Dear recommender, Sara Magalhaes,

This preprint introduces a methodology aiming at studying matings between discrete types of individuals. Specifically, the methodology allows both for the estimation of mating propensities and for performing selection among different statistical models: it can compare models that differ in their assumptions about the mating propensities, which translates into investigating whether observed mating patterns are consistent with either sexual selection, assortative mating, or both. The author has implemented the new methodology in a free software called InfoMating.

The author performs a rigorous methodological development which is conceptually interesting and which could benefit many empiricists. I want to highlight the conceptual interest (he is not just trying to sell a free software). The methodological development is directed connected to evolutionary biology and could interest people working on sexual selection, mate choice and/or speciation.

I had provided extensive review comments on an earlier version of this manuscript for a journal. At the time, I was particularly criticising its presentation (and not the formal content, which was very good from the first version I got a chance to read). I am very pleased to say that the author has done a great job, substantially revising his paper in light of my (and others') comments. Albeit remaining a paper rather slow to digest, the presentation is now much clearer and all my main criticisms have now been addressed. So I do recommend this paper for… publication(?)… or whatever PCI Evol Biol will do with it! If the author wants to submit this, I think it would now have good chances in many general journal in biology that is not reluctant to publish a few greek letters.

If it is possible, I think that minimal effort could further improve the paper. What is presented is mostly very good aside one part that I still find confusing: I do not quite understand how the analysis of the simulated data was done. I assume that this is the result of a model averaging, but I do not see where the models considered have been described. In each simulation assay, the author simulated the data according to one model (as described in Table 2), and I am tempted to assume that the estimation of mating propensities was based on all 5 models shown in the table, but as far as I can see, this is not mentioned. It is also unclear how often the best model turns out to be the one used to simulate the data. Doing so would be one way to actually fulfil the expectation introduced at the beginning of this section ("To test how well the exposed methodology is able to distinguish the different classes of models").

Also, with respect of the simulation, I would have like to see just a little bit more. In particular, since the author concludes "[t]he proposed framework should ideally work under scenarios where the availability of individuals is not affected by the matings that have already occurred", a question is naturally popping up in my mind: What happens when this assumption is violated? In some taxonomic groups, such as mammals, this assumption is actually very doubtful. Perhaps not in the sea snails that the author studies, but even in largely abundant species the assumption is likely to be violated if there is local competition for mates, or if the number of potential mate encountered is low (due to spatio-temporal constraint on mate sampling). For these reasons, I would have like to see results of an additional simulation analysis based on more limited number of individuals. This would have shown how robust the inference method is in the case of the departure of the assumption of constant frequency of types. This would have thus helped empiricists to anticipate with respect of their favourite species if InfoMating is a tool that may be suitable for them or not. Perhaps InfoMating allows for one to run such tests easily, I do not know. Also, it would be fair to mention local competition and spatio-temporal constraints in the conclusion, as just having more individuals than mating pair really does not cut it.

Another small caveat is that

Aside these minor criticisms, I have a small list of trivial comments:

-L61: I would drop the "However" as there is not explicit contradiction.
-L86: I would put caps on "pair total indices".
- Eq2: the A and t(s) do not quite match the convention used in other equations and do not seem defined.
-L165: "withIN all the other".
- L166: perhaps "When random mating occurs", instead of "When mating is at random".
-L246: "they mate equal in average" is not quite idiomatic.
-Fig2: Say that rows are males and column females.
-L334: What is a "joint isolation index"? This falls from the sky and should be introduced (not simply referred to another paper).
-L356: "When c = 1", not "With c = 1".
-L382-383 and elsewhere: two different signs are used to represent alpha and it makes things a little confusing since they do not look alike at all (at least in my pdf viewer).
-L403: "at least some is" seems grammatically incorrect.
-L430: dash missing between "so" and "called".
-L551-553: Some extra dots are present in the text of the table legend.
-Fig3: Why "Rep3" in the model name?
-L617: Is SfemC-2Pc actually better than D-2P-Rep3?
-L639-641: I would drop this confusing sentence because sexual selection is retained in the final best model.
-L656: space missing before "or".

Congratulation to the author!

All the best,

Alexandre Courtiol