simulations for implementation purposes). For each model, population genetic diversity (Tajima's π) was calculated in 100-Kbp sliding windows along the genome, and window means and standard deviations were calculated across replicate simulations.

Single-site dynamics

To aid in the interpretation of our full simulation results, we also implemented a set of simulations to investigate evolutionary dynamics at a single site over time. For the purposes of these simulations, we reduced the chromosome size to 5 Mbp and population size to 1000 individuals. Repeated simulations were loaded from a shared neutral burn-in and allowed to run for a further 2000 generations, and then a single beneficial mutation with a selection coefficient of 0.5 was introduced in the middle of a chromosome in a single individual. The mutation either affected only sporophyte fitness or only gametophyte fitness. We repeated simulations until the mutation fixed 50 times for each set of parameters and each mutation type, then calculated average fixation probability and time to fixation.

RESULTS

A *shadie* simulation can be implemented with only a few lines of human-readable Python code (Figure 5). By comparison, the Eidos source code that *shadie* produces to implement these models in SLiM ranges from 337–546 lines of code, depending on the selected life cycle model. The Eidos code has been optimized for speed and memory usage, which proved important when modeling reproduction that could theoretically involve very large numbers of sperm or eggs. By developing efficient operations in Eidos, we were able to reduce RAM usage by >50× between our earlier implementations and later optimized models.

Full simulations

Population genetic diversity varied significantly among the different life cycle models in simulations performed under both the "standardized" and "realistic" parameter settings, and all deviated from the results under a WF process (Figure 6). The effect of strong positive selection at two regions of the genome caused selective sweeps, reducing diversity in these regions as well as at neighboring neutral regions due to genetic linkage, which decayed toward the neutral background diversity rate with increasing distance from the selected sites.

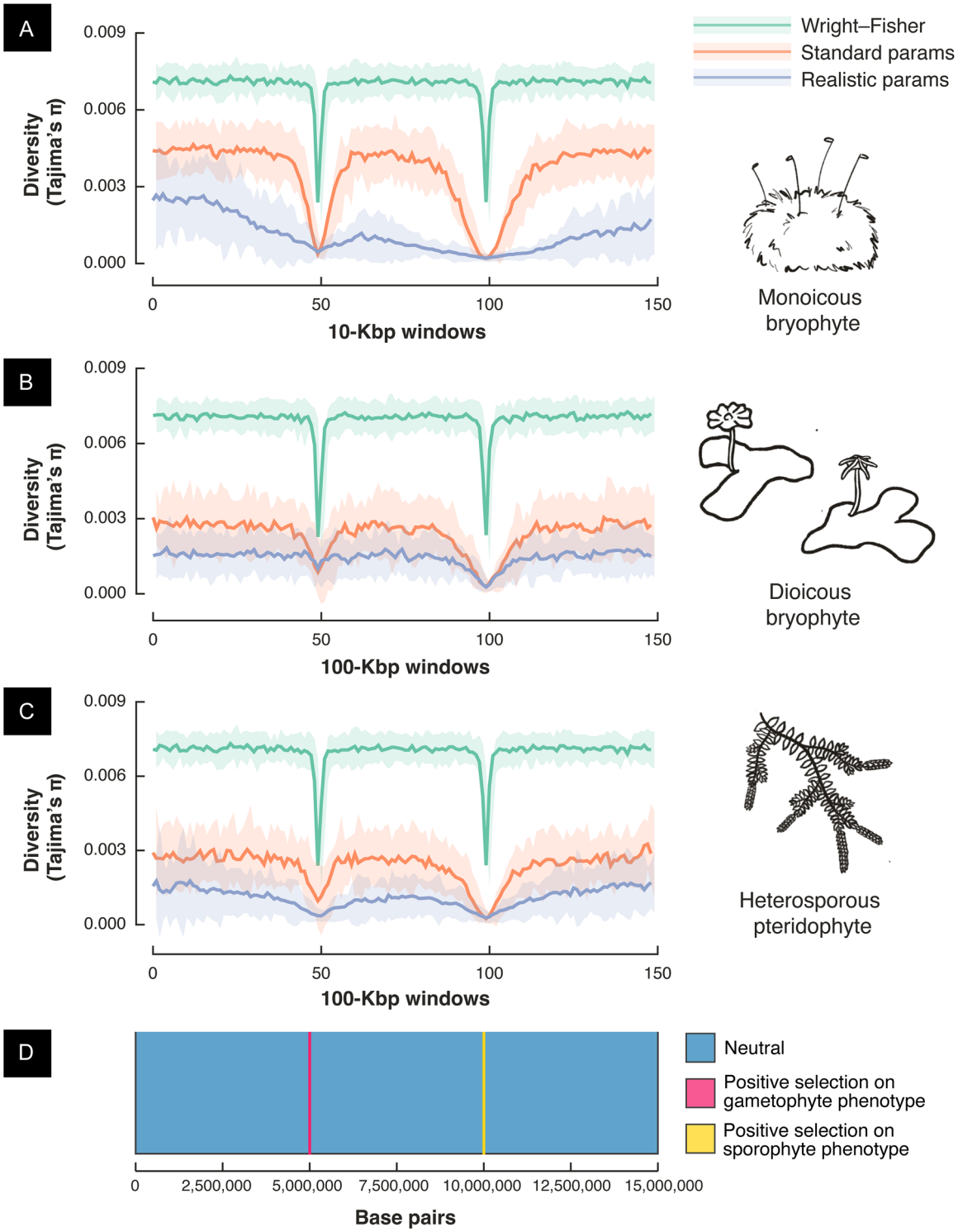

FIGURE 6 The results of the full simulations. Population genetic diversity is shown in 100-Kbp windows (mean ± 2 SD from 20 replicate simulations) across three models: (A) monoicous bryophyte, (B) dioicous bryophyte, and (C) heterosporous pteridophyte. Each panel also shows the result for a standard Wright–Fisher model in green for comparison. Simulations using standardized parameters are in orange; simulations using realistic parameters are in blue (see Table 1 for parameter values). Chromosome structure is shared across all simulations, depicted in (D) with neutral regions in blue, the gene with positive selection on the haploid gametophyte phenotype in pink, and the gene with positive selection on the diploid sporophyte phenotype in yellow. Both genes are 1000 bp in length, and each mutation that occurs within the gene region confers a 1% increase in fitness during the life stage in which it is expressed.

Under the standard parameter settings, results of the three plant models were overall quite similar, exhibiting a faster rate of linkage decay around the gene at 5 Mbp, which affected gametophyte fitness only, than for the gene at 10 Mbp, which affected sporophyte fitness only. Because all parameters affecting the life cycles were kept constant in these simulations, this result was a consequence primarily of the genetic model, in which the dominance coefficient defined for these genes (0.5) caused selection to have weaker fitness effects in heterozygotes of the diploid sporophyte. By contrast, under the approximation of a WF process, which does not explicitly model the gametophyte life phase, there was no difference in linkage disequilibrium between the two genes. Models also exhibited differences in their baseline levels of diversity in neutral regions in standardized simulations. This likely reflected differences in population sizes. The WF model has a carrying capacity of 3000 hermaphroditic individuals with random mating, whereas the monoicous bryophyte model has a carrying capacity of 3000 for both hermaphroditic gametophytes and sporophytes, but each sporophyte can produce 10 spores and each gametophyte has 10 archegonia, increasing the variance of reproductive success, and thus the amount of drift because N_e is effectively reduced. Diversity is further reduced in lineages with separate sexes, for which we also enforced the same carrying capacity of 3000, such that only 1500 gametophyte individuals were female on average. If population sizes were normalized by female reproductive output, baseline diversity levels would likely be more similar among the plant models, although this effect requires further investigation.

Simulations under more realistic parameterizations of the life cycle models yielded dramatically different levels of genetic diversity at both neutral and selected regions of the chromosome. The monoicous bryophyte model exhibited relatively low average diversity, but with the most extensive linkage disequilibrium among models, such that the entire chromosome was effectively under linked selection (Figure 6A). This was likely due to high rates of gametophytic cloning and gametophytic selfing in this model. By contrast, the dioicous bryophyte model exhibited low genetic diversity, but very fast rates of linkage decay back to the neutral diversity rate around selected sites. It also showed a weak signature of selection at the gene affecting gametophyte fitness that was barely distinguishable from background variation, whereas selection at the gene affecting sporophyte fitness showed a clear decrease in diversity (Figure 6B). Because the dioicous bryophyte simulations had similar cloning rates to monoicous bryophyte simulations, the differences among the two bryophyte models under realistic parameterizations must be a consequence of population sizes, sex ratios, and the lack of gametophytic selfing in dioicous bryophytes. Finally, heterosporous pteridophytes exhibited similarly low levels of overall diversity as dioicous bryophytes (Figure 6C), with a rate of linkage decay intermediate between the other two models. Intriguingly, the signatures of selection at the two genes in this model are less distinct from each other than in the bryophyte models.

Single-site simulations

In contrast to the spatial pattern of diversity along the chromosome discussed above, we also examined genetic diversity with respect to time, after the introduction of a single beneficial mutation. This enabled us to track the probability of fixation and time to fixation for each model. In simulations under the diploid WF model with dominance coefficient 0.5 and selection coefficient 0.5, beneficial mutations had a mean fixation probability of 0.53, which was lower than when selection acted only on the gametophyte stage in the WF model, where the mean fixation probability was 0.76 (implemented by setting the selection coefficient to 1.0). Fixation probabilities were greater under the WF model than any of the *shadie* life cycle models and varied significantly among the different models (Table 2). Across all simulations, fixation probabilities were consistently higher when selection affects the gametophyte stage.

Both life cycle variation and parameter differences among these processes affected fixation probabilities, as demonstrated by the fact that differences were observed in simulations under both the standardized and realistic parameter settings (Table 2). For example, under the standard parameters that exclude cloning and selfing (see Table 1), fixation probability was very high in the heterosporous pteridophyte model for mutations affecting gametophyte fitness (0.54), but very low for mutations affecting sporophyte fitness (0.059), a difference of nearly an order of magnitude. By comparison, beneficial mutations under both monoicous and dioicous bryophyte models were one-quarter as likely to fix when the gene affected sporophyte fitness versus gametophyte fitness (Table 2). All life cycle models exhibited a much larger disparity between fixation probabilities of mutations affecting sporophyte versus gametophyte fitness than were observed under a WF approximation.

Time to fixation did not vary greatly between models, although fixation time was consistently about twice as long for mutations affecting sporophyte fitness compared to mutations affecting gametophyte fitness. The rate at which diversity recovers through time around a beneficial mutation after it sweeps to fixation also does not vary among the different models (Figure 7). However, the total time for diversity to recover does vary, taking longer for models with a higher baseline level of genetic diversity, because they had more diversity to recover.

DISCUSSION

For educational and research purposes, *shadie* provides a convenient and easy-to-use wrapper around SLiM for exploring forward-in-time simulations in the context of plant life cycle models. In this article, we summarize the

TABLE 2 Settings and results for single-site simulations.[a]

	Wright–Fisher		Bryophyte: monoicous		Bryophyte: dioicous		Pteridophyte: heterosporous	
Life stage affected	**gam**	**spo**	**gam**	**spo**	**gam**	**spo**	**gam**	**spo**
Fixation probability, standard	0.76	0.54	0.39	0.13	0.41	0.12	0.54	0.056
Fixation probability, realistic	N/A	N/A	0.46	0.14	0.47	0.14	0.56	0.086
Fixation time, standard	91.0 ± 10.6	144.8 ± 37.4	67.6 ± 9.0	122.9 ± 19.3	66.0 ± 6.8	122.9 ± 19.3	58.9 ± 8.2	107.6 ± 20.2
Fixation time, realistic	N/A	N/A	56.3 ± 7.1	88.0 ± 16.0	57.3 ± 7.7	138.3 ± 21.0	63.7 ± 7.9	110.4 ± 18.0
Diversity at fixation, standard	23.7×10^{-4}	6.0×10^{-4}	13.8×10^{-4}	5.2×10^{-4}	13.2×10^{-4}	5.1×10^{-4}	6.6×10^{-4}	3.9×10^{-4}
Diversity at fixation, realistic	N/A	N/A	7.1×10^{-4}	6.7×10^{-4}	12.7×10^{-4}	4.9×10^{-4}	7.0×10^{-4}	3.3×10^{-4}

[a]10,000 generations of burn-in was used prior to the reported 10,000 generation simulation time. Parameters are the same as reported in Table 1, except the number of individuals is 1000 for both sporophyte and gametophyte populations. Fixation probability is calculated from the number of runs required for the mutation to fix 50 times. Time to fixation is reported in generations (one life stage per generation) ± 1 SD. Average diversity at time of fixation is reported for a 150-Kbp window centered on the mutation site, and therefore is not expected to equal 0.0. Selection coefficient, $s = 0.5$; dominance coefficient, $h = 0.5$ for all simulations except Wright–Fisher. gam: mutation affects gametophyte fitness only; spo: mutation affects sporophyte fitness only; for Wright–Fisher models, gam: dominance coefficient, $h = 1.0$, spo: dominance coefficient, $h = 0.5$. See Figure 7 for plots.

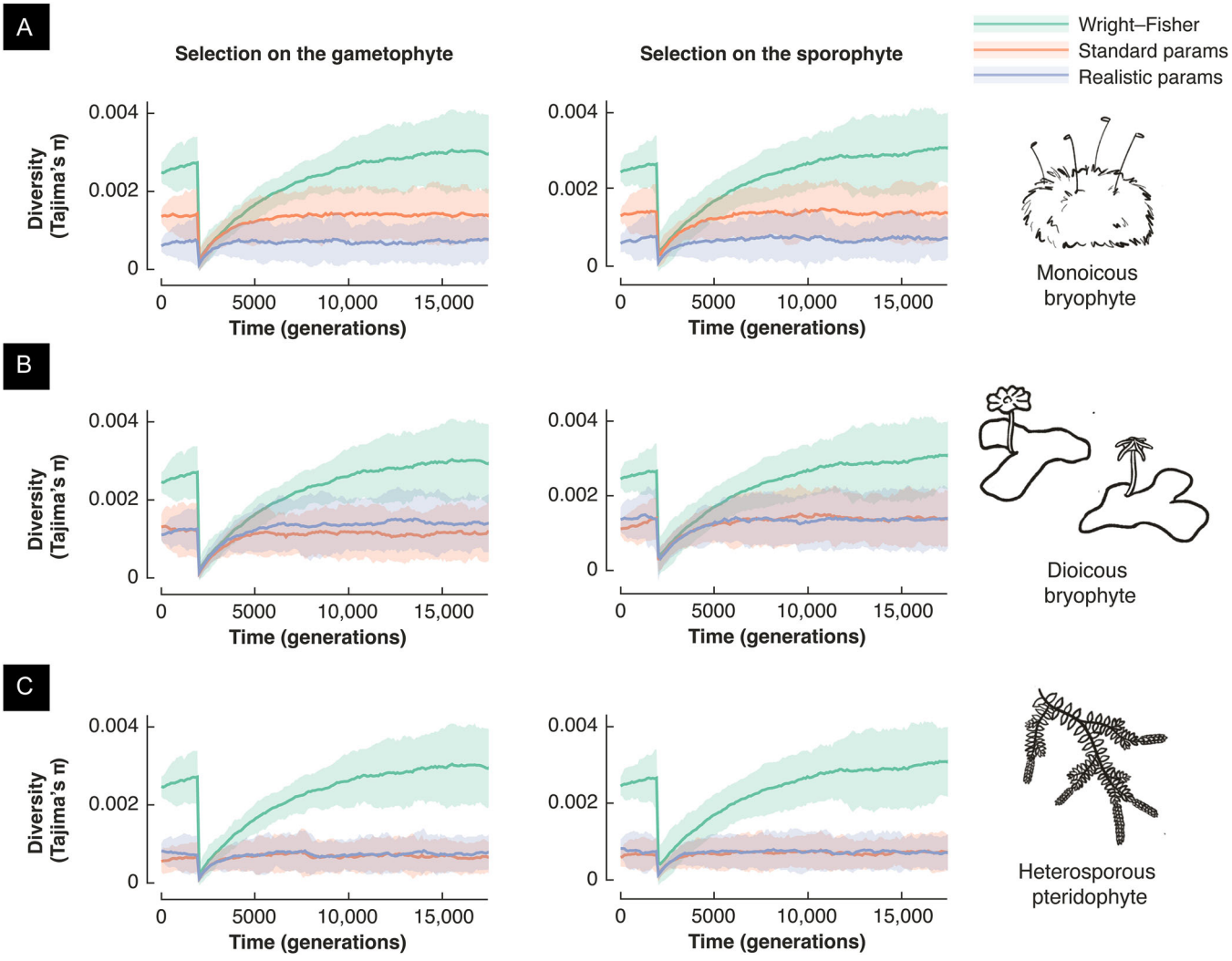

FIGURE 7 The results of the single-site simulations for (A) monoicous bryophyte, (B) dioicous bryophyte, (C) and heterosporous pteridophyte models. A beneficial mutation was introduced at 2000 generations and affected either gametophyte fitness only (left) or sporophyte fitness only (right). Standard parameters and realistic parameters are as defined in Table 1. Population genetic diversity in a 100-Kbp window around the beneficial mutation site (mean ± 2 SD from 50 replicate simulations). Time is in SLiM generations (one generation per life stage).

differences among the major plant lineages to describe how this complex biological diversity can be encoded into a set of parameterized models that can be used to configure aspects of life cycles and sexual systems to investigate their effects on population genetic statistics.

Full simulations

To test the potential effects of alternation-of-generations life cycles on signatures of selection, we used *shadie* to perform simulations under different plant models using a chromosome with two genes: one that affects fitness of the gametophyte phenotype and one that affects fitness of the sporophyte phenotype. Genome-wide patterns of genetic diversity varied among chromosomes simulated under each model, as well as between these models and the expectations of a WF process, even when all parameters were set identically (e.g., cloning and selfing rates at zero). These differences were manifest both in the average diversity at neutral genic regions and in the rate of decay of linkage disequilibrium around selected regions. It is important to note that in a chromosome with more realistic gene density, genetic linkage would likely cause a lower overall diversity to be maintained across the genome than was exhibited in our simulations, due to overlapping areas of linkage disequilibrium.

Patterns of genetic linkage also varied in all *shadie* models between the two genic regions on the same chromosome, showing how the efficacy of selection can vary among life phases (Figure 6). At sites experiencing repeated beneficial mutations, a clear decrease in genetic diversity, sometimes referred to as a "trough," around the selected site is expected (Ronen et al., 2015). The rate of linkage decay (the "width" of the trough) is affected by the selection coefficient of the mutations, while the magnitude of decrease in diversity ("depth" of the trough) is affected by how much time has passed since a mutation has fixed (Sella et al., 2009). Based on the results of our full simulations, if a beneficial mutation occurs during the gametophytic (rather than the sporophytic) life stage, then the strength of selection is effectively reduced (i.e., the effect of selection resembles that of a sporophytic mutation with a lower selection coefficient), but the mutation may fix more quickly.

This has important implications, as it is generally assumed that most adaptations caused by strong positive selection will occur during the dominant life stage of an organism. In bryophytes, this could lead to the expectation that selective sweeps will be common, because the dominant phase is haploid, and indeed, both bryophyte models exhibited stronger sweeps at the haploid-affecting gene region. By contrast, pteridophytes may be expected to experience softer selective sweeps, because selection is more likely to act on the sporophyte stage, where dominance effects occur. This expectation, however, is overly simplistic, as demonstrated by simulations performed under more realistic model parameterizations, which led to substantially greater variation in genome-wide patterns of genetic diversity than were observed between the two genic regions in the standardized simulations.

The effects of the more realistic model parameters can first be interpreted through their expected impacts on the effective population size during each life stage. For example, in dioicous bryophytes and heterosporous pteridophytes, the strongly biased gametophytic sex ratio greatly reduces the effective population size of the gametophyte phase, thus reducing the relative strength of selection versus drift acting on alleles during this life stage. Similarly, the high rate of gametophytic selfing in monoicous bryophytes greatly reduces the effective population size of the sporophyte population, because most sporophytes growing from the same gametophyte are identical. These factors contribute not only to the relative strength of selection during each phase, but also to the total extent to which genetic diversity is reduced by drift.

Not all outcomes of our nonWF simulations can be interpreted so simply. In most cases, when simulations involve selection at multiple genic regions, the results can deviate significantly from expectations based on simpler analytical or coalescent models. This may be particularly true when selection is paired with other processes that violate standard WF assumptions, such as cloning, selfing, and variable offspring numbers. Our simulations demonstrate that different patterns of genome diversity are produced by different life cycle models and parameter settings, but comprehensively understanding the effect that variation in a specific parameter will have on the genome requires further investigation (for which *shadie* should prove useful). Establishing how selection on the gametophyte life stage is expected to affect genome evolution, and how these effects are further impacted by other aspects of life history, will be important for contextualizing the results of empirical studies of gametophyte selection (Immler, 2019; Beaudry et al., 2020).

Single-site simulations

Our single-site simulations (Figure 7) demonstrate differences in the temporal dynamics of models with alternation of generations. Different models under different parameter settings showed a consistent pattern of mutations, taking about twice as many generations to fix when they affected only sporophyte fitness, compared to only gametophyte fitness. This finding is consistent with observed patterns of linkage disequilibrium in our full simulations, which show shallower troughs (smaller decrease in genetic diversity) around the gene affecting gametophytic fitness compared to sporophytic fitness. This finding is also consistent with the expectation that genetic diversity has had more time to recover at these sites. These differences affect the potential to detect selection based on diversity statistics. The small differences among models in terms of the effect of selection on genetic diversity through time contrasts with the spatial patterns of genomic diversity (Figure 6), where the effects of linked selection vary considerably among the different models.

Our single-site simulations also revealed that fixation probability is consistently higher for mutations affecting gametophyte fitness compared to sporophyte fitness (Table 2). Classical models of beneficial mutation dynamics (Moran, 1960; Kimura, 1962) expect fixation probability, P, of a single beneficial mutation $\approx 2^{\star}s^{\star}N_e/N$ where s is the selection coefficient, N_e is effective population size, and N is census size, for sufficiently large values of s and N ($s > 0.01$ and $N > 100$; Patwa and Wahl, 2008). If the selection coefficient for a beneficial mutation affecting only the gametophyte life stage is s, then the selection coefficient for a beneficial mutation affecting only the sporophyte life stage is $h^{\star}s$, where h is the dominance coefficient. Thus, based on our simulation parameters ($s = 0.5$, $h = 0.5$, $N = 1000$), we would expect the fixation probability for a mutation affecting sporophytic fitness to be exactly half that of a mutation affecting gametophytic fitness as long as N_e/N remains constant.

Our fixation probability results deviate drastically from these classical expectations, a disparity that is consistent with the findings of other studies. In models of haploid–diploid populations (populations with a mixture of haploid and diploid individuals, with asexual reproduction), Bessho and Otto (2022) found that the presence and degree of asexual reproduction (i.e., cloning) is expected to alter fixation probability and that the relative reproductive output of haploid versus diploid life stages greatly affects the strength of natural selection versus genetic drift. Of course, our models also violate many assumptions of these classical models, including variable N (and therefore, variable N_e), and the alternation-of-generations life cycle itself, in addition to the presence of cloning in our realistic parameter simulations.

In our WF simulations, we set $h = 0.5$ to simulate the equivalent of a mutation affecting sporophyte fitness only and $h = 1.0$ to simulate the equivalent of a mutation affecting gametophyte fitness only (i.e., no dominance effects). The expected result is that fixation probability should be twice as high when $h = 1.0$. However, we recover fixation probability $P = 0.76$ when $h = 1.0$, versus $P = 0.54$ when $h = 0.5$ (Table 2). Our WF simulations violate the assumptions of these classical models by implementing a carrying capacity, K, that allows N to fluctuate at or near K instead of holding N fixed. This clearly results in different values of N_e/N between simulations where $h = 1.0$ and $h = 0.5$. This suggests that the effect of alternation-of-generations life cycles on patterns of genome evolution may be understood through the relationship of life cycle features such as cloning and selfing on this N_e/N ratio and a better understanding of the consequences of fluctuating population sizes, which has been explored (as reviewed in Patwa and Wahl, 2008), but is not yet well understood.

Simulation parameters

Selecting appropriate parameter values for genetic simulations can be challenging. In fact, even obtaining realistic ranges for parameter values from empirical studies in the literature proved surprisingly difficult, underscoring the need for more population demographic studies in plants. The proportions of sporophytes to gametophytes, archegonia to antheridia, females to males, and megasporangia to microsporangia can all vary substantially (Appendix S2). Modeling every aspect of these life cycles at a realistic scale is unfeasible (for example, thousands of sperm are produced by each monoicous gametophyte), and is redundant when many individuals are mitotically identical. A goal of this work was thus to identify the parameters that are likely to most strongly influence evolutionary outcomes, and to design efficient simulations involving these parameters. We expect that the list of parameters (Appendix S1) will change over time as further development of *shadie* brings new insights into the many ways in which plant life cycles can impact evolution.

Future directions

Our hope is that by providing a convenient and reproducible method to implement plant life cycle models in SLiM through *shadie*, forward-in-time simulations will become more approachable for new users, allowing different research groups to build on past work, to become more familiar with SLiM, and even to extend our models for their own needs. s*hadie* can be used to explore many topics of interest to plant scientists, including dispersal dynamics, epistatic interactions, the effects of expression overlap between the gametophyte and sporophyte, inbreeding load due to gametophytic and sporophytic selfing, and many others. *shadie* also provides an effective method for testing theoretical assumptions. For example, *shadie* could be used to test the predictions of theoretical studies such as Bengtsson and Cronberg (2009), which predicts how the effective population size of bryophytes is affected by the size of gametophyte versus sporophyte population sizes, and Haig (2016), which predicts that pteridophytes will be subject to greater inbreeding depression than bryophytes based on their life cycles.

For instructors, *shadie* presents an opportunity to teach students about the complex and nuanced life cycles of the major plant lineages with an interactive interface that simultaneously introduces them to evolutionary biology concepts and basic coding in Python. Students can use *shadie* to generate hypotheses that they then test in a guided laboratory setting, or simply to explore the outcomes of changing various parameters on simple (Wright–Fisher) versus more complex models of evolution. The *shadie* documentation contains resources for educators and students on the topics of plant life cycles, plant evolution, *shadie*, SLiM, and more, and it will continue to be developed and refined as more functionality is added to *shadie*.

shadie will be continually developed to add more features and supporting documentation, particularly teaching materials. Further development of *shadie* will focus on

enhancing user experience and improving the user interface, as well as adding features that extend the capabilities of *shadie* for use as a research tool. Planned features include the ability to fine-tune differential gene expression between haploid and diploid life stages, an extended post-simulation analysis module for calculating additional statistics such as d*N*/d*S* and the McDonald–Kreitman test, as well as tools for developing new summary statistics that may be more appropriate for alternation-of-generations models.

AUTHOR CONTRIBUTIONS

E.S.S. and D.A.R.E. wrote the *shadie* program. E.S.S. and B.C.H. designed the models with input from D.A.R.E. E.S.S. and D.A.R.E. designed and performed the simulations with input from B.C.H. E.S.S. designed and illustrated Figures 1–5 with input from B.A.A., B.C.H., and D.A.R.E. E.S.S. and D.A.R.E. analyzed the data and generated data visualizations. All authors contributed to the preparation of the first draft of the manuscript. E.S.S., D.A.R.E., and B.C.H. prepared the revisions with input from B.A.A. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors are indebted to two anonymous reviewers, whose feedback greatly improved the manuscript. We thank Peter Ralph (University of Oregon) for his prompt and exceedingly useful advice on the *pyslim* Discussions page on GitHub. We also thank Carlos Maya-Lastra (Columbia University) for his plotting and simulation assistance, and Jonathan Morris for editing and comments on the manuscript. This work was supported by the National Science Foundation (NSF DEB-2046813 awarded to D.A.R.E.).

OPEN RESEARCH BADGES

This article has been awarded Open Materials badge. All materials are publicly accessible via the Open Science Framework at https://github.com/elissasoroj/shadie. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki

DATA AVAILABILITY STATEMENT

All code used in this study is publicly available at https://github.com/elissasoroj/shadie, including *shadie* Python source code, documentation, and Jupyter notebooks allowing users to reproduce the simulations and analyses reported in this article. At the time of writing, *shadie* uses SLiM version 3.7, available at https://messerlab.org/slim/, to perform forward-in-time simulations.

ORCID

Elissa S. Sorojsrisom https://orcid.org/0000-0001-9938-0194
Benjamin C. Haller https://orcid.org/0000-0003-1874-8327
Barbara A. Ambrose https://orcid.org/0000-0002-9617-0148
Deren A. R. Eaton https://orcid.org/0000-0002-5922-2770

REFERENCES

Beaudry, F. E., J. L. Rifkin, S. C. Barrett, and S. I. Wright. 2020. Evolutionary genomics of plant gametophytic selection. *Plant Communications* 1(6): e100115.

Bell, P. R. 1995. Incompatibility in flowering plants: Adaptation of an ancient response. *Plant Cell* 7(1): 5–16.

Bengtsson, B. O., and N. Cronberg. 2009. The effective size of bryophyte populations. *Journal of Theoretical Biology* 258(1): 121–126.

Bessho, K., and S. P. Otto. 2022. Fixation and effective size in a haploid–diploid population with asexual reproduction. *Theoretical Population Biology* 143: 30–45.

Boavida, L. C., and S. McCormick. 2010. Gametophyte and sporophyte. *In* Encyclopedia of Life Sciences (ELS). John Wiley & Sons Ltd., Chichester, Sussex, United Kingdom. Available at: https://doi.org/10.1002/9780470015902.a0002038.pub2 [accessed 18 October 2021].

Chettoor, A. M., A. R. Phillips, C. T. Coker, B. Dilkes, and M. M. Evans. 2016. Maternal gametophyte effects on seed development in maize. *Genetics* 204(1): 233–248.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper and Row, New York, New York, USA.

de Groot, G. A., B. Verduyn, E. R. Wubs, R. H. Erkens, and H. J. During. 2012. Inter- and intraspecific variation in fern mating systems after long-distance colonization: The importance of selfing. *BMC Plant Biology* 12(1): 3.

Eaton, D. A. 2020. Toytree: A minimalist tree visualization and manipulation library for Python. *Methods in Ecology and Evolution* 11(1): 187–191.

Eppley, S. M., P. J. Taylor, and L. K. Jesson. 2007. Self-fertilization in mosses: A comparison of heterozygote deficiency between species with combined versus separate sexes. *Heredity* 98(1): 38–44.

Ewens, W. J. 2004. Mathematical population genetics, Vol. 1: Theoretical introduction. Springer, New York, New York, USA.

Haig, D. 2016. Living together and living apart: The sexual lives of bryophytes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1706): 20150535.

Haig, D., and A. Wilczek. 2006. Sexual conflict and the alternation of haploid and diploid generations. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361(1466): 335–343.

Haller, B. C., and P. W. Messer. 2019. SLiM 3: Forward genetic simulations beyond the Wright–Fisher model. *Molecular Biology and Evolution* 36(3): 632–637.

Haller, B. C., K. J. Galloway, P. W. Messer, and P. L. Ralph. 2019. Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Molecular Ecology Resources* 19(2): 552–566.

Hoban, S., G. Bertorelle, and O. E. Gaggiotti. 2012. Computer simulations: Tools for population and evolutionary genetics. *Nature Reviews Genetics* 13(2): 110–122.

Honys, D., and D. Twell. 2004. Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biology* 5(11): R85.

Igic, B., and J. W. Busch. 2013. Is self-fertilization an evolutionary dead end? *New Phytologist* 198(2): 386–397.

Immler, S. 2019. Haploid selection in "diploid" organisms. *Annual Review of Ecology, Evolution, and Systematics* 50: 219–236.

Jesson, L. K., D. S. Perley, A. P. Cavanagh, J. A. Cameron, and D. S. Kubien. 2012. Mating and fitness consequences of sexual system in the moss *Atrichum undulatum* sl (Polytrichaceae). *International Journal of Plant Sciences* 173(1): 16–25.

Kelleher, J., K. R. Thornton, J. Ashander, and P. L. Ralph. 2018. Efficient pedigree recording for fast population genetics simulation. *PLoS Computational Biology* 14(11): e1006581.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47(6): 713.

Kingman, J. F. C. 1982. The coalescent. *Stochastic Processes and Their Applications* 13(3): 235–248.

Kluyver, T., B. Ragan-Kelley, F. Pérez, B. E. Granger, M. Bussonnier, J. Frederic, K. Kelley, et al. 2016. Jupyter Notebooks: A publishing format for reproducible computational workflows. *In* F. Loizides and B. Schmidt [eds.], Positioning and power in academic publishing: Players, agents and agendas, 87–90. IOS Press, Amsterdam, the Netherlands.

Kumar, K. 2001. Reproductive biology of pteridophytes. *In* B. M. Johri and P. S. Srivastava [eds.], Reproductive biology of plants, 175–214. Springer, Berlin, Germany.

Lamborn, E., J. E. Cresswell, and M. R. MacNair. 2005. The potential for adaptive evolution of pollen grain size in *Mimulus guttatus*. *New Phytologist* 167(1): 289–296.

Lenormand, T., and J. Dutheil. 2005. Recombination difference between sexes: A role for haploid selection. *PLoS Biology* 3(3): e63.

Mascarenhas, J. P. 1989. The male gametophyte of flowering plants. *Plant Cell* 1(7): 657–664.

Moran, P. A. P. 1960. The survival of a mutant gene under selection. II. *Journal of the Australian Mathematical Society* 1: 485–491.

Patwa, Z., and L. M. Wahl. 2008. The fixation probability of beneficial mutations. *Journal of the Royal Society Interface* 5: 1279–1289.

Peters, M. A. E., and A. E. Weis. 2018. Selection for pollen competitive ability in mixed-mating systems. *Evolution* 72(11): 2513–2536.

Project, J., M. Bussonnier, J. Forde, J. Freeman, B. Granger, T. Head, C. Holdgraf, et al. 2018. Binder 2.0: Reproducible, interactive, sharable environments for science at scale. Proceedings of the 17th Python in Science Conference. Website: https://conference.scipy.org/proceedings/scipy2018/pdfs/project_jupyter.pdf [accessed 18 October 2021].

Rieber, L. P. 2005. Multimedia learning in games, simulations, and microworlds. *In* R. Mayer [ed.], The Cambridge handbook of multimedia learning, 549–567. Cambridge University Press, New York, New York, USA.

Ronen, R., G. Tesler, A. Akbari, S. Zakov, N. A. Rosenberg, and V. Bafna. 2015. Predicting carriers of ongoing selective sweeps without knowledge of the favored allele. *PLoS Genetics* 11(9): e1005527.

Sella, G., D. A. Petrov, M. Przeworski, and P. Andolfatto. 2009. Pervasive natural selection in the *Drosophila* genome? *PLoS Genetics* 5(6): e1000495.

Shaw, J., and S. C. Beer. 1999. Life history variation in gametophyte populations of the moss *Ceratodon purpureus* (Ditrichaceae). *American Journal of Botany* 86(4): 512–521.

Shaw, A. J. 2000. Population ecology, population genetics and microevolution. *In* A. J. Shaw and B. Goffinet [eds.], Bryophyte biology, 369–402. Cambridge University Press, Cambridge, United Kingdom.

Shead, T. M. 2014. *toyplot*, Version 0.19.0 (2014-08-25). U.S. Department of Energy, Sandia National Laboratory, Albuquerque, New Mexico, USA. Website: https://www.osti.gov/servlets/purl/1335046 [accessed 23 September 2021].

Sigel, E. M., E. Schuettpelz, K. M. Pryer, and J. P. Der. 2018. Overlapping patterns of gene expression between gametophyte and sporophyte phases in the fern *Polypodium amorphum* (Polypodiales). *Frontiers in Plant Science* 9: 1450.

Stoeckel, S., S. Arnaud-Haond, and S. A. Krueger-Hadfield. 2021. The combined effect of haplodiplonty and partial clonality on genotypic and genetic diversity in a finite mutating population. *Journal of Heredity* 112: 78–91.

Szövényi, P., S. A. Rensing, D. Lang, G. A. Wray, and A. J. Shaw. 2011. Generation-biased gene expression in a bryophyte model system. *Molecular Biology and Evolution* 28(1): 803–812.

Tanurdzic, M., and J. A. Banks. 2004. Sex-determining mechanisms in land plants. *The Plant Cell* 16(Suppl 1): S61–S71.

SUPPORTING INFORMATION

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

Appendix S1. Detailed description of all model parameters. For *shadie* models, there is one generation per life stage (i.e., one gametophyte [haploid] + one sporophyte [diploid] cycle = 2 generations in simulation time). An updated parameter glossary is maintained in the online *shadie* documentation.

Appendix S2. Realistic parameter values, obtained from the literature where possible, and otherwise estimated by the authors.

How to cite this article: Sorojsrisom, E. S., B. C. Haller, B. A. Ambrose, and D. A. R. Eaton. 2022. Selection on the gametophyte: Modeling alternation of generations in plants. *Applications in Plant Sciences* 10(2): e11472. https://doi.org/10.1002/aps3.11472