

## MS PCI EvolBiol by Etoundi et al

The questions addressed in the paper by Etoudy et al. are of high interest and refers to an evolutionary topic of general interest. The dataset is impressive and the statistical treatments (based on previously published methodological frameworks) are innovative and insightful in the complex context of worldwide Corbiculata genetic variation. The paper brings new insights into the biogeographic origins of androgenetic Corbicula lineages, both in the native and invasive ranges of the species. I believe that this work deserve publication and recommendation in PCI EvolBiol.

I feel however that the ms could (should) be improved on different aspects to make it (even) more interesting and readable (especially for non-expert readers). I propose that the authors consider the following comments (four major points followed by a number of minor points) to improve their ms.

1/ The study context has been nicely formalized in the Introduction section which I globally like very much. I feel however that the addition of a (very first) general section about variation in sexual systems with a particular emphasis on the rare and intriguing androgenetic reproduction system would give a more generic perspective to this work; see for instance Rey O, Facon B, Foucaud J, Loiseau A, Estoup A. 2013 Androgenesis is a maternal trait in the invasive ant *Wasmannia auropunctata*. Proc R Soc B 280: 20131181.  
<http://dx.doi.org/10.1098/rspb.2013.1181>.

2/ The Results section is difficult to read especially for non-expert readers. The authors should change its structure to make it easier to read. At the moment results include too many details that are presented somewhat as a list which “submerges” the reader. Try to follow a kind of two-step presentation. For each result section, start first by describing “the big picture” of the results so that readers get the main message(s). Then you can go into details that might be of importance for specialist readers.

The illustrations (i.e. figures) that should be kept in the main text should preferentially correspond to those that integrate results on multiple-loci. Concretely, keep Fig 1 (cited from the introduction which is fine to me), then keep Fig 3 and add Fig S3 in the main text, as the two later figures provide some nice synthetic summaries/illustrations of results. You might also include in the main text one figure of the locus-haploweb treatments (choose the “best” locus), but other locus-haploweb figures are redundant in what they show and should be hence included in the Supplementary Materials.

(see below for presentation details of the figures that should be improved to make them easier to interpret)

3/The discussion section is nice. My only point is that the authors should avoid pointing recurrently to all their tables and figures (see a symptomatic example with lines 403 – 405). This is expected in the Results section but not in the discussion section.

4/ As a last main point I have to say that I am (very) surprized by the absence of triploid molecular signatures in the analysed dataset, a feature which would be expected as indicated by Fig 1b. How do the authors explain that? This point might be evoked somewhere in the Discussion section.

Other following points correspond to minor points that might also help improving the ms:

- Abstract; L23 “egg parasitism between distinct lineages”: please explain what this notion means through a few simple words.  
Generally speaking, I am not fan of the “egg parasitism” terminology used all over the ms. It is not clear to me why the term “parasitism” is used. The “parasitism” terminology” involves precise biological meanings which do not seem to apply here. I would prefer that you just refer to the genetic/evolutionary mechanism itself (mt or nuclear DNA capture) without referring to “parasitism”.
- L61: the reference Schwander & Oldroy 2016 is missing in the reference list.
- Figure S1 is great and very useful: refers to it more thoroughly in the material and methods as well as later in the ms (for example in the legends of haploweb figures). Regarding Fig S1 itself: marker of interest → say clearly that it actually means sequence haplotypes. “median-joining network”: please explain/specify. “Sharing mutually exclusive pools of alleles” is unclear → please reword and specify. Fields of recombination are clearly delineated by hand which is great: pleased do the same in your illustrations of haploweb-based figures.
- L281: it is not clear from Fig 3 where the sexual lineage *C. sandai* is located on the figure
- L320: “reveal”...really ??? Please reword to tone down this term here.
- Check all over the ms if the word “indeed” is really appropriate cf. a very “French –English” wording. Most of the time the sentence is fine without “indeed”.
- L366: Two patterns → evolutionary scenarios might be better.
- L425: “Genetic or Evolutionary” origin would be more appropriate
- Fig 2a,b c: avoid repeating legends – try to delineate fields of recombination by hand to make them more visible.
- - Fig 2c and S2: FFR2 and 4 seem to have the same colour. Use clearly different colour codes for the different FFRs.
- Fig 2d: what means FW ??? Please specify.
- Fig 3: great figure: please add a few important things in the figure. Help readers by specifying in some ways the androgenetic and sexual haplotypes in the figures. Specify the native and invasive haplotypes too. Finally please specify where *C. sandai*, *C. moltkiana*, *C. fluminalis*, and *C. African* are located in the figure to help readers to following what you mention (recurrently) in the results (and discussion) section.