Dear Dr. Ines Alvarez,
We appreciate your work, and the comments you, Dr. Rubén Torices and the two anonymous reviewers did on the previous version of our manuscript.
We considered all the issues raised by reviewers and addressed all the comments. We reduced the abstract, introduction and discussion. We now clarify the biological system by adding new results and the definition of our SNP set by adding a new analysis. We also clarified our questionings, hypotheses, main results and messages in our writings in all the different sections.
We responded to the comments below in blue, either by the "Action" or by "Answer" when appropriated.
We believe our manuscript brings novelties and advances for *Lgh* comprehension, on how to develop population genetic approach for autopolyploid plants and for understanding the population genetic consequences of a late-acting self-incompatible system as compared to self-compatible populations of the same species in similar ecological conditions.
Our manuscript must be now suitable for recommendation.

All the data supporting this new version are again openly available on Zenodo:
Stoeckel, S., Barloy, D., Portillo Lemus, L., & Becheler, R. (2023). Genotyping measures and population genetic indices for assesing reproductive modes of polyploid Ludwigia grandiflora subsp. hexapetala in western Europe [Data set]. Zenodo. https://doi.org/10.5281/zenodo.12760022

Sincerely,

Dominique Barloy and Solenn Stoeckel, on behalf of the co-authors

Sujet : PCIEvolBiol #788: Decision concerning your submission
Date : Tue, 28 May 2024 13:14:11 -0000
De : PCI EvolBiol Contact <contact@evolbiol.peercommunityin.org>
Pour : solenn.stoeckel@inrae.fr
Copie à : ines@rjb.csic.es

PCIEvolBiol
Decision concerning your submission

Dear Solenn Stoeckel,
Your article, entitled **Reproductive modes of polyploid Ludwigia grandiflora subsp. hexapetala in western Europe: the effects of a late-acting self-incompatibility system and its absence on genetic diversity within populations**, has now been reviewed.
The referees' comments and the recommender's decision are shown below. As you can see, the recommender found your article very interesting but suggests certain revisions.

We shall, in principle, be happy to recommend your article as soon as it has been revised in response to the points raised by the referees.
When revising your article, we remind you that your article must contain the following sections (see our Guide for Authors in the Help section of the PCIEvolBiol website):
1) **Data, script and code availability (if applicable)**

- **Data, statistical scripts, command lines and simulation code must be made available to readers**. They should either be included in the article or deposited **in** an open repository such as Zenodo **with a DOI**. A perennial URL can be provided if no DOI is available; please note that GitHub URL are not perennial.
- **If deposited in** an open repository, a reference to **Data, statistical scripts, command lines and simulation code**, with a DOI or a perennial URL, must be provided in the reference list and in the "Data, script and code availability" section
- The "Data, script and code availability" section must clearly indicate **where and how** data can be accessed.
- Wherever possible, data, scripts and code should be provided in machine-readable formats. Avoid PDFs other than for textual supplementary information.
- Metadata should accompany the data, to make the data understandable and reusable by the reader.

2) **Supplementary information (if applicable)**

- Supplementary information (text, tables, figures, videos, etc.) can be referred to in the article. It must be available in an open repository (such as Zenodo, Dryad, OSF, Figshare, Morphobank, Morphosource, Github, MorphoMuseuM, Phenome10k, etc. or any institutional repository, etc...) with a DOI. A perennial URL can be provided if no DOI is available.
- A reference to the supplementary information, with a DOI or a perennial URL, must be provided in the reference list and in the "Supplementary information" section.
- List all documents attached to the manuscript as Supplementary Information in the "Supplementary Information" section.

3) **Funding (mandatory)**

- All sources of funding must be listed in a separate “Funding section”. The absence of funding must be clearly indicated in this section.

4) **Conflict of interest disclosure (mandatory)**

- Authors should declare any potential non-financial conflict of interest (financial conflicts of interest are forbidden, see the PCI code of conduct).
- In the absence of competing interests, the authors should add the following sentence to the “Conflict of interest disclosure” section: “The authors declare they have no conflict of interest relating to 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 a recommender for PCI XX.”

5) **Materials and methods (mandatory)**

- Details of experimental procedures and quantitative analyses must be made **fully available** to readers, in the text, as appendices, or as Supplementary Information deposited in an open repository, such as Zenodo, Dryad or institutional repositories with a DOI.
- For specimen-based studies, **complete repository information** should be provided and institutional abbreviations should be listed in a dedicated subsection (if applicable). Specimens on which conclusions are based **must be deposited in an accessible and permanent repository**.

When your revised article is ready, please:
1) Upload the new version of your manuscript onto your favorite open archive and **wait until it appears online**;

2) Follow this link https://evolbiol.peercommunityin.org/user/my_articles or logging onto the PCIEvolBiol website and go to 'For Contributors -> Your submitted preprints' in the top menu and **click on the blue 'VIEW/EDIT' button at the right end of the line** referring to the preprint in question.
3) Click on the black 'EDIT YOUR ARTICLE DATA' button (mandatory step). You can then edit the title, authors, DOI, abstract, keywords, disciplines, and DOI/URL of data, scripts and code. Do not forget to save your modifications by clicking on the green button.
4) Click on the blue 'EDIT YOUR REPLY TO THE RECOMMENDER' button (mandatory step). You could then write or paste your text, upload your reply as a PDF file, and upload a document with the modifications marked in TrackChange mode. If you are submitting the final formatted version ready to be recommended, you should only add a sentence indicating that you posted the final version on the preprint server. Do not forget to **save your modifications by clicking on the green button**.
5) Click on the green 'SEND RESUBMISSION' button. This will result in your submission being sent to the recommender.
Once the recommender has read the revised version, they may decide to recommend it directly, in which case the editorial correspondence (reviews, recommender's decisions, authors' replies) and a recommendation text will be published by PCIEvolBiol under the license CC-BY.
Alternatively, other rounds of reviews may be needed before the recommender reaches a favorable conclusion. They may also reject your article, in which case the reviews and decision will be sent to you, but they will not be published or publicly released by PCIEvolBiol. They will be safely stored in our database, to which only the Managing Board has access. You will be notified by e-mail at each stage in the procedure.
We thank you in advance for submitting your revised version.
Yours sincerely,
The Managing Board of PCIEvolBiol

Revision round #1

Decision for round #1 : ***Revision needed***

Very interesting topic, but it needs a deep review and clarification of relevant issues

---

The manuscript deals with a very interesting topic in several areas. Knowledge about plant reproduction systems is crucial to understanding their evolution, population structure and demography. On the other hand, the invasion capacity of some species can be largely explained by their reproduction system. Therefore, this case of study is of great relevance and I believe it has great potential to be published.
However, the three reviewers who have worked on the manuscript agree on several aspects, to which I join as associate editor:
I think that the manuscript as a whole should be reduced to at least 1/3 of its original length, especially in the Abstract and the Introduction and Discussion sections. Additionally, the title should be changed to something more appropriate. In short, the manuscript requires in-depth reorganization and writing work to achieve a clearer and more synthetic text.
Regarding the content, the doubts raised by the anonymous reviewer are especially worrying and must be clarified in order to continue with the review process. The main one of all is about the identity of the system under study, since it could be two lineages instead of just one. This could be amended by checking the chromosome numbers of the two morphs in the case study.
Another important question that this reviewer raises is about the selection of the SNPs used.
In short, it is important to take into account all these issues and others that the three reviewers who have handled this work consider in order to continue with the process.

*by **Ines Alvarez**, 28 May 2024 15:07* Manuscript:**https://doi.org/10.1101/2024.03.21.586104** version: 2

We thank recommender for this synthesis and her work on our manuscript. We believe these have strengthened the communication of our results in this new version.

About the manuscript length:
We now reduced the length of the manuscript from 36 pages to 27 pages while adding new results as requested by reviewers.
- The abstract from 2 pages to 1 page
- The introduction from 7 to 4.5 pages
- The discussion from 12 pages to 8 pages

We benefited from this reduction to reorganize these three parts with the aim to better communicate the messages and arguments supported by our results and the literature. We also wrote clearer subtitles for the different subsections of the discussion.

About the title:
Action: We changed the title for "*Reproductive modes in populations of late-acting self-incompatible and self-compatible polyploid Ludwigia grandiflora subsp. hexapetala in western Europe*".

About the identity of the system:
Answer: We now included a clearer summary of the previous peer-reviewed and published papers that we already quoted and, as noted by Reviewer#3, and we added new results as requested by reviewers and recommender, that must extensively answer this worry.
We now report the number of chromosomes on karyotypes we did on plants sampled across 7 populations, two S-morph and five L-morph, along a transect of ~500 km East-West, each sampled population distant from the other of 50 to 150 km. These new results add to previous observations and evidences that L- and S-morph individuals belong in western Europe to the same species *Ludwigia grandiflora* subsp. *hexapetala*.

In details, Dandelot (2004) has already independently reported differences in flower sizes between fruitful and fruitless *Lgh* populations in France. Dandelot et al (2005) resolved taxonomic identification of *Ludwigia* species present in France and showed that '*A morphological and cytogenetic study allowed to confirm the presence of two different taxa in France. The diploids (2n = 16) correspond to L. peploides* subsp. *montevidensis (Spreng.) Raven, and the decaploids (2n = 80) to L. grandiflora* subsp. *hexapetala (Hook. & Arn.) Nesom & Kartesz.*' Cytogenetic observations (karyotypes) were carried out on fruitful (Atlantic coast) and fruitless (Provence) *L. grandiflora* subsp. *hexapetala* in different areas in France by Dandelot et al (2005), where we also sampled our populations, that were then respectively identified as S-morph and self-compatible (SC) and L-morph and self-incompatible (LSI) individuals in further studies (Portillo-Lemus et al. 2021, 2022). Barloy et al. (2024) counted the chromosome number of one S-morph plant sampled near Nantes (France) and Bou Manobens et al. (2019) counted the chromosome number of one L-morph plant sampled in Catalunya. Both karyotypes presented 80 chromosomes, as expected in *Lgh*.
In addition, Portillo-Lemus et al. (2021) and Portillo-Lemus (2021) performed cross-fertilization between L-morph (LSI) and S-morph (SC) individuals in a common garden using both open pollinations and hand-controlled cross-pollinations, using 105 individuals. We already quoted these peer-reviewed and published data in the introduction when describing the complex biology and system of *Lgh* in France and northern Spain. This work obtained full (100% seeds) fruit production regardless of type of cross (S-morph x S-morph and L-morph x

S-morph and S-morph x L-morph) using 105 individuals sampled within seven different populations of the two floral morphologies sampled from seven western European populations among the 53 we studied here (details in Portillo-Lemus et al. 2021). These populations were also sampled for the current study (see underlines populations on Figure SI 1) and one per these seven populations was karyotyped to count their number of chromosomes (2n=10X=80). All seeds, independently of their origins, germinated and even resulted in viable plants able to produce a next generation (F1 generation, Portillo-Lemus et al. 2021). Again, seeds from F1 plants germinated and gave fruitful F2 plants which produced the same number of seeds (100% of seeds and fruits; zero abortion) and viable plants (Portillo-Lemus, PhD thesis, 2021). Crosses between S-morph and L-morph resulted in the same fruit and seed set as found when crossing two S-morphs. Success of sexual reproduction to produce viable and fertile first- and second-generation descendants argue also that S-morph (self-compatible) and L-morph (self-incompatible) individuals from fruitful and fruitless populations with different floral morphologies in France belong to the same species *L. grandiflora* subsp. *hexapetala*.
Then, to again verify deeper this system, we carried out interspecific crosses between different *Ludwigia* species with different ploidy levels (diploid x tetraploid, tetraploid x diploid, diploid x decaploid, decaploid x diploid, tetraploid x decaploid, decaploid x tretaploid) in controlled conditions. These results were peer-reviewed and published in Barloy et al (2024), that we cited in the first version of our manuscript when presenting *Lgh* biological system. Crosses between *L. peploides* subsp. *montevidensis*, a progenitor of *L. grandiflora* subsp. *hexapetala*, and *Lgh* only produced some fruits. Within these few fruits, only few seeds germinated and the survivors resulted in chlorotic (probably due to maternal incompatibility) plantlets. The most vigorous plants ended up dying 56 days after their seedling, thus before reproducing. This result argues again that the successful rates of crosses (100%) observed when crossing S-morph (self-compatible) and L-morph (self-incompatible) individuals with different floral morphologies coming from fruitless and fruitful populations in France belong to the same species, *Lgh*, and doesn't match what we obtained when crossing two true different sister-species.

Action:
In addition to these published results, we now report in this manuscript the number of chromosomes counted on karyotypes of plants across seven populations of different floral morphologies (two S-morph and five L-morph), on individuals maintained on our common garden and used for controlled pollinations in Portillo-Lemus et al. 2021 and 2022. All plants independently of their floral morphs and reproductive modes had 80 chromosomes, confirming that L- and S-morph plants in western Europe correspond to *Ludwigia grandiflora* subsp. *hexapetala*. We therefore added the following part to our manuscript:
Introduction (lines 116-120):
"*Lgh presents heteromorphic flowers corresponding to two floral morphs: L-morph individuals develop long-styled flowers and S-morph individuals develop short-styled flowers, that cross and result into 100% viable and fertile F1 and F2 descendants (Portillo-Lemus 2021, Portillo-Lemus et al. 2021) while inter-species crosses only result in a low number of chlorotic and unfertile descendants (Barloy et al. 2024).*"
M&M (lines 176-184):
"*In addition to the crossbreeding results in which all L- and S-morph individuals succeeded to cross, giving full fruit set and 100% viable first- and second-generation descendants (Portillo-Lemus et al. 2021), we here counted the chromosome numbers on karyotypes of two S-morph individuals sampled in two fruitful populations and of five L-morph individuals sampled in five fruitless populations to validate that L- and S-morph individuals belong to Lgh. Between 50 to 150 kilometers separated two consecutive samples (populations underlined Figure SI1). To prepare the karyotypes, we used the method detailed in Barloy et al (2024) that already*

karyotyped a S-morph individual sampled near the French Atlantic coast. The same method was used in Bou Manobens et al. (2019) to karyotype a L-morph individual sampled in Catalunya."
Results (lines 346-350):
"*Karyotypes of L-morph plants from two populations and of S-morph plants from five populations all presented the same number of chromosomes (2n=80, Figure SI2) confirming that L and S-morph individuals in France and L-morph individuals in northern Spain (Bou Manobens et al. 2019) belong to L. grandiflora subsp. hexapetala (2n=10X=80, Barloy et al. 2024).*"
Discussion (lines 476-481):
"*The five L-morph and the two S-morph water primrose populations we karyotyped, sampled in seven populations separated by 50 to 150 km from one to the other in France, all had 80 chromosomes, corresponding to the species Ludwigia grandiflora* subsp. *hexapetala. We didn't find yet individual with 48 chromosomes corresponding to Ludwigia grandiflora* subsp. *grandiflora. These results agree with previous observations of Dandelot et al (2005) and Barloy et al. (2024) in France, Bou Manobens et al. (2019) in northern Spain and Armitage et al. (2013) in Great Britain.*"

About the SNP set:
Action: We now provide a statistical proof and validation that tetraploidy is the best model over all our data to explain the distribution of the number of sequences in our 795 genotyped individuals, as expected from Barloy et al. (2024) and the way we developed this SNP set so it matches *Lpm* and *Lgh* genomes. We computed the Akaike criterion information on the likelihood of the distribution of the number of sequences considering diploid, tetraploid, hexaploidy, octoploid and decaploid genotypes.
We added a description of this approach, similar to the one proposed by Burnham & Anderson (2002), in Material and Methods, lines 202-211. We added the detailed result in the open data file, tab "AIC of the best ploidy". We wrote a sentence reporting this result lines 350-352 in the Result section and included the corresponding Figure SI3.

Review by Rubén Torices, 08 May 2024 11:03
Understanding the role of clonal reproduction and the reproductive system in structuring genetic diversity within and between populations is crucial for comprehending the diversity of a significant portion of today's plant species. The analytical complexity increases substantially when dealing with a polyploid organism, as in this case. The study discussed in this article offers a valuable contribution, paving the way for future research on similar systems. It enhances our understanding of the synergistic effect on genetic diversity of two key components in many plants' reproduction: the capability for asexual reproduction and uniparental sexual reproduction. Given the ambitious scope of this work, it is understandably extensive. However, the introduction and discussion sections occasionally become very long, providing some details that are not directly related to the results of this study.
Answer: We warmly thank Rubén Torices for his constructive comments and concrete proposals. It helped us redirecting our argumentation and the communication of our results that required a huge amount of work and cogitation while dealing with such an unusual biological system, that, we believe, definitely contribute understanding the ecology and evolution of western European populations of this species.

Below I answer all the questions posed and after the answers I have added some general and specific comments I have about the MS.
Title and abstract

Does the title clearly reflect the content of the article? [ ] Yes, [X] No (please explain), [ ] I don't know
The title is a bit vague. The work is very broad and it is certainly difficult to include a title that briefly and very precisely describes the content of the work.
Action: Right. We change the title for "*Reproductive modes of late-acting self-incompatible and self-compatible polyploid Ludwigia grandiflora* subsp. *hexapetala in western Europe*"

Does the abstract present the main findings of the study? [X] Yes, [ ] No (please explain), [ ] I don't know
It does, but I believe that this abstract should be strongly reduced.
Action: We reduced by the abstract from 2 pages to 1 page.

Introduction
Are the research questions/hypotheses/predictions clearly presented? [X] Yes, [ ] No (please explain), [ ] I don't know
Does the introduction build on relevant research in the field? [X] Yes, [ ] No (please explain), [ ] I don't know
Materials and methods
Are the methods and analyses sufficiently detailed to allow replication by other researchers? [] Yes, [X] No (please explain), [ ] I don't know
I believe this MS could benefit from a dedicated section in MM on the model species. While the introduction does contain an extensive paragraph about the species, it should be equally important to provide detailed information on the species, including its taxonomic status, major life history traits, and other relevant aspects.
Action: we included all this information in the paragraph in introduction line 107 to 136. We need this detailed information on the species to introduce our question and hypotheses, so we left this information in the introduction.

Are the methods and statistical analyses appropriate and well described? [X] Yes, [ ] No (please explain), [ ] I don't know
Results
In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? [X] Yes, [ ] No (please explain), [ ] I don't know
Are the results described and interpreted correctly? [X] Yes, [ ] No (please explain), [ ] I don't know
Discussion
Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? [X] Yes, [ ] No (please explain), [ ] I don't know
Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [X] Yes, [ ] No (please explain), [ ] I don't know
General comments
1. Potential inconsistency between the number of genotypes and the rate of clonality
I found some of the results somehow contradictory. This study provided substantial evidence supporting a high rate of clonality. But at the same time, almost 60% of the individuals sampled were genetically different. This makes me wonder whether these two outcomes are reconcilable. Given the observed high rate of clonality, would it not be reasonable to anticipate a higher frequency of clones?
Answer: Yes, they are reconcilable. There is an extensive literature on this point, please see some of the available syntheses that we cited in our introduction lines 72-78: Halkett et al. 2005, Arnaud-Haond et al. 2007, Stoeckel et al. 2021. In short, let's imagine a population of parents

that only produce clonal descendants. If each parent produces one clonal descendant, the ancestral genotype diversity is (at least) maintained despite the population being fully clonal, with no lineage becoming dominant. This diversity, the genetic divergence between genets and between ramets within each genet can even increase (but really slowly, see Reichel et al. 2016) by mutation accumulation. To observe a dominant clone at high frequency, it implies an uneven random drawing during life cycle and reproduction of the initial clonal diversity and of the mutations that may occur (either by genotypic drift or selection). In addition, recombined genotypes arrive every generations due to the not-so-rare events of sexuality – estimated from 10% to 40% in our populations, resulting into new genotypes. Rarer sexual events (one effective descendant on the total effective population size) are yet sufficient to maintain and even increase the clonal diversity within population along generations (see Bengtsson 2003, Ceplitis 2003, Reichel et al. 2016). That's why one must investigate genotypic diversity with other population genetic indices to properly estimate rates of clonality, as we did in this study.

2. The effect of the river basins on the genetic structure of this species.

As Ludwigia grandiflora subsp. hexapetala is an aquatic plant with a high clonal rate is expected that that the river basin might have a strong influence on how the genetic diversity is structured so that the expectation of genetic similarity should be higher within than between river systems. I wonder how this lack of independency between sampled sites was taken into account in both the statistical analyses and the estimation of genetic indices.

Answer: Yes, it is yet considered in population genetic analyses (hierarchal analysis of genetic identity, see Balloux et al. 2003 and Reichel et al. 2016 for equations). Population genetic indices are computed at the scale of a local population, and at the scale of the whole population sampled (e.g. Fis, Fst, Fit) in GenAPoPop (Stoeckel et al. 2024).

3. Statistical power of the developed SNP marker set

The authors argue that the statistical power of the developed SNP marker set to identify true ramets was exceptionally high given the low probability of two samples being identical purely by chance. However, these estimates assume panmixia which is somewhat unrealistic. To what extent can these expectations vary when mating is structured within populations or even within patches? As the interpretations of this study hinge on the reliability of these markers to distinguish true repeated genotypes, I believe that this potential bias should be addressed."

For instance, I question whether the statistical power to distinguish true ramets remains constant across populations, even when there can be significant variation in the mating systems, i.e., SC vs. SI populations.

Answer: pid-panmixia and pid-sibs are traditionally used to assess the statistical power of a marker set to distinguish between individuals (Gillois 1966, Waits et al. 2001). The same approach is used in forensics and on trials, even considering the case of twins (Evett & Weir 1998). This metric can be used to compare the "power" of different markers or between species with different marker set. They are needed to state if genotyping will make sense or not considering the question.

In the case of partial clonality and reproductive modes in general, pid are used in a logical negation reasoning: pid by panmixia (or psex when applied to each genet) gives the probability to observe two individuals by chance with the same genotype under such condition (random sexuality, see Jacquard 1969). Thus, if this probability is low (low chance to observe two individuals with the same genotypes under random sexuality), it means that when truly observing two individuals with the same genotype in our data, they are clones, and that the marker set is statistically sufficient to distinguish between two individuals coming from a sexual event (even if their parents were sibling for pid-sib) using this marker set. We also traditionally use in sexual organisms the probability of identity considering only mating between sibs as a

low range of the true pid value if we don't know the real mating system (see Waits et al. 2001). At the scale of the whole sampling study, pids indicate if the genetic identities will be really due to local specificities or due to statistical limitation of the marker set. Here, our pids values demonstrate that our 36 SNPs set are sufficient to distinguish between identical individuals that would have happened by chance under sexuality, and so they are so low that it would be parsimonious to consider that two identical individuals are ramets of a same genet. We hope this long explanation clarifies the interests of such metrics. As pids are very commonly used in population genetics studies, and as we were requested to decrease our manuscript length, we didn't add more explanations within our manuscript.
Now, about considering patches, it means that we have to define a prior sampled "population" range that match possible pollination distance to allow for mating. We sampled the 15 individuals hierarchically along a transect of 40 m, a distance that can be covered by any insect pollinator. Lines 240-243 of the previous version read "*At each location, we collected 15 stems (hereafter, 'individuals') along a linear transect of 40 meters. Along each transect, we randomly collected three stems at coordinates* $X_1 = 0m$; $X_2 = 10m$; $X_3 = 20m$; $X_4 = 30m$ *and* $X_5 = 40m$ *within a one meter-square quadrat.*" In this context, local pid values come from the structure of coancestry and reproductive mode with a sufficiently-resolutive marker set, as sexual crosses may have occurred.

Some other minor comments:
L31. I disagree with the use of acronyms for species names.
Answer: We agree (on acronyms in general) and understand but *Ludwigia grandiflora* subsp. *hexapetala* is a very long name and we had to use it a lot. Our manuscript is unfortunately yet so long that Recommender and Reviewers asked us to shorten it. Moreover, *Lgh* is commonly used in European official policy documents. We propose to keep the acronym, and only hope for a reappraisal of the *Ludwigia* nomenclature that will help in the direction of clarity and conciseness.
L490. SI4?
Action: We mislabeled the figure SI5 into SI4. We now correctly changed the label in the supplementary information.
L556 Do you mean asexual or sexual dispersal propagules?
Action: we added "vegetative" propagules, line 473.

Review by anonymous reviewer 1, 16 Apr 2024 12:49
I have reviewed the manuscript entitled "Reproductive modes of polyploid Ludwigia grandiflora subsp. hexapetala in western Europe: the effects of a late-acting self-incompatibility system and its absence on genetic diversity within populations" by Stoeckel et al. I enjoyed very much the authors' efforts to disentangle the genetic properties of plant populations made of individuals generated from different reproduction modes. In addition, the study organism has not a diploid genome, increasing even further the complexity of the system. Overall, I applaud the boldness of the authors and the skills developed here to face such a challenging study system.
We thank reviewer#2 for their reading and feelings about our manuscript.

The major problem of this manuscript is the style in which it is written. Sometimes, I felt like I was reading a methodological paper looking for a biological problem to stress its value, when in my opinion, it should be the other way around. Some symptoms of this problem can be found in the following issues:
1. The length of the sections, including a 2-page abstract, a 10-page introduction and a 14-page discussion, are disproportionate, which denote structural problems dealing with the contents.

2. Indeed, the introduction includes a series of theoretical paragraphs describing different biological topics of interest in this manuscript, whereas the discussion picks results one after another that are discussed in some cases out of context. Overall, it is hard to know and follow the rationale of the authors behind this confusing strategy.
Action: We now reduced abstract to one page, introduction to a bit more of 4 pages and discussion to 8 pages, with better clarification of our scientific results, novelties and messages. This action addresses issues 1 and 2.

3. The methods read like an endless list of methodological approaches that are not connected with the main questions, perhaps because such specific objectives, which should be tightly related to specific methods addressing them, are not well defined in the introduction.
Action: We reorganized and changed our writings all along the manuscript to reduce its length. The new version of the manuscript addresses this general comment. Most of our material and methods paragraphs explicit their purposes using formula like "to better understand the correlation between population genetic indices and estimates of reproductive modes" (lines 307-308) or using explicit title like "allele dosage", which interests were explained in the introduction to be major methodological step before estimating reproductive modes in autopolyploid populations using genetic diversity.

Overall, I see a lot of potential in this manuscript, but the authors need to identify the major and specific goals of this study. Then, the authors will be able to select the appropriate theoretical background to end up in the major and specific goals, which will define the set of methods required to address such goals. Finally, the discussion should also point to a given direction by interpreting the results in the context described in the introduction, as these two sections are tightly connected for the sake of coherence, readability and overall comprehension.
I do not provide comments on specific issues because I think that the authors have a major task ahead by reorganizing and rewriting the manuscript.

Review by anonymous reviewer 2, 15 May 2024 18:47
The manuscript uses molecular markers to understand the reproductive system and clonal capacity of the invasive Ludwigia grandiflora subsp. hexapetala in Western Europe. To this end, the authors developed a set of SNPs from which they selected 36 and sampled a large number of populations to find out whether the two morphs described in the species differ in their vegetative and sexual reproduction. The topic is of great interest both to the field of reproductive biology and to the management of invasive species, which requires knowledge of their reproductive systems. The main results show that although cloning is the main form of reproduction, sexual reproduction is also present in the populations, which is different from the results obtained in other areas using other molecular markers.
Answer: We thank anonymous Reviewer#3 for this factual summary.

In spite of the interest of the topic, I have some important concerns about the manuscript. In previous papers, the authors had described the presence of two stylar morphs in this species finding that they were interfertile, one being self-incompatible and the other compatible. This is a very unusual system for several reasons: differences in the reproductive system associated with the stylar morphs; the presence in one of them of the most unknown type of self-incompatibility of all those described (late-acting) and presence of only monomorphic populations. Because of the unusual and interesting nature of the system, I have read the authors' previous work, as well as others mentioned by them. These readings have raised doubts about the study system. The taxon is part of a polyploid, highly clonal complex that shows high rates of hybridisation. Some taxa in this complex are invasive in Europe. In that complex, L.

grandiflora grandiflora and L. grandiflora hexapetala are very similar (according to Grewell et al., 2016) but differ in minor phenotypic differences and in chromosome number. These two taxa and others in the complex are interfertile and F1 hybrids are viable (Zardini et al., 1991; Grewel 2016). The study by Zardini et al. 1991 (Systematic Botany 16: 242-244) provides quantitative data indicating that L. grandiflora grandiflora and L. grandiflora hexapetala differ in both style length and flower size. The authors point out that the two morphs show differences not only in style length but also in flower size, which makes me doubt whether they are the same or different taxon. It is striking that the two morphs have not been described in their place of origin, despite extensive studies of the reproductive system of the complex. The absence of polymorphic populations is also remarkable. Finally, the authors described late-acting self-incompatibility in one of the morphs that is intra-incompatible. In species with LSI, plants may be intercompatible, as this system differ from the conventional SI systems associated with different floral morphs in the population (Gibbs, 2014 for a review). All these facts strongly suggest to me that they can be different lineages and not two morphs of a single taxon. Note that it is not uncommon for closely related species to interbreed, even if they are self-incompatible (e.g. Cistus). Assuming the baseline situation of the two morphs, which was previously published by the authors, the data obtained could be of interest, as this is a taxon with a high invasive capacity in Europe, so knowing how it reproduces is essential for establishing eradication protocols. However, the first step towards good management of invasive species is proper identification. The authors have been working with Ludwigia for a long time, so these questions have probably already been raised. However, I think that in order to avoid any doubt for the reader (as was the case for me), some aspects should be made clear in the manuscript. Therefore, it would be necessary to check the chromosome number of the two morphs studied, because if they are different, the molecular analyses would have to be reconsidered. It is possible that the authors already have these data, so they can include them in a section on the species studied. However, they will need these data for some of the LS and SS populations studied.

Action: See our answer to Recommender above. We now better communicate previous results about the 100% fruitset and the 100% viability of first- and second-generation descendants obtained from controlled crosses performed between L- and S-morph individuals, that contrast with fruitset and viability of descendants of crosses performed between individuals of two sister-species. Lines 116-120 now read: "*Lgh presents heteromorphic flowers corresponding to two floral morphs: L-morph individuals develop long-styled flowers and S-morph individuals develop short-styled flowers, that cross and result into 100% viable and fertile F1 and F2 descendants (Portillo-Lemus 2021, Portillo-Lemus et al. 2021) while inter-species crosses result in chlorotic and unfertile descendants (Barloy et al. 2024).*"

We also added new results as requested by Reviewer#3 (lines 346-350 and Figure SI3): we counted the number of chromosomes in two S-morph individuals and five L-morph individuals, that add to the S-morph reported in Barloy et al. 2024 and to the L-morph reported in Bou Manobens et al. 2019 that now we also cite. All these plants show 80 chromosomes, the number of chromosomes of *Lgh*, while *Lgg* has 2n=6X=48 chromosome (see Barloy et al. 2024 for the corresponding photos of *Lgg* karyotypes using the same method we used here).

Another concern is that throughout the manuscript the authors assume that all LS populations are self-incompatible and SS populations are self-compatible. They have only sampled 15 individuals/population (some of which are clones, so the number of genetically distinct individuals is smaller) in which they have observed whether are LS/SS and assumed differences in the reproductive system. I think this assumption could be made in the discussion (in the section starting on line 652) but not in the results. The difference in the reproductive systems of the morphs seems to come from an earlier study with hand pollination in only 7 populations.

In species with polymorphism in the self-incompatibility system, variations between populations and even between individuals within a population are common. Therefore, if the authors only know the morph type, they should not assume the self-incompatibility system as they have not studied it in these populations. The results show that of the 53 populations studied, 40 have only LS plants and 13 have only SS plants, but by only looking at 15 plants per population (and some were ramets; e.g. one population had only 3 genets) they cannot assume that all plants in the population are similar. Therefore, I suggest that the results should be presented based only on what is known about the individuals studied (LS/SS), removing any reference to self-incompatibility until the discussion.

Action: We now clearly write our approach lines 159 to 171.

We now refer to individual and populations as "L-morph (LSI) individuals" or "supposed to be LSI", and "S-morph (SC)" or "supposed to be self-compatible" all along the manuscript. These modifications clearly state that we supposed the self-compatibility status of individuals and populations from 1/ the binary distribution of the floral morphs fully matching self-compatibility using 105 individuals sampled from seven western European populations among the 53 we studied here, with a null probability to obtain such a binary distribution with a random or even relaxed association of self-compatibility and floral morphologies (Portillo-Lemus et al. 2021), 2/ the fact that 105 among the 795 individuals of the current study were directly tested (for some of these populations and individuals) and 3/ the fact that we either observed fruits or not in these monomorphic populations: S-morphs we conjectured to be self-compatible were all fruitful in the falls; L-morphs we supposed to be self-incompatible were all with no fruit in the falls, matching again our reasonable and parsimonious hypothesis.

Another important concern relates to SNPs. The species described as 'alloautodecaploid' would more accurately be described as an allopolyploid with a complex origin.

Action: We removed this term used once line 154.

In addition, GISH (genomic in situ hybridisation) is useful but not the most appropriate tool for investigating such origins. For genomes with complex origins, comparison of genome assemblies is essential to draw accurate conclusions.

My main concern is with the selection and validity of the SNP set. The study used RADseq from a pool of individuals from two species. This sampling design can be problematic for species with asexual reproduction. If all individuals in a population are clones, the result will only provide heterozygous markers for a single genotype. These markers may not be informative for other populations where different genotypes are fixed.

Action: We now clarified lines 187-201. We RAD-sequenced two pools, one per species, of 15 individuals each, sampled across five *Lgh* and three *Lpm* populations.

Answer: No, there weren't all from the same genet: see the huge diversity of genets within the 53 populations (more than half the individuals have unique multi-locus genotypes, and the remaining half belong to multiple different genets, Figure 1), and we excluded pure heterozygous SNPs for this precise reason, as routinely done when working on partially-clonal genomes. We kindly remind that we kept all SNPs that amplified on all 795 individuals across 53 populations and at a large geographical scale (Figure SI1). So they ARE informative and thus, it is not a speculative theoretical discussion as presented by Reviewer#3.

In addition, the filtering steps to select high quality SNPs are not clearly described. A crucial filtering step is the removal of linked SNPs. If the SNPs in the final set are linked, the resolution is significantly reduced.

Action: We added line 194 "one per aligned sequence". Please see our linkage disequilibrium values that also address this comment (Figure 3 and in the supporting data document).

There is also a lack of clarity regarding the selection of the "autotetraploid" part of the genome. The authors do not appear to have sequenced other species (apart from Lpm), so they cannot be sure that the markers are not shared with other related species.

Action: We now detailed the three independent but convergent results supporting the autotetraploid state of our SNP set, lines 202-211.

To ensure beyond the literature evidence and the way we developed our SNPs that these SNPs were in a tetraploid state in *Lgh* genome, we computed Akaike's information criterion from the maximum likelihood of the best genotype considering the distribution of observed (sequenced) alleles as a function of the ploidy level, using the similar approach proposed by Burnham & Anderson (2002).

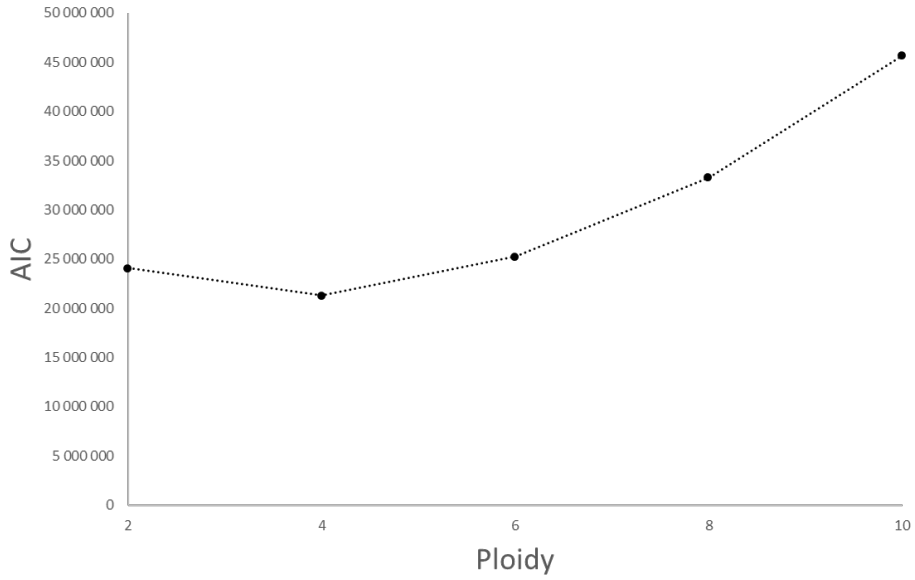
$$AIC(ploidy) = \sum_{l=1}^{l=36} \sum_{i=1}^{i=795} -2 \log \mathcal{L}\left(genotype_{i,l} | ploidy, data_{i,l}\right) + 2K$$

where data are the distribution of sequenced alleles among A, C, G and T, and K is the number of possible genotypes given the ploidy and the four possible alleles (A, C, G and T).

This result, added with the sequencing approach we developed (SNPs shared by *Lpm* and *Lgh*), added with the previous results to decipher the genomic contributions of *Ludwigia sp.* to *Lgh* genome (Barloy et al. 2024), again argue that tetraploidy is the best model for these SNPs.

We added these results to the open data that supports our work.

Ploidy	AIC
2	24056742.7
4	21307975.4
6	25245336.4
8	33292165.5
10	45666746.4



Finally, the title does not reflect the content of the article because the authors did not specifically study the SC/SI systems of the populations studied. I find the discussion speculative in places (for example, the paragraph starting on line 783).

Action: Right, we studied L-morph and S-morph individuals that were previously demonstrated on 105 individuals (15 per population, across seven populations that are also investigated in the current study) to perfectly match self-compatibility and self-incompatibility. We now added a sentence specifying this important precision in Material and Methods, lines 164-175, that reads

"We visually identified floral morphologies of flowers found along each transect within the sampled Lgh populations. In a previous study on seven populations among the 53 studied here (underlined population names in Figure SI1), one hundred and five sampled individuals resulted to identify a binary distribution of floral morphology with formally-identified self-incompatibility types: all L-morph individuals were LSI typed and all S-morph individuals were SC typed (Portillo-Lemus et al 2021) while all these individuals succeeded to cross and give viable and fertile plants. Interestingly, these two types of populations spatially distribute in monomorphic populations along different rivers. We supposed for this study the LSI versus SC status of individuals and populations using their floral morphologies: L-morph individuals were supposed to develop a LSI system and S-morph individuals being SC. To support this conjecture, as done in Portillo-Lemus et al. (2021), we checked the fruitset in each of the 53 sampled and genotyped populations: Low and even no fruitsets were found in L-morph individuals and populations, while full fruitsets were found in S-morph individuals and populations, in agreement with our conjecture."
We also added this precision in the legend of Figure SI1 where we underlined the seven population sites in which we extensively tested 15 individuals for their self-incompatibility status. One individual from each of these population was also used to count the number of chromosomes.
The new manuscript with reduced discussion clarifies this point, every time we discuss LSI and SC.

Minor
-Line 439. Change 99 individuals into 99 ramets
Action: Done.
-Change reproductive insurance to reproductive assurance.
Action: Done, we change lines 128, 544 and 582.
-Lines 491-505. It would be much clearer to the reader if these indices were in the form of a table separating the two morphs.
Answer: The requested table is already publicly available in the open data in a simple format (xlsx) that is joint to our manuscript, in the tab "Genetic indices per pop". Results are yet separating L-morph and S-morph populations. This is a huge table (due to a huge dataset and amount of work) reporting 1728 cases along 54 lines and 32 columns. This table is mentioned in the supplementary information as Table SI2.
-Lines 645-650. Can the difference in the number of individuals sampled affect these results? 195 vs 600
Answer: No, the difference in the number of individuals sampled doesn't affect these results: As indicated in Material and Methods (in the previous version lines 392-396, now lines 335-339), we used non-parametric Kruskal-Wallis statistical tests and post-hoc pairwise tests for multiple comparisons of mean rank sums (Nemenyi's test) to compare the 13 versus 40 measures of each population genetic indices, that consider the two different number of observations.