

Hybridization can facilitate species invasions, even without enhancing local adaptation

Mohsen B. Mesgaran^a, Mark A. Lewis^b, Peter K. Ades^c, Kathleen Donohue^d, Sara Ohadi^a, Chengjun Li^a, and Roger D. Cousens^{a,1}

^aSchool of BioSciences, The University of Melbourne, Victoria 3010, Australia; ^bMathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada T6G 2G1; ^cSchool of Ecosystem and Forest Sciences, The University of Melbourne, Victoria 3010, Australia; and ^dDepartment of Biology, Duke University, Durham, NC 27708

Edited by Alan Hastings, University of California, Davis, CA, and approved July 5, 2016 (received for review April 6, 2016)

The founding population in most new species introductions, or at the leading edge of an ongoing invasion, is likely to be small. Severe Allee effects—reductions in individual fitness at low population density—may then result in a failure of the species to colonize, even if the habitat could support a much larger population. Using a simulation model for plant populations that incorporates demography, mating systems, quantitative genetics, and pollinators, we show that Allee effects can potentially be overcome by transient hybridization with a resident species or an earlier colonizer. This mechanism does not require the invocation of adaptive changes usually attributed to invasions following hybridization. We verify our result in a case study of sequential invasions by two plant species where the outcrosser *Cakile maritima* has replaced an earlier, inbreeding, colonizer *Cakile edentula* (Brassicaceae). Observed historical rates of replacement are consistent with model predictions from hybrid-alleviated Allee effects in outcrossers, although other causes cannot be ruled out.

species colonization | mating system | model | *Cakile* | sea-rockets

Colonizing species rarely encounter empty environments and may, as a result, come into contact with close relatives. The most intensively discussed case of related colonizer–resident interactions, although by no means the only example, is the entry of modern humans into regions occupied by Neanderthals (1), eventually replacing them. Colonizing and resident species can interact in various ways, with consequences for their population dynamics. Competition, for example, is likely to impede successful establishment and population growth of newcomers; in contrast, resident species may facilitate new colonizers through hybridization or by shared interactions with organisms from other trophic levels. Although some interactions have been identified and are widely considered to predominate, there may be other interactions whose significance is yet to be explored. These interactions may be important in some or many cases, and a better knowledge of them might alter our interpretations of empirical observations. Here, we will investigate one such potential interaction, arising from hybridization between plant species sharing animal pollinators.

Hybridization can introduce adaptive alleles that contribute to expansion into new habitats (2, 3) and may result in new taxa. Local adaptation, resulting from changes in fitness, is the only positive role for species hybridization considered in the invasion literature; the same is true for admixtures of species genotypes from multiple sources. However, is it possible that there could be other benefits from hybridization, purely demographic ones? Furthermore, are these additional benefits capable of altering invasion dynamics over similar timescales and magnitudes? If so, such interactions would need to be excluded before adaptation can merely be assumed.

Many colonizing populations are susceptible to the demographic challenge of Allee dynamics, either because the initial founding population is small (4, 5) or because rare, long distance dispersers, beyond the present range, will initiate low-density satellite populations (6). Allee effects could slow range expansion or even halt it completely (7). In plants, a small founding population can experience Allee effects because of a scarcity of compatible mates,

low pollinator visitation, or both (8, 9); the term “pollen limitation” is often used as a generic term when the exact mechanism is unclear. In the extreme case of a single arriving adult, population persistence would normally be impossible unless the species is capable of asexual reproduction or self-fertilization [“Baker’s Law” (10)]. Here, we propose a positive role for hybridization in species invasions and range expansion, a purely demographic mechanism without the requirement for any new adaptation to result.

If a mate-limited colonizing species is capable of hybridizing with a species already present, the colonizing species could overcome the otherwise insurmountable limitations imposed by Allee effects. Early-generation interspecific crosses (i.e., hybrids) could enable the colonizer genes to establish. Pure colonizer-type individuals can subsequently arise through crossing among hybrid lineages (11) or repeated backcrossing with the colonizer parents. We further hypothesize that this hybridization-rescue effect is more likely to eventuate if the new species and hybrids are more attractive to pollinators or when the hybrids are more compatible (i.e., more likely to produce viable offspring) with the newcomer than with the established species. That is, after sufficient generations of asymmetric breeding (backcrossing to the new colonizer), plants will increasingly come to resemble the original newcomers. The arriving species will essentially have been reconstituted, at least in its nuclear genome, and its population can increase now that it has escaped the critical effects of Allee dynamics. In effect, resident populations of a cross-compatible

Significance

Species colonizations frequently occur in regions already occupied by other species of the same genus, or by other genotypes of the same species. When hybrid genotypes become more invasive than their parents, improved local adaptation is inevitably assumed. Here, we show that there is another way that hybridization could facilitate invasions, by temporarily overcoming pollen limitation when populations are small. We verify this conclusion using two contemporary sequential invaders, a plant analog of the Neanderthal–modern human system. Our findings indicate that it may be misleading to assume that all cases of invasive hybrids result from enhanced phenotypic suitability. Our hypothesis may be even more relevant to animal species because, unlike plants, animals can rarely escape mate limitation by self-fertilization.

Author contributions: M.B.M., M.A.L., P.K.A., K.D., and R.D.C. designed research; R.D.C. developed concept; P.K.A. helped in concept development; P.K.A. provided genetics expertise; M.B.M., S.O., and C.L. performed research; M.B.M., S.O., and C.L. analyzed data; and M.B.M., P.K.A., K.D., and R.D.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

See Commentary on page 9963.

¹To whom correspondence should be addressed. Email: rcousens@unimelb.edu.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1605626113/-DCSupplemental.

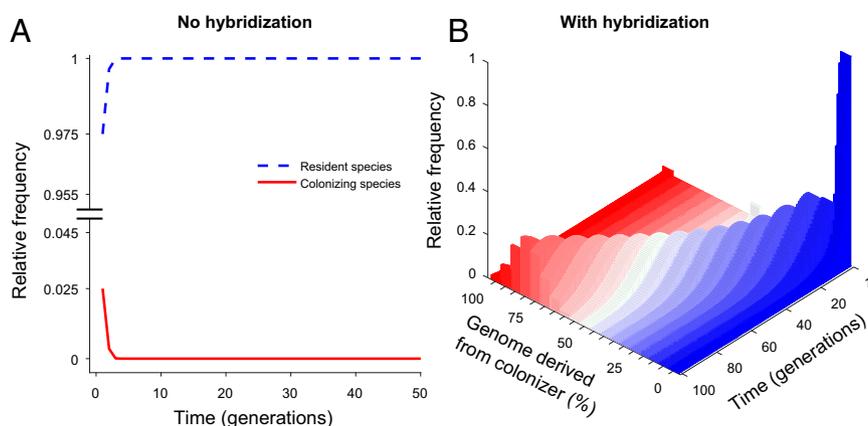


Fig. 1. Dynamics predicted by the generic model when a small number of a colonizing species arrives at a new location currently occupied by a resident species: no hybridization (*A*) and hybridization (*B*). Although in both *A* and *B*, pollinator preference parameter W is set to favor more backcrossing to the incoming genotype than to the resident species, the incoming genotype fails to establish when it does not hybridize with the resident species (*A*). However, when there is hybridization (*B*), the population steadily accumulates a higher proportion of individuals similar to the colonizer, whereas the resident genotypes decline in abundance and are finally replaced by the colonizer. Dark red indicates individuals whose genome is indistinguishable from the colonizer, dark blue indicates genomes indistinguishable from the resident species, and gray represents hybrid genotypes comprising 50% of each parental species' genome. See [Tables S1](#) and [S2](#) for lists of parameters and their values.

species, which may otherwise negatively affect the newly colonizing species through competition, become stepping stones through the landscape for a self-incompatible species via hybridization.

We develop a model that confirms that transient hybridization can overcome Allee effects under a wide range of ecological settings in relation to breeding system, pollinator behavior, and life history traits. Moreover, we show in a case study using a simplified version of the model that the parameter values required for the demographic mechanism to be as effective as genetic changes in fitness are fulfilled by the sequential invaders *Cakile edentula* and *Cakile maritima*. Our hypothesis therefore provides a possible explanation for the rapid replacement of *C. edentula* over a large part of its invasive range by *C. maritima* in the west of North America, New Zealand, and Australia (12–15).

Results

Our hypothesis is clearly supported by the model results. In the absence of hybridization, when we introduce a small number of

seeds, the incoming species population declines rapidly to extinction (Fig. 1*A*). The same propagule pressure but with hybridization transforms the population from one dominated by resident species and resident-like genotypes (blue in Fig. 1*B*) to one dominated by colonizer and colonizer-like genotypes (red in Fig. 1*B*). Over time, the colonizer genotype is reconstituted and the resident species is eventually replaced by the new colonizer. Allee population thresholds are always smaller with hybridization than without it, proving that introgression is asymmetric in favor of the colonizer [i.e., when $W_0 > 1$ (Fig. 2*A*) or $\beta > 1$ (Fig. 2*B*)]. Even with no bias in pollinator behavior ($W_0 = 1$) or compatibility ($\beta = 1$), hybridization still reduces the Allee threshold slightly (Fig. 2) because some of the otherwise wasted pollen results in seed production via the resident species, thus contributing positively to the dynamics of colonizer genes.

Consistent with the established theory that selfing provides reproductive assurance (16), the Allee threshold for establishment of colonizing species decreases with its selfing rate for both

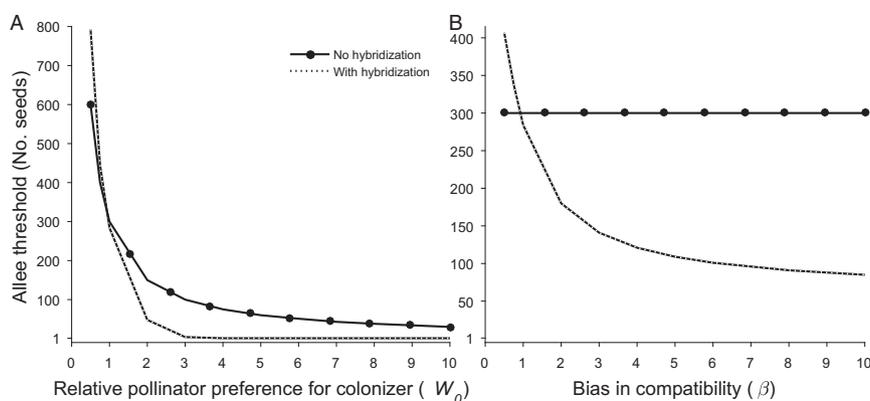


Fig. 2. The critical number of propagules for the establishment of a colonizing species, the predicted Allee threshold, as a function of either pollinator behavior (assuming equal compatibility, $\beta = 1$) (*A*) or reproductive compatibility of hybrids with the parental genotypes (assuming equal pollinator preference, $W_0 = 1$) (*B*). Note that *B* would also indicate the results expected from passive pollination, such as by wind. Compatibility, β , has no effect on the dynamics of either species when they do not hybridize: the Allee threshold is therefore constant for the no-hybridization scenario (solid line in *B*). When $W_0 = 1$ and $\beta = 1$, backcrossing is symmetric, but when β is larger than 1, F1 and later generation hybrids are more compatible with the colonizing-type individuals. In our simulations, the criterion for the Allee threshold was that the population size (seed number) after 100 generations was the same as the initial population size. See [Tables S1](#) and [S2](#) for lists of parameters and their values.

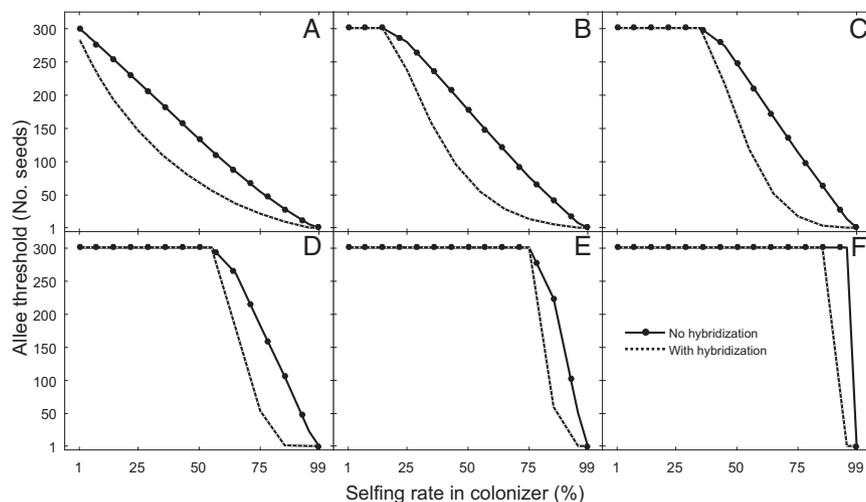


Fig. 3. Predicted Allee thresholds for the establishment of a colonizing species as a function of its self-fertilization rate, assuming that the resident species has a self-fertilization rate of 1% (A), 20% (B), 40% (C), 60% (D), 80% (E), or 99% (F). The minimum number of seeds necessary for the establishment of the colonizer, its Allee threshold, is never larger with hybridization than without. See [Tables S1](#) and [S2](#) for lists of parameters and their values.

hybridization and no-hybridization scenarios. However, Allee thresholds were always smaller for a hybridizing colonizer than the nonhybridizing one over a wider range of simulated breeding systems of both species (Fig. 3). Because the resident species adopts a more autogamous breeding system, Allee thresholds become larger in the colonizing species, but with hybridization, the incoming species could still invade the resident population (Fig. 3).

To test our predictions in a contemporary invasion, we parameterized a simplified version of our model with a single hybrid category (F1 hybrids), using empirical data from *Cakile maritima* [self-compatible (SC) (17)] and *C. edentula* [self-incompatible (SI) (17, 18)]. Why and how the established invader (*C. edentula*) has been replaced by a newly arrived species (*C. maritima*) in three parts of the world has remained a mystery over the 40 y since the replacement was first reported (13). Could hybridization with the prior SC invading plant have increased the probability of subsequent establishment of an SI species? Hybrids between the two species can be produced easily by hand (17). We have also confirmed, using genetic markers, that individuals with intermediate morphology near the leading edge of the *C. maritima* invasion in Tasmania are the results of hybridization (19).

We simulated the dynamics of *C. maritima* with and without hybridization with *C. edentula*. The results indicate that hybridization could facilitate the establishment of a small (otherwise

sub-Allee threshold) population of the primarily self-incompatible *C. maritima* (Fig. 4), with a strong correspondence between the timescales for species replacement predicted by the model simulation (Fig. 4A) and the frequency of *C. maritima* relative to *C. edentula* in herbarium specimens (Fig. 4B) (20). Without hybridization, small initial populations of *C. maritima* were predicted to fail to establish (Fig. 4C). A note of caution is required for this comparison: our model simulates the population dynamics within a single location (with no account of subsequent spread), whereas herbarium data give a (very crude) representation of invasion dynamics over a wider geographic scale (21). Our simulation results are, however, also in good agreement with the replacement time-scale observed on Lord Howe Island, where *C. maritima* replaced *C. edentula* in perhaps 20–30 y (14).

Our simulation model and the *Cakile* system provide support for our hypothesis of an overlooked, purely demographic, role for hybridization in both establishment and spread. A robust test of the hypothesis would require deliberate introductions into regions lacking one or other species, which would probably be unethical or unwise. However, in the case of SI and SC *Cakile* species in Australia, there is now clear evidence that interspecific hybridization has occurred during the establishment and expansion of *C. maritima* (19). Genetic and morphological evidence across the current *C. maritima* invasion wave-front in Tasmania

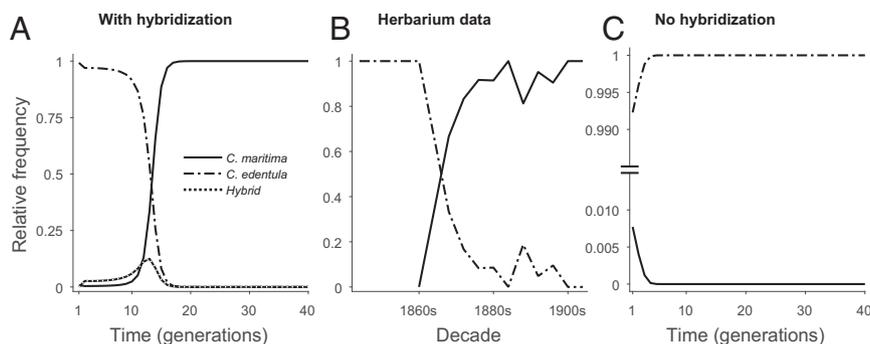


Fig. 4. Population dynamics of *Cakile* predicted by the simplified model (A and C) and evidence from herbarium samples (B). Whereas hybrid presence is predicted to be ephemeral, the predicted dynamics of *C. maritima* and *C. edentula* (A) corresponds to the pattern and time scale of change in their representation in herbarium specimens (B) collected from Victoria and South Australia. Adapted from ref. 20. C shows what would happen without hybridization. Parameters and their values are given in [Table S3](#).

(19) match what would be expected if species replacement by this mechanism had occurred. Specifically, in places where *C. edentula* has a very high relative frequency, hybrids are common and even plants that are morphologically indistinguishable from one parent often have the chloroplast genotypes of the other (19). In regions where *C. edentula* has long since disappeared and there is now little evidence of nuclear introgression, *C. edentula* chloroplasts may still be common in *C. maritima*, indicating extensive past hybridization and reconstitution of parental phenotypes. Although our artificial intercrossing provided no evidence for bias in cross-compatibility of hybrids with the two parental species (Fig. S3), pollinators clearly preferred to visit *C. maritima*-like individuals over *C. edentula* when both were present (SI Text, *Simplified Model for C. maritima and C. edentula*, *Sources of data for parameterization*). We note that introgression has been proposed before for these species (22) but based on less convincing data and without any mechanism being proposed (19).

If this demographic phenomenon does occur in nature, is it idiosyncratic—a unique combination of species sharing pollinators and invading the same habitat in the same sequence—and of little relevance to ecology and evolution in general, or is it more common? We believe that the latter may be true. First, both hybridization (3) and Allee effects are common in biological invasions (7). Evidence is mounting for the importance of hybridization as one of the processes driving invasions, although so far, the explanations have relied solely upon genetic consequences of hybridization endowing adaptive benefits (refs. 2 and 3 but see ref. 11). Indeed, the risks of extinction and genetic swamping of native species have been raised as dire consequences of hybridization (2). Allee effects or, more specifically, pollen-limited seed production, are expected to be common, because more than 80% of plants rely on pollen transfer for reproduction (23) and ~50% of plant species are obligate outcrossers (24). Pollen limitation, resulting from low availability of both compatible mates and pollinators, seems to be more common in introduced species than their native counterparts (25, 26). First, our model shows that the presence of another species can alleviate such pollen limitation and reduce Allee effects through hybridization with related species. Second, our model shows that the rescue effect of hybridization will be stronger when there is an asymmetry in the direction of introgression (backcrossing): this phenomenon seems to be common in both natural and artificial hybridizations. Pollinators rarely commit to random foraging bouts but rather discriminate among plant types (27); such behavior can result in assortative mating and produce the required asymmetry. Intercrossing success (i.e., the production of viable seeds) is also known to be affected by the direction of crossing (28, 29). The interplay between these pre- and/or postzygotic processes can result in unequal transfer of parental genetic material into the genome of hybrid progeny (28, 29). Whereas the short-term demographic consequence of this asymmetry is facilitation of establishment, the long-term outcome can be species replacement as we see in *Cakile* species.

Materials and Methods

We developed a density-dependent, time-discrete, deterministic model based on the life cycle of an annual plant with no persistent seedbank and occurring in a small, isolated patch of suitable habitat. The model incorporates demography, mating systems, pollinators, and quantitative genetics through the following recurrence difference equation:

$$N_{x(t+1)} = \sum_y M_y \theta_y F_y(M) \left[s_y H(x:y, y) - (1 - s_y) \sum_z C(y, z) V_y \phi(M) H(x:y, z) \right],$$

where $N_{x(t+1)}$ is the number of seeds from genotype x in generation $t + 1$, and $M_y = N_{y(t)} S_y(N)$ denotes the total number of female adults with genotype y , surviving from $N = \sum_k N_{k(t)}$ total seeds according to a Beverton–Holt type recruitment function, $S(N)$ (Eq. S1). θ_y is the per capita ovule production of female genotype y in the absence of neighboring plants, which decreases

with total adult population size ($M = \sum_k M_{k(t)}$) according to a rectangular hyperbolic function $F_y(M)$ (Eq. S3) because of resource competition and other negative interactions between plants. A fraction s_y of flowers produces seeds through selfing, whereas the $1 - s_y$ nonselfed proportion relies on pollinators for seed production. Not all pollen from other genotypes will result in viable offspring; hence the compatibility of genotype y with genotype z is given by $C(y, z)$ (Fig. S2). The relative frequency of genotype y , adjusted for pollinator preference (Fig. S3), is V_y , whereas the probability of pollination is related to the total adult population size following a Holling type III functional response, $\phi(M)$ (30).

We used the number of chromosomes as a proxy for determining the total number of genotype classes (which is $2n + 1$) and the chromosome combination of the two parental species to characterize x (Fig. S4). That is, a pure individual of the incoming species, denoted by $x = 0$, has no alien chromosome from the resident species whereas the karyotype of an F1 hybrid embodies an equal number of chromosomes from the two parental species, so it is shown as $x = 0.5$. The proportion of ancestry from the incoming species decreases as x approaches 1, whereas that of the resident species increases, with $x = 1$ indicating a pure individual of the resident species with no alien chromosome from the incoming species. A diploid individual then undergoes gametogenesis following the principle of independent assortment of chromosomes, assuming no crossing over, where a haploid gamete randomly receives a mixture of chromosomes from the two parental species. By taking the convolution over the gamete types that can be produced by two crossing diploid individuals, we obtained the probability $H(x:y, z)$ of two parents with genotype y and z producing an offspring with genotype x . In essence, our quantitative genetic model is similar to the hypergeometric phenotypic model used previously to study sympatric speciation (31, 32) and risk of extinction by hybridization (33). We mapped genotypic class x to a phenotypic character using a Gaussian model (30) (Fig. S5). We simulated the population dynamics of 19 genotypic classes corresponding to a diploid species with $2n = 18$ over 100 generations. This value was chosen to correspond to *Cakile* spp. ($2n = 18$) which we consider later. A smaller or larger number of chromosomes will respectively either increase or decrease the speed at which species replacement will occur but without altering the qualitative dynamics. Full details of the model and simulations are provided in SI Text, *Generic Model to Simulate the Dynamics of Hybridizing Annual Species*, *Generic model formulation* and *SI Appendix*.

We compared the population dynamics of the newcomer under two scenarios: (i) hybridization with an established species; and (ii) an established relative present but no hybridization. Fecundity and survival parameters were chosen to be constant and equal for the two species and for all hybrids to ensure no fitness advantage or disadvantage from hybridization. These assumptions can be altered according to the question being investigated. As a measure of establishment success, we estimated “Allee thresholds” as the minimum number of colonizing individuals required to ensure positive population growth for each scenario. A facilitation effect for hybridization would be apparent as a smaller Allee threshold than that observed for the no-hybridization scenario. We investigated the sensitivity of the model predictions over a wide range of parameter values related to initial population size, mating systems of the interacting genotypes, intercrossability, pollinator behavior, and pollination-plant density relationship (Table S2).

To verify the findings of the above theoretical model, a reduced version including only three genotypes, the two parental species, and a single hybrid class (34), was parameterized using empirical data from *Cakile maritima* and *C. edentula* (see Fig. S6 for example simulation). Parameter values related to fecundity were obtained from a common garden experiment. The selfing rates and cross-compatibility of the genotypes were quantified in two artificial crossing experiments (Fig. S1). Pollinator visitation rates were obtained by monitoring a population where the two species had similar relative frequencies. For hybrids, we used the average phenotypic values of the two parental species when no data existed. For all other parameters, we used published data and if no data were available, we ran the model over a wide range of parameter values to ensure consistent outcomes (e.g., Fig. S7). See SI Text, *Simplified Model for C. maritima and C. edentula* for detailed descriptions of the estimation of parameters.

ACKNOWLEDGMENTS. We thank Juliette Bouhours for advice on the development of the model and Sally Otto for comments on an earlier draft. Funding was from Australian Research Council Discovery Grant DP140100608 (to R.D.C., M.A.L., and K.D.). S.O. received a Victorian International Research Scholarship. C.L. was funded by the China Scholarship Council. M.A.L. acknowledges funding from the Natural Sciences and Engineering Research Council and a Canada Research Chair.

- Papagianni D, Morse MA (2015) *The Neanderthal Story: How Modern Science is Rewriting Their Story* (Thames & Hudson, London).
- Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proc Natl Acad Sci USA* 97(13):7043–7050.
- Hovick SM, Whitney KD (2014) Hybridisation is associated with increased fecundity and size in invasive taxa: Meta-analytic support for the hybridisation-invasion hypothesis. *Ecol Lett* 17(11):1464–1477.
- Simberloff D (2009) The role of propagule pressure in biological invasions. *Annu Rev Ecol Syst* 40:81–102.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Mol Ecol* 17(1):431–449.
- Taylor CM, Hastings A (2005) Allee effects in biological invasions. *Ecol Lett* 8:895–908.
- Keitt TH, Lewis MA, Holt RD (2001) Allee effects, invasion pinning, and species' borders. *Am Nat* 157(2):203–216.
- Elam DR, Ridley CE, Goodell K, Ellstrand NC (2007) Population size and relatedness affect fitness of a self-incompatible invasive plant. *Proc Natl Acad Sci USA* 104(2):549–552.
- Elliott SE, Irwin RE (2009) Effects of flowering plant density on pollinator visitation, pollen receipt, and seed production in *Delphinium barbeyi* (Ranunculaceae). *Am J Bot* 96(5):912–919.
- Cheptou P-O (2012) Clarifying Baker's Law. *Ann Bot (Lond)* 109(3):633–641.
- Huxel GR (1999) Rapid displacement of native species by invasive species: Effects of hybridization. *Biol Conserv* 89(2):143–152.
- Cousens RD, Ades PK, Mesgaran MB, Ohadi S (2013) Reassessment of the invasion history of two species of *Cakile* (Brassicaceae) in Australia. *Cunninghamia* 13:275–290.
- Barbour MG, Rodman JE (1970) Saga of the West Coast sea-rockets: *Cakile edentula* ssp. *californica* and *C. maritima*. *Rhodora* 72(791):370–386.
- Mills K (2013) Replacement of *Cakile edentula* with *Cakile maritima* in New South Wales and on Lord Howe Island. *Cunninghamia* 13:291–294.
- Boyd RS, Barbour MG (1993) Replacement of *Cakile edentula* by *C. maritima* in the strand habitat of California. *Am Midl Nat* 130(2):209–228.
- Busch JW, Delph LF (2012) The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Ann Bot (Lond)* 109(3):553–562.
- Rodman JE (1974) Systematics and evolution of the genus *Cakile* (Cruciferae). *Contrib Gray Herb Harvard Uni* 205:3–146.
- Thrall PH, Young AG, Burdon JJ (2000) An analysis of mating structure in populations of the annual sea rocket, *Cakile maritima* (Brassicaceae). *Aust J Bot* 48(6):731–738.
- Ohadi S, et al. (2015) Genetic structure along the strandline: Unravelling invasion history in a one-dimensional system. *J Biogeogr* 43(3):451–460.
- Rodman JE (1986) Introduction, establishment and replacement of sea-rockets (*Cakile*, Cruciferae) in Australia. *J Biogeogr* 13(2):159–171.
- Hyndman RJ, Mesgaran MB, Cousens RD (2015) Statistical issues with using herbarium data for the estimation of invasion lag-phases. *Biol Invasions* 17(12):3371–3381.
- Cody ML, Cody TWD (2004) Morphology and spatial distribution of alien sea-rockets (*Cakile* spp.) on South Australian and western Canadian beaches. *Aust J Bot* 52(2):175–183.
- Ollerton J, Winfree R, Tarrant S (2011) How many flowering plants are pollinated by animals? *Oikos* 120(3):321–326.
- Igic B, Kohn JR (2006) The distribution of plant mating systems: Study bias against obligately outcrossing species. *Evolution* 60(5):1098–1103.
- Burns JH, Ashman T-L, Steets JA, Harmon-Threalt A, Knight TM (2011) A phylogenetically controlled analysis of the roles of reproductive traits in plant invasions. *Oecologia* 166(4):1009–1017.
- Bufford JL, Daehler CC (2014) Sterility and lack of pollinator services explain reproductive failure in non-invasive ornamental plants. *Divers Distrib* 20(8):975–985.
- Jones KN (1997) Analysis of pollinator foraging: Tests for non-random behaviour. *Funct Ecol* 11(2):255–259.
- Ruhsam M, Hollingsworth PM, Ennos RA (2011) Early evolution in a hybrid swarm between outcrossing and selfing lineages in *Geum*. *Heredity (Edinb)* 107(3):246–255.
- Sweigart AL, Willis JH (2003) Patterns of nucleotide diversity in two species of *Mimulus* are affected by mating system and asymmetric introgression. *Evolution* 57(11):2490–2506.
- Morgan MT, Wilson WG, Knight TM (2005) Plant population dynamics, pollinator foraging, and the selection of self-fertilization. *Am Nat* 166(2):169–183.
- Doebeli M (1996) A quantitative genetic competition model for sympatric speciation. *J Evol Biol* 9(6):893–909.
- Barton NH (1992) On the spread of new gene combinations in the third phase of Wright's shifting-balance. *Evolution* 46(2):551–557.
- Ferdy JB, Austerlitz F (2002) Extinction and introgression in a community of partially cross-fertile plant species. *Am Nat* 160(1):74–86.
- Wolf DE, Takebayashi N, Rieseberg LH (2001) Predicting the risk of extinction through hybridization. *Conserv Biol* 15(4):1039–1053.
- Willmer P (2011) *Pollination and Floral Ecology* (Princeton Univ Press, Princeton).
- Tinbergen N, Impekoven M, Franck D (1967) An experiment on spacing-out as a defence against predation. *Behaviour* 28(3/4):307–320.
- Stephens PA, Sutherland WJ, Freckleton RP (1999) What is the Allee effect? *Oikos* 87(1):185–190.
- Goodwillie C, Ness JM (2013) Interactions of hybridization and mating systems: A case study in *Leptosiphon* (Polemoniaceae). *Am J Bot* 100(6):1002–1013.
- Hall RJ, Hastings A, Ayres DR (2006) Explaining the explosion: Modelling hybrid invasions. *Proc Biol Sci* 273(1592):1385–1389.
- Keddy PA (1981) Experimental demography of the sand-dune annual, *Cakile edentula*, growing along an environmental gradient in Nova Scotia. *J Ecol* 69(2):615–630.
- Feldman T, Morris WF, Wilson WG (2004) When can two plant species facilitate each other's pollination? *Oikos* 105(1):197–207.

Supporting Information

Mesgaran et al. 10.1073/pnas.1605626113

SI Text

Generic Model to Simulate the Dynamics of Hybridizing Annual Species.

Generic model formulation. We first construct a deterministic population model for a single annual species, of a single genotype, and then show how it can be extended to two species, to include more cooccurring genotypes that arise because of hybridization and later-generation backcrossing. Throughout we use population density and population size (or number) interchangeably, assuming that for a fixed area, changes in density are equal to changes in number.

Single species. Suppose that a population consists of N diploid seeds at time t , $N(t)$. Seeds germinate and are recruited into the population as seedlings and assumed to be spread evenly within the receiving patch of habitat. The proportion surviving to maturity is a function of both density-dependent and density-independent factors, collectively represented by $S(N)$. Each surviving plant has the potential to produce θ ovules in the absence of competition from other plants, but its per capita ovule production decreases with population density, following a negative density-dependent function $F(M)$, where M indicates number of adult plants: $M = NS(N)$. Each plant can self-fertilize a fraction s of its ovules, whereas the remaining $(1 - s)$ fraction is available for outcrossing. Because of pollinator or mate limitations, however, not all $1 - s$ will be outcrossed and some will be aborted. We assume that the probability of successful pollination is positively related to the population density (i.e., there is an Allee effect) following $\phi(M)$. A positive correlation between population density and pollinator visitation rate has been observed in many studies and has been implemented in population dynamics models (35). The difference equation with both negative (ovule production and survival) and positive (pollination success) density regulation for a single species' population dynamics is therefore given by

$$N(t+1) = M(N(t))\theta F(M(N(t))) [s + (1-s)\phi(M(N(t)))]. \quad [\text{S1}]$$

For density-dependent and density-independent recruitment of seeds into the population, we use a Beverton–Holt type model:

$$S(N) = \frac{L}{1 + bN}, \quad [\text{S2}]$$

where L denotes the proportion of seeds recruited independent of population density, whereas b is the rate of decrease in survivorship of a plant as density increases. We used a rectangular hyperbola to describe ovule reduction as related to density of mature plants, M :

$$F(M) = \theta \left[1 - \frac{dM}{1 + dM} \right], \quad [\text{S3}]$$

where d is a coefficient: the larger the value of d , the stronger the negative effect of density on reproduction. Given that $M = NS(N)$, Eq. S3 can be written as a function of N :

$$F(M(N)) = \theta \left[1 - \frac{\frac{LdN}{1+bN}}{1 + \frac{LdN}{1+bN}} \right], \quad [\text{S4}]$$

which can be simplified to

$$F(M(N)) = \frac{\theta(bN + 1)}{N(b + Ld) + 1}. \quad [\text{S5}]$$

For pollination probability, or “pollinator foraging choice” (30), we used Holling type III functional response to relate the probability of pollination to population size:

$$\phi(M) = \frac{(\rho M)^\gamma}{1 + (\rho M)^\gamma}, \quad [\text{S6}]$$

and, again,

$$\phi(M(N)) = \frac{\left[\frac{\rho NL}{1 + bN} \right]^\gamma}{1 + \left[\frac{\rho NL}{1 + bN} \right]^\gamma}. \quad [\text{S7}]$$

This type of functional response is expected for generalists that frequently switch between flower types (35) or if pollinators engage in an area-restricted search strategy (36). The reciprocal of ρ gives the population size at which 50% of plants will be visited (and pollinated) by pollinators. The parameter γ determines the shape of the curve: when γ is large, there is little change in probability of pollination at low densities compared with a small γ , which gives a higher pollination rate at the same low densities. When $\gamma = 1$, the model reduces to a type II or saturating functional response. Eq. S7 can be simplified to $\phi(M(N)) = 1 - \frac{1}{1 + [\rho NL / (1 + bN)]^\gamma}$, and Eq. S1 becomes

$$N(t+1) = \underbrace{\left(\frac{N(t)L}{1 + bN(t)} \right)}_{\text{No. mature plants}} \underbrace{\left(\frac{\theta(bN(t) + 1)}{N(t)(b + Ld) + 1} \right)}_{\text{Per capita fecundity}} \times \left[s - (1-s) \underbrace{\left(1 - \frac{1}{1 + \left[\frac{\rho N(t)L}{1 + bN(t)} \right]^\gamma} \right)}_{\text{Probability of pollination}} \right]. \quad [\text{S8}]$$

This model results in component Allee dynamics (37) by negatively relating recruitment and per capita fecundity to population density, while positively scaling pollination probability.

Two species. Now suppose that an equally competitive congeneric species already occurs in the habitat (we will call this a resident, which could be either a native or a previously established invader) and can interact with the above species in both positive and negative ways. The assumption is made that both species are homogeneously distributed across space. A negative interaction occurs because of competition between the two species for limited abiotic resources and is reflected in their reproduction, $F(M(N))$, and survival, $S(N)$, functions. The intensity of competition, embedded within $F(M(N))$ and $S(N)$, is assumed to depend on the total number of both species regardless of the species identity. This equality in competitiveness of species is required to ensure that any difference between hybridization and nonhybridization scenarios is solely attributable to hybridization and not to an a priori fitness differential. However, the model can accommodate species inequality in any traits, as shown in Eq. S16. A positive interaction can occur when there is a positive relationship

between the chance of pollination and the *total* population size, such that M on the right hand side of Eq. S6 is replaced by the combined number of individuals from the two species. However, the two species may also compete for pollinators and, depending on their relative attractiveness to pollinators, may experience differential outcrossing success, accounted for by the V_x function (see below). Using subscript x , where $x = 0$ denotes the invader and $x = 1$ indicates the resident species, the difference equation for two interacting species is given by

$$N_x(t+1) = M_x(\underline{N}(t)) F_x \left(\sum_{k=0}^1 M_k(\underline{N}(t)) \right) \times \left[s_x - (1 - s_x) V_x(\underline{N}(t)) \phi \left(\sum_{k=0}^1 M_k(\underline{N}(t)) \right) \right], \quad [\text{S9}]$$

where \underline{N} is the total density summed across all genotypes; $M_x(\underline{N}) = N_x S(\sum_{k=0}^1 N_k)$ with $k = 0$ (invader) and $k = 1$ (resident). $V_x(\underline{N})$ is the proportion of pollinator visits to species x . If both species are equally preferred by pollinators (i.e., visited randomly) then the likelihood of a visit to any given species is proportional to its frequency in the population, i.e., $V_x(\underline{N}) = M_x(\underline{N}) / \sum_{k=0}^1 M_k(\underline{N}) = R_x$, where R_x is the relative frequency of adult plants. However, the two species may differ in their attractiveness such that pollinators visit W plants of the invasive species for every one plant of the resident species. Then the proportion of visits to the invasive species is

$$V_o = \frac{WR_0}{WR_0 + R_1}, \quad [\text{S10}]$$

and to the resident is

$$V_1 = \frac{R_1}{WR_0 + R_1} = 1 - V_o, \quad [\text{S11}]$$

where R_0 and R_1 are the relative frequencies of invader and resident plants in the mixed population ($R_0 + R_1 = 1$), whereas V_o and V_1 are adjusted frequencies after accounting for the bias in pollinator preference (see Fig. S1 for a graphical explanation of the pollination submodel). When the two species hybridize and produce viable offspring, the population will consist of three genotypes in the next generation: two parental pure species plus the F_1 hybrid, so $x = \{0, 0.5, 1\}$. We use $x = 0.5$ to indicate that the hybrid genotype is intermediate relative to its parents. In the later generations, more intermediate genotypes will be produced through backcrossing and segregation (accounted for in Eq. S12), so x can then take any values between 0 and 1. However, as a workable assumption, we consider variable x as a discrete variable with G genotypic classes. We use the number of chromosomes as a proxy for determining G (the total number of genotype classes) and the chromosome combination of the two parental species to characterize x . That is, suppose that species A (e.g., invasive) and B (resident) both have $2n = 18$ chromosomes and can freely cross with each other. Individuals from this mixed population can then be categorized in 19 genotype classes ($G = 2n + 1$) based on the number of chromosomes that they receive from the two species [i.e., (18,0), (17,1), ..., (1,17), (0,18)]. The first element in (\bullet, \bullet) indicates the number of chromosomes from species A , A_n , whereas the second element is the number of chromosomes from species B , B_n . An individual with an (18,0) chromosome combination is regarded as a pure individual of species A , whereas (0,18) denotes a pure individual of species B . We map chromosome combinations to variable x using $x = A_n / 2n$, which gives $x = 1$ for species A and $x = 0$ for species B (the use of A_n in the numerator is arbitrary). In this paper, we assign $x = 0$ to the incoming species and $x = 1$ to the resident

species. Crossing-over can result in more intermediate genotypes, where a single chromosome can comprise various proportions of the genome from the two species, but for simplicity, we do not consider this possibility in our model. It should not affect our qualitative conclusions.

A diploid genotype with a total of ℓ alien chromosomes ($\ell \leq n$ and $\ell \neq 0$) undergoing gametogenesis can produce $\ell + 1$ haploid gametes (an F_1 hybrid which has n alien chromosomes can produce the maximum number of gamete types, which is $n + 1$). A gamete can possess ν alien chromosomes where ν range from 0 up to ℓ (i.e., $\nu \in \{0, \dots, \ell\}$) with probability given by

$$P(\nu|\ell) = \frac{\binom{n}{\nu} \binom{n}{\ell-\nu}}{\sum_{i=0}^{\ell} \binom{n}{i} \binom{n}{\ell-i}}. \quad [\text{S12}]$$

This formula is based on the number of ways that chromosomes can line up at the cell equator during the metaphase of the first meiotic division. As shown in Fig. S2, with $\ell = 4$, various arrangements are possible: (i) all alien chromosomes (red) can occur on the same strand, (ii) one alien chromosome on the left strand and three on the right, and (iii) two alien chromosomes on either side. Within each of these arrangements, many different combinations (i.e., positioning of aliens along the native strand) can take place. Indeed, $\binom{n}{\nu}$ in the numerator of Eq. S12 gives the number of possible ways that ν alien chromosomes (out of ℓ) can be positioned on the left strand, whereas $\binom{n}{\ell-\nu}$ gives the number of positions for the remaining alien chromosomes (i.e., $\ell - \nu$) on the right strand. Note that when the number of alien chromosomes on the left and right strands are the same (i.e., $\nu = \ell - \nu$), the number of combinations for that gamete type doubles. The denominator in Eq. S12 simply gives the total number of all possible combinations.

Now, suppose that a female genotype y with ℓ_y alien chromosomes is crossed with a male genotype z with ℓ_z alien chromosomes. We obtained the probability of producing offspring x (with $\forall \ell_x \in \{0, \dots, L\}$, where $L = \ell_z + \ell_y$) from parents with genotypes y and z using a discrete convolution:

$$H(x:y, z) = \sum_h y_h z_{x-h+1}. \quad [\text{S13}]$$

In this formula, $y = P(\nu|\ell_y)$ and $z = P(\nu|\ell_z)$ are two vectors containing the probabilities for gamete types that can be produced by parents y and z , respectively, following Eq. S12, and h ranges over all legal subscripts for y_h and z_{x-h+1} , specifically, $h = [\max(1, x + 1 - L), \dots, \min(x, L)]$. Note that we transform ℓ_x to x using $x = 1 - \ell_x / 2n$, which represents the proportion of alien chromosomes within the genome ranging from 0 (pure invasive species) to 1 (pure resident species). Our genetic model gives rise to similar qualitative outcomes as the hypergeometric phenotypic model (31, 32). The difference equation model that allows for hybridization then takes the form:

$$N_x(t+1) = \sum_{y=1}^G M_y(\underline{N}(t)) \theta_y F_y \left(\sum_{k=1}^G M_k(\underline{N}(t)) \right) \left[s_y H(x:y, y) - (1 - s_y) C(y, z) V_y \phi \left(\sum_{k=1}^G M_k(\underline{N}(t)) \right) H(x:y, z) \right], \quad [\text{S14}]$$

where $G = 2n + 1$ indicates the total number of genotypes in the population. The function $C(y, z)$ in Eq. S14 indicates the compatibility (i.e., the likelihood of producing viable offspring

once pollinated) of female genotype y with the male genotype z and is calculated using

$$C(y, z) = 1 - (\alpha - 1)[y^\beta - 2y^\beta z^\beta + z^\beta], \quad [\text{S15}]$$

where parameter α indicates the compatibility between the two parental species, whereas β introduces asymmetry in the direction of backcrossing (asymmetric compatibility). The model (Eq. S15) assumes that the compatibility between two genotypes increases as these two genotypes come to share more genome from the same species. For the specific cases of $y = z = 0$ (cross between two pure individuals of the incoming species) and $y = z = 1$ (cross between two pure individuals of the established species), the model gives $C(y, z) = 1$ but $C(y, z) \neq 1$ if $y = z \neq 0$ or $y = z \neq 1$. Note that $C(y, z) = 1$ always holds if $\beta = 0$ or $\alpha = 1$. When $\beta > 1$, F_1 and later generation hybrids are more compatible with the incoming-type individuals, whereas for $\beta < 1$ compatibility is biased toward the resident species and for $\beta = 1$ introgression is symmetric (Fig. S3). Studies with both natural and artificial crossing have shown that the intercrossing success rate can be affected by the direction of crossing (38), and our model accounts for such an asymmetry by incorporating the bias parameter β .

To relate genotype x to a phenotypic character T (any model parameter), we assumed that trait values (e.g., selfing rate) are normally distributed among genotypes, following Hall et al. (39) with minor modifications:

$$T(x) = T_{\max} \exp\left(-\frac{(x - x_{\max})^2}{2\sigma^2}\right), \quad [\text{S16}]$$

where T_{\max} is the maximum value of a given trait, whereas x_{\max} is the genotype which acquires that maximum trait value. The parameter σ^2 represents the variance, which can be solved to give the minimum value of T for a specific genotype. For example, we may choose to assign the highest selfing rate to genotype $x = 0$ (newcomer species) and then calculate σ^2 such that the genotype $x = 1$ (resident species) obtains the lowest selfing rate among all genotypes (Fig. S4).

A list of model parameters is given in Table S1. To summarize, some of the assumptions of our model are as follows:

- i) There is no persistent seedbank.
- ii) Both density-dependent and density-independent factors act at the same time to determine the survival rate. For example, if disturbance occurs first whereby many seedlings die, then the density-dependent mortality would be minimal.
- iii) Self-fertilization occurs first, and the remaining flowers then out-cross.
- iv) Pollination success is only determined by population size and attractiveness of genotypes (the frequency of pollen from any genotype in the pollen pool is proportional to the relative abundance of adult plants of that genotype in the mixed population and the *per capita* pollen production is therefore the same for all genotypes).
- v) There is no difference between chromosomes when mapping genotype to phenotype (i.e., it is the number of chromosomes that matters, not their type)
- vi) There is no crossing over.
- vii) Self-fertilization incurs no inbreeding depression. Inbreeding depression could easily be incorporated into the model as another multiplier (for example, of selfing rate s) in Eq. S1 and its multigenotype derivation (Eq. S14), but the results would not change qualitatively.

Simulations. We simulated the population dynamics of 19 genotypic classes corresponding to a diploid species with $2n = 18$ over 100 generations. All simulations were run using Matlab (version 2014a);

code is given in *SI Appendix*. Table S2 provides details of parameter values (and range) used for various simulations related to the effect of pollinator behavior, mating system, compatibility between genotypes under hybridization, and no-hybridization scenarios.

Simplified Model for *C. maritima* and *C. edentula*.

Model assumptions. To test the hypothesis that hybridization could have facilitated the invasion of *C. maritima* without enhancing fitness, we used a reduced version of the generic model. This simplified model simulates the invasion dynamics of only three genotype classes (34): *C. maritima*, *C. edentula*, and their F_1 hybrid. The variable x thus takes three values: 0 for *C. maritima*, 1 for *C. edentula*, and 0.5 for hybrids. The main reason for reducing the number of genotypes is that there is no information on the demography and biology of intermediate genotypes (later generation hybrids). Furthermore, this simplification makes the model more tractable without affecting the qualitative conclusions. We used intercrossing compatibility rates obtained from artificial hybridization experiments rather than using Eq. S15.

The model was parameterized using both the literature and our own experimental or observational data; for a few parameters, for which there were no estimates, we evaluated a range of plausible parameter values to test whether their precise values affect the qualitative results. If no parameter estimate was available for the hybrid genotype, we used the average of the two parental species, because hybrids tend to be intermediate in most characters relative to their parents (17). See Table S3 for a list of parameter values and their sources: although some parameters are in favor of *C. maritima* (e.g., higher fecundity and survival) compared with *C. edentula*, this inherent fitness advantage alone is insufficient for *C. maritima* to drive a well-established *C. edentula* population to extinction (Fig. 4).

Initial population sizes. Almost no data exist on the number of arriving propagules in accidental introductions (4); clearly, the number could be as few as one or it could be several orders of magnitude greater, depending on the dispersal vector and chance. We have no information on the founding population sizes (number of seeds arrived) of *C. maritima* when the species was first introduced to Australia, or even for its dispersal from one beach to the next during subsequent spread. However, we know that *C. edentula* had time to saturate most of its preferred habitats by the time *C. maritima* arrived as the second invader. The number of *Cakile* adult plants that can be found on a beach is highly variable, so it is equally hard to say what a plausible carrying capacity would be. Highly disturbed beaches may function as sink populations and never support more than a few individuals, whereas a sheltered beach may be a source population of many thousands of plants and a net exporter of seeds. In our simulations, we fixed the carrying capacity of a beach at 500 adult plants (regardless of the species; this value is within the range of our field observations for small beaches) and then introduced 1–600 seeds of *C. maritima* (with unit increment).

We simulated the dynamics of the three genotypes over 40 generations. To estimate the Allee threshold for *C. maritima*, we compared the number of *C. maritima* seeds after 40 generations with the initial size of the founding population: the minimum initial seed number that resulted in a positive population growth was taken to be the Allee threshold. When *C. edentula* was present on the beach, its population size was set at the carrying capacity (i.e., ~500 individuals). However, we tested situations other than an equilibrium state, wherein *C. edentula* occurred with a lower number of individuals as if its habitat had recently been disturbed (Fig. S5). The results indicate that hybridization could facilitate the establishment of a small (otherwise sub-Allee threshold) population of the primarily self-incompatible *C. maritima* (Fig. S5).

Sources of data for parameterization. The number of flowers (θ) produced by each species and their hybrids was obtained from a common garden experiment in pots containing a sandy soil.

Plants were derived from seeds of *C. maritima* and *C. edentula* and the progeny of putative hybrids (later confirmed as hybrids by simple sequence repeat markers) collected from five, six, and two locations, respectively, across their invaded ranges in Australia. The number of flowers produced by each plant over its lifetime was determined when plants were no longer flowering, by counting the number of scars and pedicels left on the racemes. *Cakile maritima*, *C. edentula*, and putative hybrids produced an average of 1,195 (SE = 72; $n = 54$), 906 (SE = 59; $n = 63$), and 1,486 (SE = 128; $n = 38$) flowers, respectively. These numbers are within the ranges estimated in beach populations in Tasmania.

Data on intercrossing compatibility were determined in two glasshouse experiments, where we first produced F_1 hybrids and then established a series of reciprocal crosses between the F_1 hybrids and their parental species. The probability of producing viable offspring when crossed was calculated from the number of viable (germinable) seeds divided by the total number of crosses made between genotypes. Selfing rates were determined by excluding pollinators. No viable seeds were produced when *C. maritima* was selfed by hand, although a very small number of putative spontaneous selfs occurred on caged plants; it was therefore not possible to measure inbreeding depression. Fig. S6 shows the average compatibility for the various conspecific crossings implemented in these experiments; as can be seen, backcrosses are almost equally compatible in either direction. However, hybrids are more likely to backcross with *C. maritima* than with *C. edentula* because they have larger flowers than *C. edentula*: in our common garden experiment, mean petal dimensions (displayed section of petal only) were 5.42×2.01 mm for putative hybrids, 6.19×2.45 mm for *C. maritima*, but only

4.35×1.19 mm for the inbreeder *C. edentula*, which, on some occasions, may not even produce all four petals (17).

We monitored pollinator visits to 47 *C. maritima* and 37 *C. edentula* plants on a beach in Tasmania where both species were approximately in equal abundance. *C. maritima* has a more conspicuous display, with more open flowers per raceme and larger flowers (see above). Both species were mainly visited by honey bees (*Apis mellifera*); other common visitors were European bumblebees (*Bombus terrestris*) and cabbage white butterflies (*Pieris brassicae*). Although hybrids and their back-crosses may have been present, we chose plants that had all of the characteristics of their assumed parent species. We recorded the time from the start of recording until the first visitor and the number of flowers subsequently visited on that plant. We analyzed these time-to-event data with the Proc Reliability procedure of SAS to fit a lognormal distribution (version 9.3; SAS Institute). The mean times to first visit were 110 s for *C. maritima* and 830 s for *C. edentula*, implying that *C. maritima* was visited by pollinators 7.5 times as frequently as *C. edentula*. Pollinators also visited more flowers per visit on *C. maritima* than *C. edentula*: using a Poisson distribution, the average number of flowers visited per pollinator was 11.7 for *C. maritima* and 6.2 for *C. edentula*.

All details of experiments outlined above can be obtained from R.D.C. For ρ and γ , parameters that describe the relationship between population density and probability of pollinator visit, we investigated the sensitivity of the model predictions over a wide range of parameter values (Table S3), because we had no empirical estimates for these two parameters. As shown in Fig. S7, a model with hybridization always gave smaller Allee thresholds over all tested ranges of these two parameters (Fig. S7).

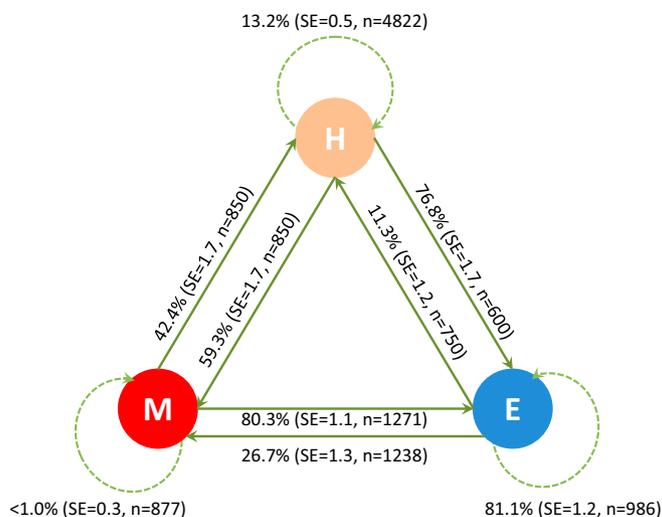
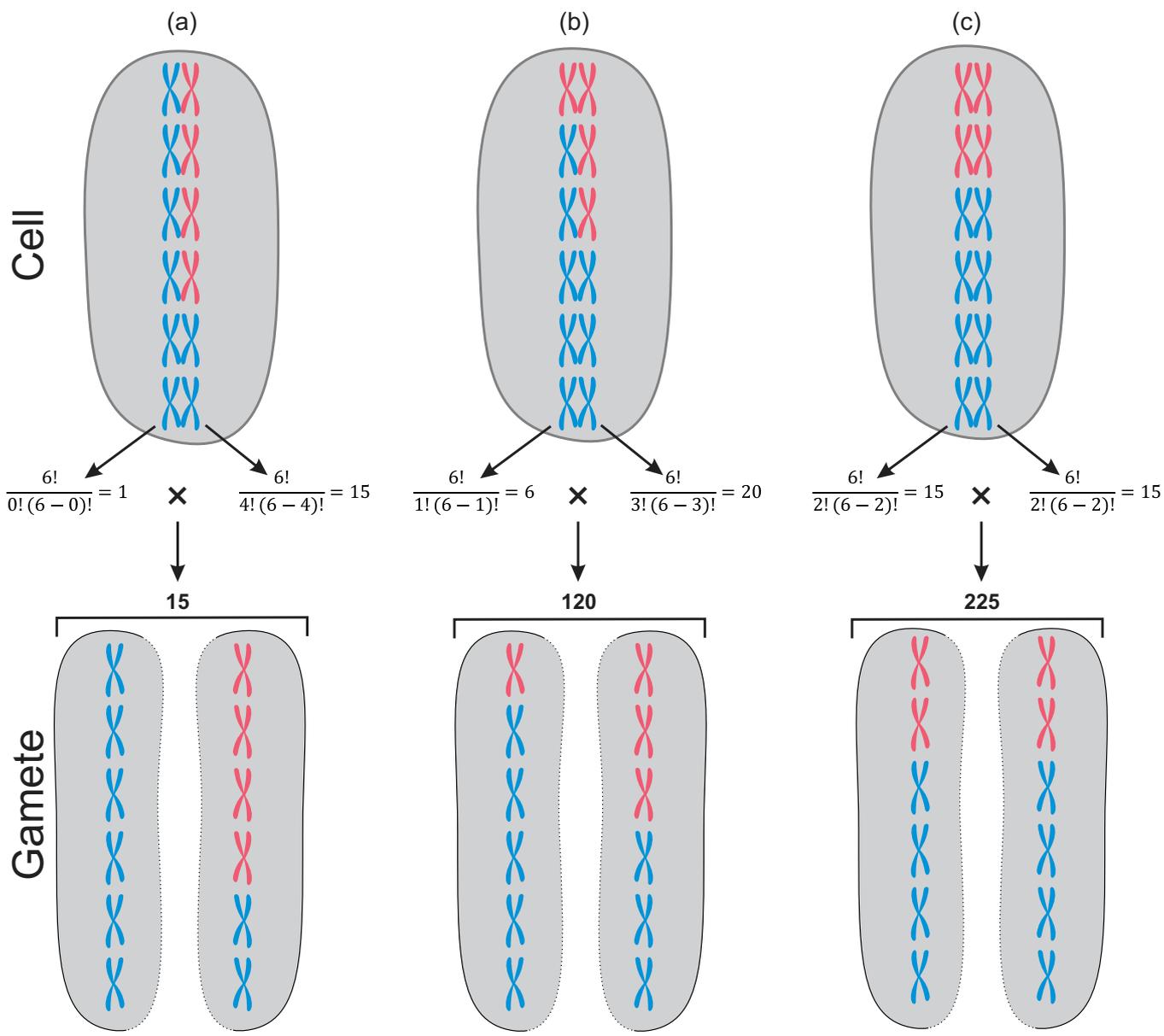


Fig. S1. Probability of successful crossing (i.e., viable seed production) for different directions of hand-crossing. The arrow shows the direction of pollen transfer from a donor to a recipient plant: *C. maritima* (M), *C. edentula* (E), and F_1 hybrids (H). Spontaneous self-fertilization rates are shown as dashed lines. n indicates number of crosses.



No. alien chromosomes	No. combinations	Probability
0	15	$15/720 = 0.0208$
1	120	$120/720 = 0.1667$
2	2×225	$450/720 = 0.6250$
3	120	$120/720 = 0.1667$
4	15	$15/720 = 0.0208$
Sum	720	1.0

 Native
 Alien

Fig. S4. Some chromosome arrangements for a hypothetical diploid genotype with $2n = 12$ where eight chromosomes descend from the resident species (blue) and four ($\ell=4$) from the alien species (red). All alien chromosomes may be located on a single strand (A) or occur on the two strands (B and C). The number of ways that alien chromosomes can be positioned within each strand is given by $\binom{n}{\nu}$ and $\binom{n}{\ell-\nu}$ for left and right strands, respectively (Eq. S12), and the total number of combinations is the product of these two values. For example, there are 15 different ways to position 4 alien chromosomes on the right strand while having no alien chromosome on the left strand. This arrangement will result in two gamete types (one with no alien chromosomes and one with four) with the equal probability of 0.0208. The inset table provides the probability values for each of the five gamete types (i.e., $\nu \in \{0, \dots, 4\}$) that can be produced by a diploid genotype with $n = 6$ and $\ell = 4$.

Table S1. List of parameters used in the generic model

Parameter	Definition	Range
L	Density-independent survival (i.e., fraction of seeds that germinate and survive to maturity)	$L \in [0,1]$
B	Density-dependent coefficients for seedling survivorship to maturity	$b \in \mathbb{R}^+$
θ	Per capita flower production	$\theta \in \mathbb{R}^+$
d	Density-dependent coefficients for flower production reduction	$d \in \mathbb{R}^+$
s	Selfing rate (i.e., proportion of flowers that produce seeds through self-fertilization)	$s \in [0,1]$
ρ	Pollination switching density parameter: $1/\rho$ gives the population size at which the chance of pollination is 50%	$\rho \in \mathbb{R}^+$
γ	Switching exponent parameter in pollination equation (shape parameter): when $\gamma=1$, the Eq. S6 will collapse to type II functional response	$\gamma \in \mathbb{R}^+$
α	Compatibility of the two parental species when they intercross	$\alpha \in [0,1]$
β	Compatibility bias parameter	$\beta \in \mathbb{R}^+, \beta \neq 0$
n	No. of chromosomes in haploid	$n \in \mathbb{N}_1$
ℓ	Total no. of alien chromosomes in a diploid genotype	$\ell \leq n$
ν	No. of alien chromosomes in a haploid gamete	$\nu \in \{0, \dots, \ell\}$
x	Index variable indicating the identity of a genotype (e.g., $x=0$ represents the newcomer species, whereas $x=1$ denotes the resident species). Total no. of elements (genotypes) in x is $2n+1$.	$x \in \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-2}{n}, \frac{n-1}{n}, 1\right\}$
R	Relative frequency of a given genotype in the population	$R_x \in [0,1]$
V	Relative frequency of a given genotype in the population after adjustment for pollinator preference	$V_x \in [0,1]$
W	A weight parameter that adjusts the relative frequency of genotype x according to its degree of attractiveness to pollinators	$W_x \in \mathbb{R}^+$
y	An index variable indicating the identity of a female genotype (i.e., pollen recipient)	As x
z	An index variable indicating the identity of a male genotype (i.e., pollen donor)	As x
T_{max}	Maximum value of a given trait (model parameter)	Variable
x_{max}	The genotype to which the maximum value of a trait (T_{max}) is assigned	As x
σ^2	Variance of a phenotypic trait across genotypes	$\sigma^2 \in \mathbb{R}^+$

Table S2. Parameter values (and ranges) used for simulations of the generic model

Parameter	Value/range			
	Invasion dynamics*	Pollinator preference [†]	Bias in compatibility [†]	Mating system [†]
L_x	0.1	0.1	0.1	0.1
b_x	0.0001	0.0001	0.0001	0.0001
θ_x	1,000	1,000	1,000	1,000
d_x	0.1	0.1	0.1	0.1
s_x	0.01	0.01	0.01	$s_0 = [1...99\%]$ $s_1 = [1, 20, 40, 60, 80, 99\%]$
ρ_x	0.035	0.035	0.035	0.035
γ_x	3.5	3.5	3.5	3.5
α	0 and 0.5	0 and 0.5	0 and 0.5	0 and 0.5
β	1	1	[1...10]	1
n	18	18	18	18
W_0	5	[1...10]	1	1
N_0	50	[1...300]	[1...300]	[1...300]
N_1	1,950	300	300	300

Parameter values used for sensitivity analysis are shown in boldface.

*Parameter setting used for the invasion scenarios presented in Fig. 1.

[†]Parameter setting used for estimating Allee thresholds presented in Fig. 2.

[‡]Parameter setting used for estimating Allee thresholds presented in Fig. 3.


```

%% Example: The below code will produce Figure 1B as in the main text of paper.
% Make sure all four "functions" are in the same folder. These functions are:
% 1- "hybrid_facil.m" (this is the main function that excutes the model: you only need to work this
% function);
% 2- "crossing.m" (this function gives the identity of offspring genotypes);
% 3- "geno2pheno.m" (this function maps genotype to phenotype);
% 4- "selfingrate.m" (this function uses "geno2pheno.m" function to gives the selfing rate of genotypes)

% Mohsen B Mesgaran (mohsenm@unimelb.edu.au): May 2016
%%
clear % clear memory
close all % close all figures;
%% Parameters: see "hybrid_facil" function for a description of all the listed parameters
n_chromosome = 9;
n_colonizer = 50;
n_resident = 1950;
L = 0.1;
b = L/1000;
d = 0.1;
theta = 1000;
rho = 0.035;
gamma = 3.5;
s_0 = 0.01;
s_1 = 0.01;
x_max = 0;
W_max = 5;
alpha = 0.5;
beta = 1;
time = 99;
n = 2*n_chromosome+1; % number of genotypes
%% Run the model
SeedBank =
hybrid_facil(n_chromosome,n_colonizer,n_resident,L,b,d,theta,rho,gamma,s_0,s_1,x_max,W_max,alpha,beta,time);
SeedBank = [[n_colonizer; zeros(n-2,1); n_resident]'; SeedBank]; % Add initial population size to the matrix SeedBank
RF=bsxfun(@x./y,x./y,SeedBank,sum(SeedBank,2)); % convert numbers to relative frequency

%% Figure configuration
set(0,'units','centimeter')
scsz = get(0,'screensize');
figWidth = 15;
figHeight = 15;
fc = figure('Units','centimeters','Position',[(scsz(3)-figWidth)/2 (scsz(4)-figHeight)/2 figWidth figHeight], 'Color', [1 1 1]);
%% Color for bars (color-coding genotypes)
red = [ones(1,fix(n/2)); linspace(0,0.9,fix(n/2)); linspace(0,0.9,fix(n/2))];
blue = [linspace(0.9,0,fix(n/2)); linspace(0.9,0,fix(n/2)); ones(1,fix(n/2))];
col2 = [red; [0.97 1 .97]; blue];
col3=col2*.9;
%% plot the 3D bar graph
b1 = bar3(RF,1);
ylim([1 time+2])
xlim([0.25 19.75])
zlim([0 1])

```

```

for i=1:n
set(b1(i),'facecolor',col3(i:),'EdgeColor', col2(i:),'LineWidth',0.05);
end
set(gca,'FontSize',8,'YTick', [1 20:20:100], 'XTick', 1:n,'XTickLabel',...
    {'100' " " " '75' " " " " '50' " " " " '25' " " " '0'}, 'XDir','normal');
axis square
x2 = xlabel('Genome derived\nnewlinefrom colonizer (%)','HorizontalAlignment' , 'center');
set(x2,'rotation',-26,'FontSize', 11);
x2.Position = [-6.5719  29.4396  -0.8152];
y2 = ylabel('Time (generations)');
set(y2,'rotation',34,'FontSize', 11);
y2.Position = [4.5  -10.6476  -1.0063];
z1 = zlabel('Relative frequency','FontSize', 11);
z1.Position = [-16.9728  27.1791  -0.2612];
ax = gca;
grid off
view(42,34);
annotation('textbox','String', 'Population dynamics of 19 genotypic classes over time as result of hybridization between
colonizing and resident species',...
    'LineStyle', 'none','FontWeight','bold','FontSize', 11,'Position',[.1 .85 .75 .05],'HorizontalAlignment','center');
%% Save figures in three different formats
set(fc,'PaperUnits','centimeters','PaperSize',[figWidth figHeight],'PaperPosition',[0 0 figWidth
figHeight],'PaperPositionMode' , 'manual');
print(fc,'Fig1_PNAS', '-depsc' , '-painters');
print(fc,'Fig1_PNAS', '-dpdf' , '-painters');
print(fc,'Fig1_PNAS', '-dpng', '-r300' );

```

```

function SeedBank =
hybrid_facil(n_chromosome,n_colonizer,n_resident,L,b,d,theta,rho,gamma,s_0,s_1,x_max,W_max,alpha,beta,time)
% n_chromosome = number of haploid chromosomes
% n_colonizer = initial population size of colonizer genotype
% n_resident = initial population size of resident genotype
% s_0 = selfing rate in colonizer [0,...,1]
% s_1 = selfing rate in resident [0,...,1]
% L = the proportion of seeds recruited independent of population density (Eq.2 in Supplementary Information)
% b = mortality rate due to density: b/L gives the max number of adult plants (sort of carrying capacity for adult plant
stage not seed stage: see Eq.2 in Supplementary Information)
% theta = the maximum number of ovules/flowers produce by a single plant (Eq.3 in Supplementary Information)
% d = rate of decrease in ovule/flower production due to density (Eq.3 in Supplementary Information)
% rho = location parameter in pollination function: 1/rho gives the plant density at which the probability of pollination
is 50% (Eq.6 in Supplementary Information)
% gamma = shape parameter in pollination function (Eq.6 in Supplementary Information)
% alpha = compatibility between parental genotypes i.e. colonizer*resident (Eq.15 in Supplementary Information).
When alpha = 0 the whole model is reduced to a "competition model" with two species
% beta = bias parameter in backcrossing (Eq.15 in Supplementary Information)
% time = number of generations/years
% x_max = genotype with the highest value of attractiveness (see Eq.16 in Supplementary Information)
% W_max = attractiveness of the most attractive genotype (that of the least will be 1). if W_max = 1, then all "n"
genotypes will have the same attractiveness value of 1

% Mesgaran B Mesgaran (mohsenm@unimelb.edu.au): May 2016
%% Genotyping
n = 2*n_chromosome+1; %number of genotypes
N_x = [n_colonizer; zeros(n-2,1); n_resident]'; % number of seeds from genotype x
x = linspace(0,1,n); % Vector of genotype ranging from 0 to 1 with n classes where 0 indicates colonizer and 1 resident
[Sx , Sx_x] = selfingrate(s_0,s_1,n); % this function gives the selfing rate of all "n" genotypes (based on Eq.16 in
Supplementary Information) in vector "Sx" and n replicates of them in "Sx_x"
[cross_offs , self_offs] = crossing(n); % this function gives the identity of genotypes resulted from all pairwise crossings
between genotypes in the matrix "cross_offs" with n^2 by n dimension while the identity offspring from selfing is given
in "self_offs" whcih has a dimension n*n (this function is based on E.12 and Eq. 13 in Supplementary Information)
xx = repmat(x,1,n); % repeating the vector "x" for pairwise crossings
z = sort(xx(:,1),1); % vector of male plants
y = xx(:); % vector of female plants
W_x = geno2pheno(x_max,W_max,1,n); % this function maps genotype to phenotypic (here attractiveness) using Eq.
16 in Supplementary Information

% the start of the loop
for j=1:time
%% Survival to maturity (flowering stage)
N = sum(N_x(:)); % total number of seeds summed over all genotypes
S_N = L/(1+b*N); % Survival function(Eq.2 in Supplementary Information)
M_N = S_N*N_x; % number of mature plants
M_N2 = [reshape(repmat(M_N,n,1),n*n,1) repmat(M_N',n,1) ]; % repeating the vector "M_N" for pairwise crossings

%% Adjusting relative frequencies for pollinator preference (attractiveness)
M_total = sum(M_N(:)); % total number of adult plants
R_x = M_N./M_total; % unadjusted (for pollinator preference) relative frequency of each genotype
V_x = R_x.*W_x./sum(R_x.*W_x); % adjusted (for pollinator preference) relative frequency of each genotype (Eq.10
in Supplementary Information)
RV_x = repmat(V_x,1,n); % repeating the vector "V_x" for pairwise crossings

```

```

RV_x=reshape(RV_x.',[],1); % reshape it to get the proper column vector

%% Pollination
phi_M = (rho*M_total)^gamma/(1+ (rho*M_total)^gamma); % pollination functional response relating probability of
pollination to total plant density (Eq.6 in Supplementary Information)
phi_M_x = phi_M*RV_x; % probability of pollination for genotype x while accounting for the for pollinator preference

%% Probability of out-crossing
out_cross_x = phi_M_x.*(1-Sx_x); % while "s" fraction of ovules get fertilised through selfing the remaining "(1-s)"
need to be fertilized through pollination

%% Compatibility
C_yz = z.^beta.*(alpha - 1) - y.^beta.*(2*z.^beta.*(alpha - 1) - alpha + 1) + 1; % the compatibility of female genotype y
with the male genotype z (Eq.15 in Supplementary Information)
fertile_out_cross_x = out_cross_x.*C_yz ; % the probability of fertile outcrossing between genotypes after
accounting for probability pollination (phi_M_x), non-selfing fraction (1-s) and compatibility (C_yz)

%% Flower production
F_M = 1-(d*M_total)/(1+(d*M_total)); % reduction in capacity of ovule production due to density-dependent
competition (Eq.3 in Supplementary Information)
theta_M_x = F_M*(M_N2(:,2)).*theta; % total number of ovules produced by adult plants of genotype x

%% Outcrossed seed production
out_seed = theta_M_x.*fertile_out_cross_x; % number of seeds produced on genotype x through outcrossing (note that
these are not seeds with genotype x, the identity of these seeds need to be determined by function "cross_offs" which is
based on Eq.12 and Eq.13 in Supplementary Information)

%% Selfed seed production
theta_M_x = F_M*(M_N.*theta); % total number of ovules produced by adult plants of genotype x, here M_N which is
a vector of length n is used not M_N2 which has a length of n^2 (the values are the same but as in the previous section
we were dealing with outcrossing we needed to have all pairwise combinations)
self_seed = (theta_M_x.*Sx'); % number of seeds produce on genotype x through outcrossing (note that these are not
seeds with genotype x, the identity of these seeds need to be determined by function "self_offs" which is based on Eq.12
and Eq.13 in Supplementary Information)

%% Total seed production
out_seed_x = out_seed*cross_offs; %number of seeds with genotype x that has been produced through outcrossing
self_seed_x = self_seed*self_offs; %number of seeds with genotype x that has been produced through selfing
out_seed_x(isnan(out_seed_x)) = 0; % convert NaNs to zero (if for any unforeseen reason the model gives NaN values)
self_seed_x(isnan(self_seed_x)) = 0; % convert NaNs to zero (if for any unforeseen reason the model gives NaN values)
N_x = self_seed_x+ out_seed_x; % total number of seeds with genotype x after on generation
%% End of the loop
SeedBank (j,:) = N_x; % we save all these data into the new matrix SeedBank
% This matrix SeedBank has "n" columns (equal to the number of genotypes) and "time" rows (equal to the
% number of generations). The first column represents the "colonizer" while the last column is the "resident" species.
% If "alpha = 0" it means there is no hybridisation, so all the first and last columns of matrix "SeedBank" will have
"non-zero" data.
end
end

```

```
function [cross_offs, self_offs] = crossing(n)
% This function gives the identity of genotypes resulted from all pairwise crossings between
% genotypes in the matrix "cross_offs" with n^2 by n dimension while the identity offspring from
% selfing is given in "self_offs" which has a dimension n*n.
% This function is based on E.12 and Eq. 13 in Supplementary Information.
```

```
% Mesgaran B Mesgaran (mohsenm@unimelb.edu.au): May 2016
```

```
%%
```

```
n = (n-1)/2;
```

```
L = 1:n;
```

```
N=cell(n,1);
```

```
for j=1:length(L)
```

```
for i=1:L(j)+1
```

```
    K = 0:L(j);
```

```
    if K(i)==L(j)-K(i)
```

```
        N{j,1}(i) = 2*(nchoosek(n,K(i))*nchoosek(n,L(j)-K(i)));
```

```
    else
```

```
        N{j,1}(i) = nchoosek(n,K(i))*nchoosek(n,L(j)-K(i));
```

```
    end
```

```
end
```

```
end
```

```
P = cellfun(@(x) x./sum(x),N,'uni',0);
```

```
for i=1:n
```

```
    l = length(P{i});
```

```
    P{i}(1,l+1:n+1)=0;
```

```
end
```

```
gmt=[[1 zeros(1,n)]; cell2mat(P)];
```

```
gmt2 = fliplr(flip(gmt));
```

```
mios = [gmt; gmt2(2:end,:)];
```

```
%% For outcrossed flowers
```

```
cross_offs = cell(2*n+1,1);
```

```
for i=1:2*n+1
```

```
    for j=1:2*n+1
```

```
        cross_offs{i}(j,:) = conv(mios(i,:),mios(j,:));
```

```
    end
```

```
end
```

```
cross_offs = cell2mat(cross_offs);
```

```
%% For selfed flowers
```

```
self_offs = zeros(2*n+1);
```

```
for i = 1:2*n+1
```

```
    self_offs (i,:) = conv(mios(i,:),mios(i,:));
```

```
end
```

```
end
```

```

function F = geno2pheno(u,T,z,n)
% This function maps genotype to phenotype using Eq.16 in Supplementary Information. The output
% vector F is scaled to vary from T to 1 (if T<1) or 1 to T (if T > 1). The function uses the "solve"
% to estimate the value of sigma in Eq. 16.
% u = the location of the genotype with largest or smallest phenotypic value
% T = is the phenotypic value of a trait
% if T > 1 u is the location for the genotype with largest phenotypic value
% if T < 1 u is the location for the genotype with smallest phenotypic value
% if T = 1 all genotypes are equally fit
% z = can take 3 values: 1, 2 and 3
% if z = 1, the output F is of length of n (= number of genotypes);
% if z = 2, the output F is of length of n^2 and includes the phenotypic values of genotypes acting%
as female in all pairwise combinations of genotypes
% if z = 3, the output F is of length of n^2 and includes the phenotypic values of genotypes acting%
as male in all pairwise combinations of genotypes
%
% Mesgaran B Mesgaran (mohsenm@unimelb.edu.au): May 2016
%% Total number of genotypes
G = linspace(0,1,n); % genotypes vector
E = repmat(G,1,n);
Y=[sort(E(:,1)) E(:,2)];
Ym = Y(:,1);
Yf = Y(:,2);

if z ==1;
    G = G;
elseif z == 2;
    G = Yf;
else z == 3;
    G = Ym;
end
if T < 1;
    if u <= .5;
        X = 1;
    else X = 0;
    end
    syms v
    v=solve(T*exp(((X-u)^2)/(2*v^2)))==1,v);
    v = double(v);
    v=v(v>0);
    F = T*exp(((G-u).^2)/(2*v^2));
elseif T > 1;
    if u <= .5;
        X = 1;
    else X = 0;
    end
    syms v;
    v=solve(T*exp(-((X-u)^2)/(2*v^2)))==1,v);
    v = double(v);
    v=v(v>0);
    F = T*exp(-((G-u).^2)/(2*v^2));
else T == 1;
    F = ones(1,length(G));

```

```
end
  F = reshape(F,length(G),1);
end
```

```

function [Sx, Sx_x] = selfingrate(s_0,s_1,n)
% this function gives the selfing rate of all "n" genotypes (based on Eq.16 in Supplementary Information)
%in vector "Sx" and n replicates of them in "Sx_x"
% n = number of genotypes
% s_0 = selfing rate in colonizer [0,...,1]
% s_1 = selfing rate in resident [0,...,1]

% Mesgaran B Mesgaran (mohsenm@unimelb.edu.au): May 2016
[m , e] = meshgrid(s_0,s_1);
Sx_x = [m(:) e(:)];

Lm = length(s_0);
Le = length(s_1);
Sx = zeros(n,Lm*Le);

for i=1:Lm*Le
    if Sx_x(i,1)>=Sx_x(i,2)
        Sx(:,i) = tdg(0,Sx_x(i,1)/Sx_x(i,2),1,n)*Sx_x(i,2);
    else
        Sx(:,i) = tdg(1,Sx_x(i,2)/Sx_x(i,1),1,n)*Sx_x(i,1);
    end
end

Sx_x= repmat(Sx,n,1);
end

```