Evolution of flowering time in a selfing annual plant: Roles of adaptation and genetic drift

Laurène Gay, Julien Dhinaut, Margaux Jullien, Renaud Vitalis, Miguel Navascués, Vincent Ranwez, and Joëlle Ronfort

10.1101/2020.08.21.261230
Submitted by Laurene Gay 2020-08-21 17:26

Abstract

Resurrection studies are a useful tool to measure how phenotypic traits have changed in populations and they allow testing whether these traits modifications are a response to selection caused by an environmental change. Selfing, through its reduction of effective size, could challenge the ability of a population to adapt to environmental changes. Here, we used a resurrection study to test for adaptation in a selfing population of Medicago truncatula, by comparing the genetic composition and flowering across 22 generations. We found evidence for evolution towards earlier flowering times by about two days and a peculiar genetic structure, typical for highly selfing population, where some multilocus genotypes (MLGs) are persistent through time. We used the change in frequency of the MLGs through time as a multilocus fitness measure and built a selection gradient that suggests evolution towards earlier flowering times. Yet, a simulation model revealed that the observed change in flowering time could be explained by drift alone, provided the effective size of the population is small enough (<150). These analyses suffer from the difficulty to estimate the effective size in a highly selfing population, where effective recombination is severely reduced.

Keywords: Adaptation | Selfing | Climate change | Selection gradient | Flowering time

Round #1

Author's Reply:
Dear Dr Haag,

Please find enclosed a revised version of our manuscript. We are very grateful to you and the reviewers for the comments and suggestions that have improved the manuscript substantially. We tried to answer to all of them (see the point-by-point reply below). We provide a track-changes version where the changes in the main text and supplementary files are highlighted in bold. We hope that you will find this updated version of our manuscript suitable for recommendation by PCIEvolBiol and would be happy to take any further comments if you judge it would improve the manuscript.

Laurène Gay, on behalf of all the coauthors

---

by Christoph Haag, 2020-10-26 22:11
Manuscript: 10.1101/2020.08.21.261230

Revise

Dear Dr Gay,

Thank you for submitting your preprint "Evolution of flowering time in a selfing annual plant: Roles of adaptation and genetic drift" to PCI Evol Biol. Your work has now been considered by three reviewers, whose comments are enclosed. As you will see, the reviews are largely positive, and, based on these reviews as well as my own reading, I am happy to further consider your preprint for recommendation. However, before reaching a final decision, I would like you to revise your manuscript according to the recommendations by the reviewers.

Besides the more minor points (which also should be considered carefully), I think there are two main issues that need particular attention:

- First, the introduction (and perhaps also some other sections) would profit from some streamlining. In my opinion, this does not mean that you should entirely drop the discussion of the effects of selfing on the efficacy of selection. But this section should be reduced in length and care should be taken to clearly state the objective of the study early on without raising issues (e.g., comparison between selfers and outcrossers) that are not subsequently addressed.

We agree that the Introduction was too long. As detailed below in response to Reviewer 1, we tried to refocus the Introduction on addressing why adaptation could be challenging for selfing population, but dropped some of the examples that made the focus drift towards a comparison between selfers and outcrossers. This resulted in a reduction by ¼ of this section (page 2 of the manuscript, lines 99-170).

> Incidentally, from my own reading, I also think that the last part of page 1 (where you give some more detail on the different possible approaches to investigate the influence of selection on phenotypic change) would profit from some reformulation: I found this part difficult to follow and its purpose is not entirely clear to me: Do you want to provide details on some of the approaches or do you want to explain why you used only some bot not others in your study?

The aim of this paragraph was to stress out that it is essential to test for selection in order to conclude that a phenotypic change is adaptive. We detail the methods to test for selection both to establish the state of the art and to discuss the experimental requirements. In particular, this section is important to define the effective size, a parameter that is key for our study in selfing populations. We tried to add a few details to shed more light on the purpose of this section (starting line 69, end of the 2nd paragraph, page 1).

> Moreover, the statement that natural populations cannot be replicated may also need to be nuanced (replication might in principle be possible across different populations or using independent samples from the same population).

Samples from the same population would not be independent in terms of evolutionary trajectory. Yet, assuming that the parameters are similar across populations (same effective size, same selective pressures…), looking at the evolutionary trajectories across different populations could provide some kind of replication. This is actually what we did when we considered the regional scale. We clarified this in the last paragraph of page 1 (line 74: “Pattern tests of neutrality, for example, are based on comparing phenotypic or allele frequency changes across replicates in experimental populations, or across populations,

assuming that they are independent replicates of the evolutionary process (same effective size and selective pressure, no migration)"). We also refer to 'pattern tests' in the Methods, page 6 section "Regional analysis" lines 532-533.

- Second, the analysis of the frequency changes of the multilocus genotypes needs some clarification, both in terms of potential effects of excluding rare genotypes and in terms of confidence intervals given (likely) non-normal distribution of residuals.

We answer this point below in reply to Stefan Laurent's comment.

If you submit a revised version, please include a letter in which you describe how you have responded to each of the referees' comments.

Best withes, and apologies again for the delayed decision, Christoph Haag

Additional requirements of the managing board:
As indicated in the 'How does it work?' section and in the code of conduct, please make sure that:
-Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.
-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.
-Details on experimental procedures are available to readers in the text or as appendices.
-Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders."

Reviews

Reviewed by Pierre Olivier Cheptou, 2020-10-20 11:18

The study by Gay et al. reports empirical data on the evolution of flowering time in a highly selfing species: Medicago truncatula. The authors used several approach to investigate the question. In particular, they used a resurrection approach with seeds from 1987 and 2009. The aim of the study is to disentangle the role of drift and selection in the shift observed as well as estimating selection gradient of flowering time. The study is interesting and the different experiments (pop centered, regional) is consistent with a shift in flowering time. Below, my comments:

1-The introduction discuss the question of adaptation face to environmental change. While the text is rich and well referenced, I found that the introduction is a bit long. There is a long discussion on whether outcrossing/selfing traits influences adaptation. The logical consequence would be to compare outcrossing/selfing populations. Since the study does not compare outcrossing and selfing populations, I think this part should be greatly reduced. Also, the statement that bottlenecks are more frequent in selfers (if true !!) would be more striking if the references were reporting empirical data. To my knowledge, Schoen and Brown (1991) and Ingvargsson 2002 hypothesize that it is the case but did not demonstrated that selfers suffer from higher bottlenecks. In the following paragraph, I found confusing to assert that "self-fertilization mays have facilitated adaptations to agricultural practices" when discussing the role of mating system on adaptation. Is it because the traits were preadapted or because the genetic architecture of selfers facilitates adaptation? In short, the introduction should be more focused to introduce the question short term adaptation of flowering time in the face of warming.

The discussion of the influence of selfing on adaptation was not meant as a comparison between selfers and outcrossers but rather aimed at stressing out why it is uncertain that selfing populations will rise to the challenge of adaptation, in particular when facing climate change. We agree that some of the references cited (Morran 2019, Hereford 2010, references about domestication) may move the focus away from the question of short term adaptation of selfing populations to climate change and emphasise too much the comparison between selfers and outcrossers. We streamlined the text and removed some of the examples in an attempt to keep the Introduction concise and more focused on the aim of this study: short term adaptation of selfing populations to climate change (page 2, lines 99-170).

2-Sum of temperature. The individual flowering time is converted in sum of temperature. The basal temperature is assumed to be 5°C, based on Moreau et al (2007). Would it be possible that Tb has evolved during the two decades? Would the conclusions be different if flowering time were measured as the number of days? At least, the possibility of a shift in Tb should be discussed as I found contradictory to evaluate adaptation to warming but keeping Tb constant.

It is true that we had not considered a potential shift in Tb. We first reiterated the analysis using numbers of days rather than degree.days. As detailed in the following table, the shift in flowering time remained significant. We added this conclusion in the Results section (page 6, 1st paragraph of Results, lines 564-566).

In addition, we discuss the possibility of a shift in Tb on page 10 at the end of the 1st and 2nd paragraphs (lines 898-901 and 926-928).

Table 1BIS Results of the linear mixed model used to analyse the effect of sampling year and treatment on flowering time (**measured in number of days**) in the cap Corsica population, taking into account the family effect (genetic effect). Effect values on mean flowering time are given for fixed effects (in italics) and variance components are given for random effects (with standard errors in brackets). The family effect was nested into year (1987 or 2009) and treatment (T1: short vernalization treatment; T2: long vernalization treatment), leading to four variance components. For each component, the degrees of freedom, likelihood ratio (χ^2) and *p*-values are reported. None of the interactions considered in the complete model [1] were significant: between year and treatment (LRT $\chi^2 = 1.8$; df = 1; $p = 0.178$); between block and year ($\chi^2 = 0.0006$; df = 1; $p = 0.981$).

Tested effect on flowering time	Mean effect or variance component (SE)	df	χ^2	*p*
year	*-1.75*	*1*	*5.5*	*0.019*
treatment	*-8.35*	*1*	*34.2*	5.10^{-9}
block	0.503	1	55.2	1.10^{-13}
family | year × treatment	1987-T1: 10.396 1987-T2: 9.053 2009-T1: 20.786 2009-T2: 18.376	10	43.6	2.10^{-7}
error	6.433	1081		

*in days. Note that this effect is close to the result found using flowering time measured in degree.days (28.76 degree.days correspond to two days).

3-Maternal effects. If I understood well, the results on the studied populations are corrected for maternal effects (one generation to refresh seeds stock) but the results of regional analysis are based on the F1 generation (without correcting for maternal effects). I was interested by the amplitude of the shift: two days in the cape Corsica populations but five days in the regional analysis. This may be a "true result" or an effect of correcting for maternal effects. Did the authors measure the flowering date in the F1 of the cape Corsica populations. I would suggest to mention this result in the discussion. Is it possible that the difference in flowering date reported have changed in Cape Corsica population because of the F1 generation in greenhouse? My feeling is that these results are, as such, interesting. We often see this pattern of a lower amplitude after one generation. If it was only noise, the first generation should exhibit either lower or higher difference than the F2. Epigenetic components of flowering could have played a role in adaptation to warming and these effect cannot be distinguished from true quantitative genetic effects if parts of these effects last more than one generation. Do the same MLG (from 1987 and 2009) have the same fitness? Because the authors have the chance to have the same MLG, it would be interesting to look at this relationship to investigate maternal effects.

The plants used in the regional analysis were inbred lines just like in the Cape Corsica population experiment, multiplied during a generation in standardized conditions in the greenhouse. This is explained in the Methods (page 5 paragraph 'regional analysis' lines 536-544) and we added a sentence to underline that the protocol is the same for both experiments (page 6 lines 544-545). We therefore expect similar maternal effects in both experiments (i.e. none, assuming that one generation in the greenhouse is sufficient to control for this effect).

4-Genetic analysis. If I understood well, the test for selection versus drift is based only on conserved multilocus genotypes, i.e. a fraction of the population. Why doing this choice? Why not using a Qst/Fst approach that would take into account all the individuals? (the design allows to estimate Qst, doesn't it?). In addition, I see a potential bias because it assumes that the population behaves as a fully selfing populations, which is not the case. While the authors point the potential differential selective response of outcrossers versus selfers, the results

reported are based only on the full selfing fraction of the population, which I found contradictory.

Our test for selection versus drift is based on the changes in frequency of the MLGs through time. That includes all the homozygous MLGs: those that were conserved through time but also those that disappeared or appeared between 1987 and 2009. For the later, there is an uncertainty because we ignore how many generations a given MLG was present in the population. To circumvent this problem, we reiterated the analysis using only the MLGs that were present in 1987, but it reduces the sample size considerably. We discuss this issue in the manuscript page 10 lines 831 to 839.

Our rationale behind considering the changes in frequency of the repeated genotypes was to build a more integrative measure of fitness by taking profit from the peculiar genetic structure of predominantly selfing species that produce multilocus genotypes. It allows comparing two estimates of fitness (total number of seeds produced by a plant in the greenhouse + changes in frequency of multilocus genotypes *in-situ*) which we thought was interesting. Yet, it is true that any outcrossing event will lead to an underestimation of the fitness of a given MLG, and this fitness estimate can only be measured for the fully inbred genotypes (i.e. without any residual heterozygosity). We added a figure (Fig. S2 page 16) showing the distribution of the MLGs through time, and completed the supplementary table S4 to indicate the number of MLGs each year, and how many were removed when we excluded the MLGs with residual heterozygosity (6 in 1987 and 6 in 2009), Table S4 page 16. It highlights that the selfing rate is so high in this population that excluding the MLGs with residual heterozygosity only moderately reduced the sample size, from 145 to 133 individuals (58 from 1987 and 75 from 2009).

We considered adding a Qst-Fst analysis to this study before submitting this first version, seen as our experimental design indeed allows to estimate Qst. Yet, with temporal samples, one sample is derived from the other and our "populations" are not independent. If the expectation Fst = Qst holds under a drift alone scenario for any meta-population structure, including temporal data (Whitlock 1999 Genet Res), the expectations are more hazardous for non-neutral scenarios (G. Martin, pers. Comm.). For spatial data, it is generally assumed that directional selection should lead to Qst > Fst. Qst < Fst on the other hand, is interpreted as an indication of stabilizing selection. Yet, in this case, several issues impede a neat interpretation, among which non-additive genetic architecture, limited migration and drift (Lamy, Plomion, Kremer & Delzon, Molecular Ecology 2012; Le Corre & Kremer, Molecular Ecology 2012).

In our data, there is a large amount of phenotypic variance within population, and some is shared between years. In comparison, the variance between years is relatively small, and it results in a small Qst estimate (around 2%). The Fst is 0.226, so we have Qst < Fst. This could be an indication for stabilizing selection (or uniform selection) acting to maintain flowering time at an optimum, and not changing through time. But given the caveats listed above, and the fact that the non-independence between our populations (one being derived from the other) could affect the predictions for the non-neutral models, we found it difficult to conclude on this result. We would prefer not to include this analysis in the manuscript in order to avoid overloading it and to keep the focus on the multilocus genetic structure, which for us is more appropriate for such a highly selfing population. Should you think it is worth adding the Qst estimate in the main text, we could change the manuscript accordingly.

Overall, I found the ms interesting and such long term dataset is rare. However, the ms would benefit from being more focused (particularly the introduction) in order to highlight the results and their biological interpretation.

Reviewed by Jon Agren, 2020-10-19 15:12

This study uses a resurrection experiment and simulations to explore the possible causes of changes in flowering time and genetic composition of a Medicago truncatula population across 22 generations. In the resurrection experiment, plants grown from seeds collected 22 years apart were raised in the greenhouse to produce selfed lines. These lines were then used to document possible changes in flowering time and to quantify selection on flowering time in the greenhouse. Changes in genetic composition were characterized by scoring 20 microsatellite loci (16 kept after filtering) and documenting changes in the frequencies of multilocus genotypes. The paper is well written and addresses interesting problems of wide general interest. However, I think the authors need to (a) motivate their approach to use estimates of selection obtained in the greenhouse to infer selection in the field, (b) provide more detail on the distribution of multi-locus genotypes and the power of their analysis of change in genetic composition, and (c) clarify a few details when it comes to sampling procedure (see below).

Main comments:

1. The authors appear to assume that selection quantified in the greenhouse is likely to mirror selection in the field at present and 22 years ago. This needs to be motivated.

It is entirely possible that it does not. This is why we compared our greenhouse fitness proxy (number of seeds) with the change in MLG frequencies in the field (as stated now on page 5, lines 468-471). If the selection quantified in the greenhouse mirrors selection in the field that occurred over the last 22 years, we expect that the MLGs that increased in frequency should be those with the highest fitness proxy in the greenhouse (Fig. 3A). As reported in the text (Results page 8 lines 665-671; discussion page 9 lines 828-839), the relationship, though positive, is only significant for the sample restricted to the MLGs that were already present in 1987. This could be so, because the changes in frequencies measured for the MLGs that appeared within this time period are biased. Only a much larger sample for each year could solve this problem by reducing the uncertainty in MLG frequency estimates each year. Unfortunately, it is the “historic” sample (1987) that is most limiting.

2. I suggest the authors provide more detail on the distribution multilocus genotype (MLG) frequencies, and that this information is given already at the start of the third paragraph on p. 7. They report that 60 different MLGs were detected in their sample of 145 individuals. Two MLGs were common, and 12 MLGs were shared between the two sampling years. This suggests that most MLGs were rare and perhaps only represented by a single plant? The authors may want to discuss whether their sample sizes are sufficient to characterize changes in genetic composition of a population with such skewed distributions of MLGs.

Only 5 MLGs were shared between years, as detailed page 8 lines 661-665 (it was written 4 in the previous version but we corrected this mistake). 12 is the number of MLGs that were present in 1987 and fully homozygous. We added a supplementary figure to show the distribution of MLGs through time as well as the MLGs that are fully homozygous or with

some level of residual heterozygosity (Fig. S2, page 16). In addition, we completed the supplementary table S4 to include the number of MLGs and fully homozygous MLGs per year (page 16). Should you feel that it is required, we could include these figure and table in the main text.

We agree that the distribution of MLGs is highly skewed, but this is a characteristic of predominantly selfing populations (as detailed in Jullien et al. Heredity 2019). In selfing species, it is common to sample a single individual (sampled from the most frequent MLG) to avoid genotyping large numbers of fully identical individuals. Diversity is thus generally studied at the species scale, while occulting within-population diversity. Population studies are relatively rare, even more so with samples close to 100 individuals per year. We believe that this sample size is large enough to be representative of the genetic diversity, in the sense that it can capture the rare variants.

Nevertheless, we agree that the sample size is limiting for the year 1987, seen as several MLGs were absent and we ignore when they appeared. This disturbs the relationship between the change in frequency of a MLG and its fitness, as explained page 10 line 831-839.

3. I suggest the authors clarify a few details regarding sampling:

(a) For the resurrection experiment, “100 seeds per sampling were replicated” (p. 3, second paragraph). Were these seeds from 100 different pods and thus sampled from 100 plants, or were they a random sample of 100 seeds from a pooled seed sample from each year?

The seeds used for the experiment were a random sample from a bulk of seeds preserved since 1987 or collected in 2009. Going back to the raw data made us realize that there was a mistake in the numbers of seeds. We apologize for this approximation. The actual number of seeds that were multiplied was 64 for the year 1987 and 96 for 2009. We updated the Material section with the correct numbers of seeds (page 3, lines 248-256)

(b) For the genetic analysis, leaves were sampled from “the multiplication generation in the greenhouse” (p. 4, fifth paragraph), and after filtering 145 individuals remained in the data set to be analysed. Please, state explicitly that the “multiplication generation” refers to the plants derived from the 200 field-collected seeds (presumably representing seeds from 200 plants(?); see previous comment). Were seeds from the two sampling occasions equally represented among the 145 individuals included in the analysis?

After genotyping, we had to remove 15 individuals from the year 2009 because they had >10% missing data. As a result, the total number of plants that was multiplied and genotyped for this population is 64 plants collected in 1987 and 81 plants collected in 2009. This adds up to 145, as detailed in the results section. This is now detailed in the Methods section (page 4 lines 398-399).

Minor corrections:

Abstract, line 11 from bottom. Change “population” to “populations”

corrected

p. 7, first paragraph, second line from bottom, “in both years”. From this wording, you easily get the impression that selection was quantified in two years. I suggest you add a few words to indicate that this rather refers to a similar negative relationship being observed among lines derived from each of the two years.

Changed accordingly, page 7 lines 614-621

To make text in graphs readable, font size should be increased in Figures (in particular in Figs. 3-5).

Changed accordingly

Reviewed by Stefan Laurent, 2020-10-16 11:05

In this study, the authors test whether flowering time evolved in an experimental population of Medicago truncatula and whether this change could represent an adaptation to varying environmental conditions. For this, they measure changes in flowering time in a natural population over 22 generations (2 timepoints), they quantify the association between flowering time and fitness (as approximated by the number of seeds produced), they track changes in haplotype frequencies characterized by different approximated fitness values, and finally they also measure changes in flowering time in 17 populations from the same geographical region that have been sampled twice over a comparable time range.

The authors report a significant reduction in flowering in the main population and in the regional analysis that appears to be consistent with the specific effects of climate change in the Mediterranean region (i.e. limiting summer drought occurs earlier in the year). They also report a significant association between flowering time and seed production. However, the evidence for the effect of positive selection obtained by analyzing the changes in haplotypes is at best marginal; even if the authors do a good job in describing some of the uncertainty associated with this analysis, I think that one more aspect should be exposed.

Besides my major comment, I find the manuscript clearly written, the analyses carefully conducted and presented, and the intro and discussion very well written and informative, at least for the non-expert.

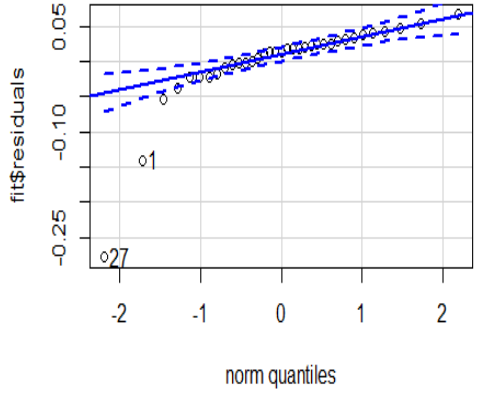
Major comment

My only major criticism refers to the results presented in Figure 3. The selection gradients measured here seem to be heavily influenced by two outlier points with low seed production and early flowering. As a result, the linear models (especially the one for MLG found in 1987) appear to be a poor fit to the data, as can probably be seen by inspecting the residuals, which are unlikely to be normally distributed. I think that the authors should report the uncertainty around the slopes and that this uncertainty should be further considered in the analyses presented in figure 4, which will likely cause the observed selection gradient to be non-significant under a larger range of Ne values. I am not sure about the best way to obtain confidence intervals for the selection gradients but I imagine that a bootstrap approach should be applicable.

We agree that these two points (corresponding to the two most frequent MLGs) may have a heavy influence on the selection gradient. It is worth pointing out that these 2 points

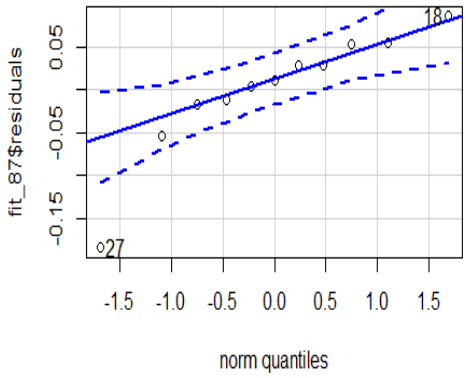
correspond to the two most frequent MLGs, and together account for 50% of the population in 1987 and 20% in 2009. It is therefore unlikely that there is large experimental error in the phenotypic estimates for these MLGs and the association between their late flowering time and their decrease in frequency is unlikely to be an artefact. The large number of families that carry these two MLGs also explains why the selection gradient at the family scale is so highly significant, when the selection gradient based on MLGs is more equivocal. We agree that the selection gradients based on MLGs do not seem to fit the data very well. The residuals of the linear models indeed deviate slightly from normality, in particular for the model including all the homozygous MLGs. Below are the qqplots (that we could include in the manuscript as supplementary figures if you think it is necessary).

Analysis using all points
n = 48; slope = -0.0002 p = 0.179




```
Shapiro test of normality
W = 0.65467, p-value = 7.9e-08
```

Analysis restricted to the MLGs present in 1987
n = 12; slope = -0.0009 p = 0.038




```
Shapiro test of normality
W = 0.84759, p-value = 0.0397
```

To account for this limit of our models, we now report in the manuscript the uncertainty on the slopes, measured as the profile likelihood confidence bounds (Fig. 3, page 8).

In our test for selection, we compared the slopes estimated from simulations of drift alone with an effective population size of 19, or with Ne ranging between 10 and 500 (Fig. 4) with the slope estimated in our data. A conservative way to take the estimation error into account could be to compare the slopes on simulated data to the upper bound for the observed slope (negative). For the data restricted to the MLGs present in 1987, the upper bound of the slope is very close to the slope estimate for the complete dataset (-0.0002), which means that the minimum effective size to explain the observed pattern by drift alone is 150 (Fig. 4A). In order to keep the manuscript relatively concise, we did not include this but rather stressed out the fact that this analysis with realized fitness (change in MLG frequency through time) severely suffers from lack of data. More temporal samples or a much larger sample size for 1987 would be required to improve the interpretation of the results. We believe that this does not change much the conclusions of the manuscript because we were aware of these shortcomings and already rather cautious in our interpretation. Our conclusions are based on the body of evidence (among which Robertson-Price selection gradient, regional analysis) rather than on this analysis alone. Should you think this is not clear enough in the manuscript, we could change it.

Minor

I agree with the authors that the N_e value estimated from the temporal Fst is very likely underestimated. Comparing the expect heterozygosity under Ne=19 with the observed He would further support the idea that larger Ne values are indeed realistic. How does the observed heterozygosity in the population compares to the theoretical expectations given by Nordborg and Donnelly (1997)? Rescaling the census number (>2000) by 1/(1+F) would lead to a less conservative Ne value for the test for selection and may allow a putative selection signal to be detected even after considering the uncertainty around the observed selection gradient.

According to Equation 16 in Nordborg and Donnelly (1997), we predict that $H_E = 1 - \frac{1}{1+4N_e\mu}$, where N_e is the effective size as estimated using the Waples method. In *Arabidopsis thaliana*, mutation rates for dinucleotide microsatellite loci have been estimated ranging between 4.96 x 10^{-5} and 2.03 x 10^{-3} (Marriage et al. 2009 Heredity). Based on these estimates, we expect that H_E should lie between 0.004 and 0.134, which is nearly three times lower than the H_E estimated using the microsatellite markers (0.351 (0.252-0.424) in 1987 and 0.623 (0.599-0.627) in 2009, Suppl. Table S4). Equation 16 is based on the Infinite Alleles Model (IAM) mutation model, which may not be the best model for SSR markers. Under the SMM (Stepwise Mutation Model), Kimmel et al. showed that $H_E = 1 - \sqrt{\frac{1}{1+2*4N_e\mu}}$ (Equation 13 in Kimmel et al. 1998 Genetics). With this equation, we expect that H_E should lie between 0.004 and 0.126. We added this comparison in the section "Changes in the genetic composition of the population" of the Results (page 7, lines 634-643) and in the Discussion (page 9 lines 770-773).