Reply to F. Guillaume

Comment: Overall, if writing a recommendation from the current manuscript, I would summarize your findings as being in line with a model of stabilizing selection with drift on a set of locomotion traits. The main evolution of the traits during stasis is a general loss of genetic variance across the board, which is compatible with a model of stabilizing selection after hybridization among divergent inbred lines, but also with drift. There is no need to invoke directional selection at this stage (see below). The transient changes in G-matrix orientation seem rather limited, although not fully explored, limiting our ability to draw strong conclusions on their importance. Similarly, the relative importance of selection and drift has not been formally tested. The study thus remains a general assessment of changes in genetic variation over time without a strong conclusion regarding the causes and consequences of the changes. I think you can strengthen your manuscript if you address the points raised here, especially regarding the changes in G and the selection regime.

Reply: We thank you for giving our manuscript a thorough read. We reply below in detail to your comments but agree with your general description of our analyses. We believe that the pattern of transient changes of G-matrix orientation between our three replicates, because of drift, is a strong and interesting result from the experiment. Obviously, however, most of the changes were of a reduction in the size of the G-matrix. We detail below why we think it is robust. The consequences of these transient variance-covariance changes were not tested, and we only refer to potential consequences in the discussion. Similarly, we agree that our claim of directional selection is an *ad hoc* conjecture from the founder isolates observations that is not strongly supported.

Comment: Directional selection: I am still not convinced that you have a strong case for directional selection in the domestication phase (from A0 to A6140), in contrast to your argumentation for it, eg. on lines 529-531, 558-559. As point in case, you provide ample evidence of a lack of directional change in the traits during that period. It is thus misleading to invoke directional selection in the Discussion to explain genetic divergence. I am not denying that there might have been directional selection during hybridization of the founders, but apparently there was no directional selection during domestication, and it is not necessary to explain your results. The manuscript should clarify this point.

Reply: We agree with you that we do not show directional selection, and that it does not explain divergence or differentiation. We have removed mentioning directional selection, except during the hybridization stage of the founder strains (new line 564). We have eliminated or reworded the other sentences in the Discussion that you mention. The first paragraph of the discussion was completely restructured.

Comment: Your results can be interpreted as stemming from the effects of stabilizing selection on an initially broad phenotypic/genotypic distribution (A0) generated by admixture of the inbred founder lines and thus large segregation variance. Stabilizing selection sifts through that initial

variance, eliminating extreme phenotypes during domestication and eventually leading to a selection-mutation-drift-recombination equilibrium. There is no need here to consider directional selection on the traits nor "effective" stabilizing selection, just stabilizing selection is enough. Please make this clearer in the text (ie, lines 529-531, 533-535, 558-559).

Reply: We agree. We have eliminated "effective" whenever mentioning stabilizing selection, and reworded the relevant sentences to clarify how genetic variance is lost after hybridization (initially perhaps because of stabilizing selection, after domestication mainly because of drift, even if we show stabilizing selection among the inbred lines).

Comment: Also, the model of stabilizing selection on a plateau and directional selection outside the plateau that you mention is, in my opinion, not different from a model of stabilizing selection with a broad adaptive surface around the optimum (at least broad relative to the phenotypic variance). This doesn't deny directional selection when the population mean is away from the optimum. This is to clarify my thoughts, not to ask you to change your phrasing but you can make the parallel with the classical model of stabilizing selection for a multi-trait optimum with a broad selection surface.

Reply: Yes, we agree, see previous replies.

Comment: **G-matrix divergence and orientation**: divergence in G-matrix orientation is not directly evaluated in the Results section but mentioned in the Discussion (lines: 557, 648), and also regarding alignment with the selection surface (lines 591-593). Most mentions of alignments are relative to gmax of A6140. You do not explicitly test for a change in the orientation of gmax in the CA populations. Assessment of changes in CA lines can bring evidence of transient changes in genetic covariances and arguments in favor of their importance relative to selection. In particular, it would help asses if "cryptic evolution of genetic covariance" has happened (line 648). I have two requests:

Reply: Thanks for this suggestion. True, the eigentensor analysis confounds both size and shape of the G-matrix divergence and differentiation. We now present evidence for divergence in the direction of generation 100 CA gmax, g2 and g3, by showing that their alignment with the ancestral gmax, as measured by the angle between them, does not differ from null expectations, meaning that evolved G-matrices no longer align with the ancestral one (new panel D in Figure 2). We further present, among the generation 50 CA populations only, evidence of transient differentiation by showing that the angle between their gmax remained aligned with the ancestral gmax in only 2 replicates out of 3 (new Figure S11).

Regarding the alignment with the selection surface, please see reply below.

Comment: clarify alignment of G's between A6140 & CA-100; reduced variance along one phenotypic dimension can be caused by the general reduction in mean variance (smaller trace) **and** changes in orientation of G. As of now, it is not clear what are the changes in G's orientation in the CA lines and how they affect their reduced variation along gmax of A6140.

Interpretation of the tensor analysis relative to changes in orientation of G's is not clear to me. A more direct comparison of G's alignement will help you clarify the changes in alignment.

Reply: We have added an extra panel to Figure 2 showing angles between gmax vectors of the A6140 population and the CA100 populations. We agree that the tensor analysis is not fully informative in terms of shape variation. This new metric allows us to test for an alignment of the different populations' gmax vectors. We conclude that there is a change in the gmax vectors due to the overall decrease of genetic variance due to drift. Similarly, we added as a supplementary figure the angle measurement with the CA[1-3]50 populations. This figure could be moved to the main text as the direction of eigenvectors does not depend on how much genetic variance there is (see previous round of revisions), though we prefer to present it as supplementary so not to confound the reader with potential scaling issues.

Comment: <u>clarify alignment of selection surface with G's of CA lines</u>; similarly, it is not clear how changes in G's orientation in CA lines affect patterns of genetic variance along axes of the selection surface nor how G-matrices align with that surface. Only mention is about alignment of y3 and gmax of A6140 (lines 523-524). A direct comparison with G's orientation in the CA lines is also warranted.

Reply: In the previous version of the manuscript we showed (1) the patterns of genetic variance along the selection surface for A6140 and CA100 populations (Figure 5), (2) the divergence axis (main axis of variance reduction between A6140 and the CA100s populations) is aligned with y3 (Fig. S12 [now Fig S13] top panel), and (3) the differentiation axis (main axis of variance difference among the CA50s populations) is aligned with y3 (Fig S12 [now Fig S13] bottom panel). Thus, we disagree that we have not provided clear evidence for a lack of an alignment between the G-matrix and the selection surface. The decrease in genetic variances along the selection axes is sufficiently explained by drift. However, we conducted a new analysis, shown below, where the angles between the gmax of ancestral and evolved populations with the selection axes are calculated. You will see that indeed, y2 aligns with the A6140 and the CA3100 gmax (and also y4 with CA1100 gmax), as the angle between them is smaller than the null expectation. You will notice, nonetheless, that the alignment occurs because only in these axes there is genetic variance (main Figure 5). y2 and y4 are not under selection, we thus believe it is misleading to (implicitly) infer that the G matrix has evolved because of selection, as it would be the case of a less careful reader. We're happy to show this analysis as a supplementary figure, if you think it is critical.
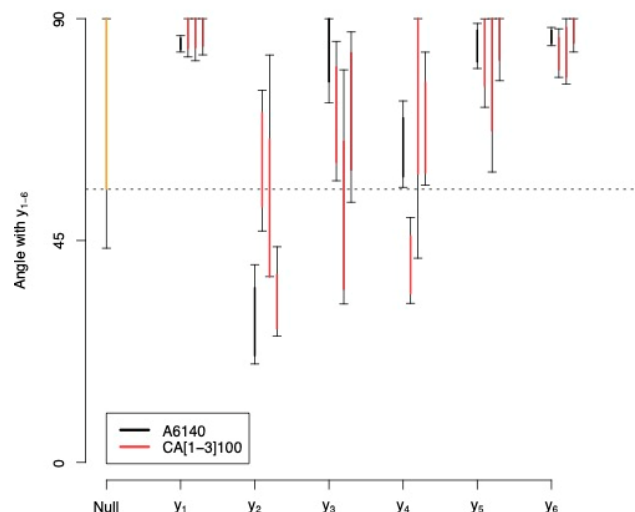
Figure legend: angles between the gmax vectors of the ancestral (A6140) and evolved (CA[1-3]100) populations with the selection surface. A6140 and CA3100 align most with y2 while CA1100 aligns most with y4. All matrices align only with axis that are under drift. Angles and null expectation as defined in the main text.

Comment: **Testing for differentiation by drift vs. selection**: It might be that drift is the most important driver after domestication, but we do not know the relative importance of stabilizing selection vs. drift because you do not formally test for it. Also, it is unclear how you tested for a loss of variance matching with drift expectations. Please provide the model you are using and how you parameterized it. It is unclear how you modeled the multivariate loss of variance along axes of the selection surface with that model.

Reply: We don't think that testing the relative importance of drift and selection is pertinent here. As you wrote above - and we agree - we observe a lack of genetic variance along the axis of selection and therefore there is no possible response to selection. Drift alone explains our results. See also the previous reply.

We modeled the loss of genetic variance due to drift as a constant per generation loss of $(VA/(2*Ne))$. This is now clarified in the figure caption.

Comment: This calls for a multivariate Fst-Qst analysis to test whether genetic changes are dominated by drift or selection, over time or between lines. My understanding is that you have genomic data for some of the lines, from which you could estimate an Fst distribution. That additional analysis is likely demanding but would greatly add to the understanding of the system. A suggested approach is driftsel by Ovaskainen et al. (A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. 2011. Genetics 189(2): 621-632; Karhunen et al. 2013. driftsel: an R package for detecting signals of natural selection in quantitative traits. Mol Ecol Resour 13(4): 746-754).

Reply: We understand your argument and we thank you for pointing us toward these resources. We are currently analyzing the genomic data and plan to publish the results in a separate manuscript. We believe that the current manuscript is already quite dense and difficult. Furthermore, including comparisons between Fst and Qst would require a discussion about the limitations of this approach (particularly in the inference of stabilizing selection), which is out of scope for this study. As we argue above we do not believe that selection has an impact on G matrix evolution.

Comment: **QTL data (lines 567-569)**: this mention of two QTL results brings more confusion than clarity. It is unclear how only two QTL of unknown effect size could be responsible for the maintenance of genetic variation in the locomotion traits. What about genome-wide mutations over 140 generations in a population with $N=10^4$? That part could be removed.

Reply: We have eliminated this discussion.

Comment: **Divergence versus differentiation**: I understand that you use the two words to differentiate the "between" and "within" generation differences in G's, while using the same method. This is confusing. Both words bear different meanings and would normally be tested with different methods. Differentiation is usually assessed with a Qst index (or Fst), for instance (ie, a ratio of variance). I would clarify the use you make of the two terms relative to your two types of comparison, or find other words.

Reply: Our terminology does not depend on the method for comparison, it is just a generally used definition in the literature of differences between and among ancestral and derived populations (see for example the results presentation in line 491). We do not see how the terminology is confusing. Indeed, the power of experimental evolution relative to comparative studies in natural populations is to know the ancestry of the populations being studied and thus to define divergence and differentiation independently of the methods being used. We have nonetheless defined both terms in the Methods section (line 319).

Comment: **Abstract**: lines 14-16, it is unclear what are you referring to with "large phenotypic scale" and "local phenotypic scale", no mention of such scales are done in the rest of the manuscript.

Reply: Thank you, we removed these sentences.

Comment: l55 : "little unexplored" -> "little explored", or "unexplored"

Reply: Corrected.

Comment: l119: please specify which population is used for the samples

Reply: We have added a table with this information (Table 1). It can also be found in our github page. Line 119 (now line 118) refers to general assay conditions for any kind of material, population or inbred line. We have clarified this point.

Comment: l124: "generation 4-6 generations" -> phrasing unclear

Reply: Reworded.

Comment: l524: "Overall there is a strong alignment of both divergence and differentiation axes with y3 (Figure S12), and thus with gmax (see above)" -> phrasing unclear, it means that "divergence and differentiation axes" are aligned with gmax, but I guess you mean to say that y3 is aligned with gmax. Also clarify where exactly we should be looking "above".

Reply: We refer to the (i) divergence and (ii) differentiation axis as the main axis of differences obtained from the tensor analysis between (i) A6140 and CA[1-3]100 and (ii) among the CA[1-3]50 populations (the e_{11} vectors of each analysis specifically). The sentence was reworded, hopefully it is clearer.

Comment: l573: "however, the strength selection" -> "however, that the strength of selection"

Reply: Corrected.

Comment: Figure 5: please add the meaning of the color code as a legend

Reply: Done.