Response to the reviewers' comments.

Editorial comments

While both reviewers and myself find the study topic interesting and timely, at present, your manuscript would need to be revised substantially to be acceptable for publication in GENETICS. The manuscript has a strong focus on maize and thus the generality of the studied scenarios and of the conclusions is limited in scope. To increase the impact of the study and for publication in GENETICS, we would require at least one of the following: 1) to increase the generality of the predictions to different crops, and/or 2) to test the predictions of the model on available gene expression data in maize (see suggestion by reviewer 2).

The first can be achieved by investigating how the characteristics of various well studied crops (different bottlenecks times, bottleneck strength, timing of domestication, initial standing variation, genetic architecture of traits, history of domestication, strength of selection shift, effect of dominance,...) do generate different output and predictions regarding gene network evolution.

The second requires to show the applicability of predictions based on a limited set of simulations with limited amount of genes, to a more complex dataset with real expression data in a species with a wealth of available characterized gene network elements underpinning key domestication traits.

After having carefully considered both options, we chose the first one, and ran simulations to explore three additional plant domestication scenarios. Although an analysis of available transcriptomics datasets from domesticated plants and their wild ancestors is an interesting research line, the RNAseq dataset from Lemmon et al. was not produced in that aim (they were looking at cis-regulation in F1 hybrids), so that it requires a complete remapping of the sequences on the reference genome and downstream analyzes. This would add substantial supplementary material to the manuscript (including a whole new set of figures, and several pages of bioinformatics methods), and it would more conveniently be the topic of an independent paper. Comparing simulations across systems however (with the perspective of analyzing corresponding datasets in a future study) is a great adding to the manuscript. We thank warmly the editor for this suggestion.

In addition, note that both reviewers also have technical comments and concerns on running the simulations that would need to be addressed in a revised manuscript. Specifically, reviewer 1 points to the possibly short burn-in phase and asks for clarifying/justifying some modelling assumptions. Reviewer 2 asks for clarifying some modelling assumptions (role of standing variation, definition of cis-regulation) and to represent the variance between simulations (repetitions) to demonstrate the robustness of the results/conclusions.

We thank both reviewers and the editor for their clear and constructive remarks. We did our best to address the reviewers' concerns, fix the inaccuracies in the manuscript, and rerun simulations when necessary. Most of the changes fall into the following categories:

* Improving model description and avoiding misunderstandings. The modeling framework is not trivial, as it couples a gene network model (which provides the genotype-phenotype map) and a population genetics model (which provides the evolutionary framework). We did our best to answer reviewers' questions, and to provide additional explanations in the paper (including two new explanatory figures).

* Discussing model limits. We acknowledge that some parameters (such as the number of time steps or the duration of the burn-in) were chosen as a compromise between realism and simulation time. We also avoided unnecessary complexity layers in the model (dominance and interactions in regulation, gene expression to phenotype mapping…), although these could be added at the cost of additional parameters to explore.

* Running new simulations to address reviewers' and editor's questions, including (i) three additional domestication scenarios, mimicking the domestication of tomato, pearl millet, and african rice; (ii) four additional network parameters (less and more time steps, less steps to compute gene expression variation, and no selection on instability) to assess the potential consequences of model choices.

Note that the simulation setup is highly computer-intensive, and pushes our hardware to its limits. For instance, as noticed by reviewer #1, the "burn-in" stage was a bit short to reach equilibrium. We thus ran longer burn-ins for the main simulation runs, but not for all conditions, as rerunning all simulations with a longer burn-in would have taken >200,000 CPU hours, which would have limited our ability to run the new simulations proposed in the revision. We therefore had to consider trade-offs between testing new ideas or new parameters, resubmitting the manuscript within a reasonable time frame, and (of course) CPU availability and the energy footprint of such a computer-intensive work. We hope that our decisions regarding such trade-offs will be considered reasonable by the editor and the reviewers.

Reviewer #1

The paper by Burban et al, describes a simulation study of gene regulatory networks to predict/model/simulate the changes of maize during domestication and study the effects of domestication on the genetic level of GRNs (higher than that of the sequence). It uses the well-known and classic Wagner's GRN model with several modifications (the most interesting one is the modification of the discrete output to the continuous output).
The paper is very well written. Introduction and abstract are fine, giving the right background for the reader to follow. Methods are clear (some comments see below) and results and discussions are well presented. I have a few comments that I'd characterize them as not very difficult to be addressed and I think they will further improve the quality of the manuscript.

We thank the reviewer for constructive comments.

Unfortunately there are no line numbers (please put line numbers if there will be a next revision cycle).

We apologize for the inconvenience (Genetics encourage a "free format" first submission, and line numbers are not added in the generated pdf). The revision was prepared according to the official Genetics guidelines, including line numbers.

abstract, "... such as the loss of genetic diversity".
Even though, this is in general of course true, I'd add: "compared to adequately large wild populations". This is because domestication can also increase genetic variability a lot. See for example the dogs.

We agree this is a general pattern in plants, and it may not apply to dogs (or other animals). We changed the abstract to “Domestication of plant species” instead of “plant and animal species”, and we modified the sentence according to the reviewer’s suggestion.

In methods: (Wagner's model)
I guess the most critical comment is why 24 steps?

We thank the reviewer for the question, which helped us to notice an inaccuracy in the methods section, and encouraged us to check rigorously the effect of the number of time steps instead of assuming it was unimportant.

The number of steps remains arbitrary: with few, most networks will be out of equilibrium; with many, simulations would waste a considerable amount of computer time, as simulation time is roughly proportional to the number of steps. A time step in the Wagner model does not have a strong biological meaning apart from being longer than the life time of transcription factors (the regulation factors being completely degraded between two consecutive “time steps”). There is no particular reason to stick to any particular number of time steps.

To our embarrassment, we realized that simulations had already been optimized prior to the previous submission, and the reported simulations used 16 time steps instead of 24, as erroneously indicated in the manuscript. We now provide the comparison of simulation results with 8, 16, and 24 time steps in a new sup fig S9, and show that the number of steps has a tiny effect on most indicators (especially when > 8). We added this sentence in the text: (which also addresses the two following remarks):
“We confirmed that the number of developmental time steps does not affect the simulation results (except if very low, < 8) (Figure S9 B and C), nor the number of time steps during which network instability was measured (Figure S9D). Selection on the network stability does not have a perceptible effect on the results (Figure S9 E).”

Wagner's model results often in non-equilibrium. According to Wagner this is interpreted as non-viable phenotypes. Here, the authors chose to just put a penalty. I'd like to see some discussion on that. I don't necessarily agree with Wagner, and I tend to believe that he named them as non-viable because he didn't know what to do with them. So some discussion is necessary. Also, some discussion on the penalty. Why to have a penalty on unstable networks?

The reviewer’s remark is relevant, and we have been into this line of thought before. (i) We agree that Wagner’s networks are prone to display a cyclic pattern, as shown in “Pinho, R., Borenstein, E., & Feldman, M. W. (2012). Most networks in Wagner's model are cycling. *PloS one*, *7*(4), e34285”; (ii) Eliminating cycling networks is probably not a neutral choice (Siegal, M. L., & Bergman, A. (2002). Waddington's canalization revisited: developmental stability and evolution. *Proceedings of the National Academy of Sciences*, *99*(16), 10528-10532), as selecting on stability alone is known to affect the properties of the network. (iii) Selecting against cyclic networks is probably unrealistic; periodic fluctuations in gene expressions are probably not too harmful, especially if cells from the same tissue are out of sync, and there are biological processes in which gene network cycling is under positive selection.

Yet, removing selection on network stability and selecting on the average expression over the last n time steps has an embarrassing side effect: some expression levels are much easier to achieve than others. For instance, extreme fluctuations (between expression levels 0 and 1, which can be obtained by simple but very strong negative feedback loops) average out to 0.5; other simple fractions (e.g. 1/n, 2/n, etc) are also easy to obtain through different cycle periods. As the expression level is treated as a continuous variable, this would

generate a non-linear effect of the fitness optimum on the corresponding gene network; for instance, a fine-tuned mixture of positive and negative regulations for phenotypes ranging between e.g. 0.34 and 0.49, then an extreme negative feedback loop for phenotypes between 0.49 and 0.51, and then back to the first topology until 0.66, etc. We considered this pattern unrealistic, and thus decided to avoid it.

We nevertheless ran simulations with s'=0 (no selection on network stability), new figure S9E, and confirmed that the differences with the default setting are barely noticeable for the variables we were interested in, except for the gain/loss connection balance which seems, for some reason, slightly less biased towards gains in absence of selection for stability. We hope that this new figure addresses the reviewer's remark.

Why authors used the four last time steps to evaluate stability? I gess two consecutive equal values would mean that we reached stability.

The reviewer is right: if the whole network is stable for two consecutive time steps, it has reached equilibrium. Note that this is not necessarily true for a single gene, as cyclic expression can follow a pattern A-A-B, for instance (at least two other genes need to be involved, following expression patterns A-B-A and B-A-A). We agree that taking four steps does not completely eliminate the problem, and in this work we do not look at gene-specific instability. We ran simulations with only 2-steps average (figure S9D), and we found no noticeable difference between the 2-steps simulations and the default setting.

It is possible that networks will reach equilibrium at step 25 or 30 or 40 for example. These networks are a bit slower but still they are stable. Why to call them unstable?

Correct. We changed the sentence to include this possibility:
" *A non-null variance characterizes networks that have not reached equilibrium at 16-4=12 time steps, either because slow network dynamics or because the network is unstable (cyclic pattern).*"

Population model:
why they build the genotype by averaging and not use a dominance coefficient for each gene of the network. Thus, they could also evaluate the effect of dominance. Isn't it easy with their simulation scheme? Also, the last sentence of this paragraph " Even if regulatory effects were additive, ..., level", needs further explanation and some examples please. Why and how non-linearities arise?

The fact that regulation interactions are additive does not preclude dominance at the gene expression level because of the non-linear (sigmoid) mapping between regulation strength and gene expression. More specifically, in our setting, strong regulation (both up- and down-regulation) is dominant. For instance, consider two alleles at gene B differing by their sensitivity to transcription factor A, regulation strengths being w1=0.2 and w2=2 for alleles 1 and 2 respectively. The expression of B will be 0.466 in the w1w1 genotype, 0.999 in the w2w2 genotype, and 0.995 in the heterozygote w1w2: the strong activator is almost completely dominant. The same pattern would be found for down-regulation, with a dominance of the strong inhibitor. When regulation is weaker (and when alleles are closer to each other), the system is closer to additivity: w2=0.5 (instead of 2) leads to w2w2=0.850, /and w1w2 = 0.690 (vs mid-homozygote 0.658). Note that the Gaussian fitness function generates dominance for fitness in an unrelated way.

Even if it is theoretically possible to add dominance or interactions at the regulation scale, this would increase the complexity of the model, and require supplementary arbitrary parameters (amount of dominance, etc), while the benefits in terms of realism are uncertain.

To clarify, this paragraph has been expanded. The manuscript now reads:
“In the ‘Wagner’ model, regulatory effects are additive; the regulatory effects of both alleles average out, and the effects of transcription factors add up. Yet, even if regulatory effects are additive, the mapping between the strength of regulation and gene expression is non-linear (sigmoid). As a consequence, the model accounts for both dominance and epistasis at the gene expression level, strong regulators (activators or inhibitors) being dominant over weak regulators. For instance, the gene expression in a loss-of-function heterozygote will be closer to the functional homozygote than to the mutant homozygote.“

For the mutation process: I'd like to see a bit more detailed explanation perhaps with a small example (all this can go to supplement, but it can help the reader who is not familiar with this mutational process).

The manuscript mentioned that “A mutation consists in replacing a random element of the W matrix by a new value drawn in a Gaussian distribution centered on the former value”, which is admittedly dense information-wise. As we consider only cis-regulation in our network model, a mutation is a change in the strength of a single cis-regulation in the network. The following explanation was added:
" *In this model, mutations affect gene cis-regulatory regions only (i.e. protein sequences do not evolve); mutations occuring in the promoter of a gene affects primarily its own expression, but the rest of the network may also be affected when this gene regulates other ‘downstream’ genes.*"

We also now include a new main figure (Figure 1) that provides an example of a mutant network, we hope it helps clarifying.

Also, some discussion about how this is connected to the "nucleotide-based" mutational process. (perhaps through the infinitesimal model? not sure).

It sounds difficult to connect the Wagner model to precise nucleotidic events. In the Wagner model, alleles are described by their effect (on regulation), as in quantitative genetics models where alleles are characterized by their effect on a quantitative trait. As a consequence, two independent nucleotidic substitutions with the same effect would be considered as the same allele. Note that any other mutational event (including indels, or even a complete transposable element insertion in the promoter -- or a heritable methylation) would behave the same way.

The lack of direct connection between DNA sequence and regulation change is not problematic in our approach, except perhaps when estimating the “neutral” molecular variance, which we aim at comparing with neutral nucleotide diversity (note that we stated “analogous measure of nucleotide genetic diversity” in the manuscript). In our model, a regulatory site is described by a single number (the strength of the regulation), and the only diversity we can measure is how this strength varies in the population. With real genomic data, diversity is measured over a DNA segment, with several SNPs; both measures are thus essentially different. Yet, every single SNP originates as a unique mutation, which might have a quantitative effect on the binding affinity of the transcription factor. Adding up all these SNP effects would therefore give a regulation “score” that is very similar to the way we quantify regulation. We therefore admit that our neutral diversity index is imperfect (in the sense that it cannot be directly compared to empirical nucleotide diversity), but it catches the same biological feature (how regulatory sites vary in the population).

We clarified in the text: “Hence, molecular variance is an analogous measure of neutral nucleotide genetic diversity of the genes of the network in the sense that it captures the same biological feature, that is how the effect of mutations (SNPs) on regulatory sites translates into variation of regulation scores among individuals of the population level.”

Domestication scenario
about the burn-in: I think they would mention that with the burn-in they bring the population to an equilibrium. I also think that 12K generation are too few. For a popsize of 20,000 I would guess at least 20,000 generations. So in my opinion 12K is just too few. Alternatively, could they demonstrate that an equilibrium has been reached?

The reviewer is right: the burn-in is a bit short, and populations have not reached an equilibrium for all indicators (especially for the variance). In theory, this is not strictly necessary, as it is doubtful that wild populations were into the very same environmental conditions for thousands of generations. Still, we ensured that a short burn-in is unlikely to affect our conclusions by running long burn-ins (24000 generations) for the three major simulation runs (default, no bottleneck, and no selection switch). We did not run long burn-in for all simulations because of computational burden (long burn-ins double up the simulation duration). When looking carefully, it is possible to see that the short burn-in simulations are not at equilibrium, especially for variance estimates (e.g. comparing fig 5A with fig 5C, maize vs pearl millet, as maize has a long burn-in and pearl millet a short burn-in).The following sentence was added in the manuscript:

“Due to computational constraints, we also had to limit the number of generations prior to domestication (Ta) for some simulations; as a consequence, “wild” populations were not necessarily at mutation-selection-drift equilibrium. However, the effect remains limited compared to the strong effects due to domestication (e.g. Figure 5A (Ta=24000) and Figure 5C (Ta=12000), or Figure S4A vs S4B prior to domestication).”

It was confusing for me why they have put explicitly selection on the genes. My guess would be that selection operates on the network itself (perhaps some combination of the genes that controls for some phenotypic trait) and then selection on the genes is just a consequence of this. I would like some further discussion on that (no need to change their simulation model).

We added a discussion point about this. The reviewer is right, direct selection on genes is not completely natural, as there are probably multiple levels of phenotypes between genes and traits that are targets of selection. Selecting on genes is, however, the usual way to proceed with Wagner model simulations (because gene expression is the only “phenotype” available). In some settings, a transition matrix can be used to translate n gene expressions into m traits that are under selection. Yet, since this translation is linear, selection on genes will still be Gaussian (although there will be correlated selection on gene combinations). This point is now mentioned in the discussion:

“ For simplicity, we considered stabilizing selection directly on the gene expression level -- a common setting in similar studies (e.g. Siegal & Bergman, 2002). This remains an oversimplification, as the relationship between gene expression, physiological characters, life history traits and fitness can be very complex. For instance, Draghi & Whitlock (2012) mapped genes into traits via a transition matrix, stabilizing selection being applied at the phenotypic level, translating into indirect selection on gene expressions. Yet, if the relationship between gene expression and selected phenotypes is monotonous, applying a multivariate bell-shaped fitness function on gene expression probably remains an acceptable approximation, assuming that the details (e.g. asymmetry) of the fitness function does not affect deeply the evolution of gene networks.”

Please provide some details on how the interaction between genes was implemented. Is it in the W matrix?

We clarified the text, and added a new figure 1 that hopefully illustrates how the Wagner model implements regulatory interactions.

Code is available, and this is great!
HOwever: please add a readme file and examples that a researcher who aims at using it to know what to do.

We are not sure to know how to interpret the reviewer's suggestion, as a README file is already provided at the root directory of the project hosted on Github (https://github.com/lerouzic/domestication/blob/master/README.md) .
Note that the repository provides two parameter files called "param0-test.txt" and "extparam0-test.txt", that can be used to run short test simulations (the corresponding command is commented in the run/make_sim.R script).

Figures and take-home messages are well presented. Perhaps in figure 4b increase the font size of the X-axis and right Y-axis.

Figure font sizes have been increased close to the limit (beyond which symbols will overlap). Note that the figures were re-organized, so that figure numbers do not match the previous version.

Reviewer #2 (Comments for the Authors (Required)):

In this manuscript the authors aim to provide predictable expectations for the change in expression plasticity during domestication. They simulate 24 genes in a regulatory network and evaluate three different scenarios: domestication (selection change + bottleneck), only selection change, only a bottleneck. Despite the low number of genes and interactions (which is understandable given computational limitations) the work is timely and important. The finding that the population bottleneck has relatively little importance (compared to the selection switch) seems very surprising. The authors point out the importance of such simulation work for empirical observations. It would be great if this connections could be strengthened through adding an empirical comparison.

We thank the reviewer for useful comments and suggestions. We acknowledge the importance of comparing theoretical results to empirical data; theory can drive data analysis and data can help improve the theory. As detailed in our response to the editor, we also had to address theoretical issues and explore the parameters space more thoroughly, and we were concerned about the size and the consistency of the manuscript. Our preference thus went into putting effort in providing a stronger piece of theory and comparisons across species using simulations, leaving the empirical work for subsequent study.

Major comments:

- I am worried that the selection switch introduced a very strong bottleneck which led to the negligible effect of the simulated bottleneck. If relative fitness is what defines the mating probability, a strong selection shift leads to disproportional fitness advantage of the distribution tail. In this scenario every individual is very unfit, but in a Gaussian fitness distribution a few individuals are magnitudes fitter than the rest and are the only ones contributing to the next generation. This might also explain why standing variation had so little influence on the outcome. Have you looked into how many individuals contribute to the next generation after the shift?

We agree with the reviewer: strong selection increases the variance in the number of offspring contributed to the next generation, and induces a bottleneck in the effective population size. We apologize for not having highlighted this in the text of the previous version of the manuscript; estimates of effective population sizes are now presented in figures 5, S4, S5 and S9. As predicted by the reviewer, Ne was indeed substantially smaller than N immediately after the selection switch, but this effect was transient (as the population adapts quickly to the domesticated environment) and did not last for more than a few hundred generations. We computed the average Ne (harmonic mean) during the whole bottleneck and indicated it in the figures.

It thus turns out that, at least in our "default" scenario, the selection-induced bottleneck is not strong enough to alter our conclusions, as the total strength of the bottleneck is not altered by more than 10% (average Ne=3247 for N=3430). We now discuss this effect in the revision, as it may be overwhelming for some specific simulations (such as the "strong selection" scenario, fig S4C, for which Ne drops to 100).

We added the following sentence in the discussion:
" *For simplicity, we parameterized the bottlenecks by setting the census size (N) to the effective population sizes (Ne) documented in the literature. Because of selection, Ne < N, which made bottlenecks slightly stronger than expected. Yet, the difference was modest (< 10%, Figure 5), and was unlikely to affect the results.*"

- The figures only show mean values. It would be very helpful to show a measure of the variance of these values (e.g. confidence intervals)

We updated the supplementary figures to display the dispersion of the indicators (10% and 90% quantiles) across simulation replicates. Most figures in the main text were too crowded to overlay dispersion, but supplementary figures displayed individual scenario and could be modified without hampering their readability.

- A comic figure displaying the model and the simulation strategy would be very useful to understand the system

We added a new figure to summarize our approach (Figure S3). New figure 1 also illustrates the gene network model.

- I think it would be useful to be defined what is referred to as cis-regulation here.

Our intention was to use a standard terminology, and we apologize for not realizing that "cis-regulation" can be understood in different ways. In practice, the only assumption made by the model is that there is no recombination between regulatory sites of a given gene. We deleted several occurences of "cis" when not necessary (i.e. "cis-regulation" was changed to "regulation"), and we clarified our definitions:

The introduction has been updated to include the following sentence at the first occurence of "cis":"*reduced variation in expression was observed at domestication candidate genes, indicating that selection primarily acts on cis-acting regulatory variants (Hufford et al. 2012): most evolutionary-relevant mutations affecting the evolution of gene expressions are located in (or in the close vicinity) of the domestication genes*". The methods now read: "*There was no recombination between regulatory sites at a given locus (the model assumes cis-regulation only)"*

- In the introduction various examples of network changes during maize domestication are given. It would be interesting to see if the predictions derived from the model hold true in real data. The publicly available data from Lemmmon et al 2014 should provide expression networks for key domestication genes in teosinte and maize, which can be compared to the simulations.

As justified in our previous comments, we decided to perform additional simulations and test scenarios for three other crops. These modifications have substantially reinforced our theoretical work. We however fully agree that an extensive and detailed analysis of empirical data would be an interesting idea, but we see it as a perspective, beyond the scope of the present paper.

Minor comments:

Introduction:
- page 4 first paragraph: Not clear to me why this paragraph is needed for the study. Could probably be cut out

We agree and shorten the first paragraph of the introduction accordingly.

Methods:
- Shouldn't the diagonal of W represent the cis-regulation? or is the idea that only one dimension of the matrix are regulators and the other dimension are regulated genes? In that case the n x n annotation feels misleading. should it then be n transcription factors and m regulated genes, where n= m? If so, can this then still be referred to as cis? To me, cis would mean a change in n altering the expression of n. I think this section would strongly benefit from clarification, as it is essential for the rest of the article.

We are sorry for the misunderstanding. We updated the text to be more specific about the Wagner model internals. In particular, we added a new Figure 1, which may hopefully help readers to grasp the meaning of the regulation matrix. In order to confirm the potential for n x n regulatory interactions, we added the following sentence in the methods: "*All genes have the potential to regulate other genes of the network (although such feedback is not mandatory).*"

We agree with the reviewer's definition of *cis*, but we are not sure to uderstand where the misunderstanding lies. All mutations occuring in the promoter of gene n alters the expression of n, which is why we refer to "cis-regulation". The promoter of n contains several potential sites of fixation for transcription factors, so mutations can affect the sensitivity of gene n to transcription factors a, b, c, etc. (including n itself, which we refer to as "self-regulation"). This promoter region can be understood as a line of the W matrix (as suggested by e.g. Siegal & Bergman 2002, in the following figure https://www.pnas.org/content/pnas/99/16/10528/F1.medium.gif). We would be glad to know if the reviewer finds the new figure insightful enough, or if we need to provide further details in the text.

Population model:
Do I understand correctly that one row of W is drawn at random from each parent and the mean of each row was calculated to form the offspring?

The simulation software stores two alleles (= two "rows") for each individual at each locus, and one allele is transmitted randomly to the offspring. These two alleles are averaged out when computing the expression phenotype. The text has been clarified in the revised version:

"*The genotype of an individual was defined by both inherited gametes; the W matrix from which the expression phenotype was calculated was obtained by averaging out maternal and paternal haplotypes*."

Results:
Variances across simulations should be addressed, rather than only reporting means.

We now report (when figures were not too crowded) both the mean and the 10% - 90% quantiles (quantiles being more informative than the standard deviations when the distribution is not Gaussian).

Discussion:
- I am wondering if the decreased plasticity was selected on or if it is rather the result of domestication. Maybe this aspect could be addressed in the discussion.

This was indeed not completely clear in the previous version, with the sentence "... *phenotypic canalization, simulated as the evolution of selection pressure towards decreased plasticity paralleling environmental stability of phenotypes*."

Among the 6 plastic genes in the Teosinte scenario, 2 remained plastic, 2 became stable (and thus, were selected against plasticity), and 2 became unselected (and thus, could remain plastic or lose their plasticity without any consequences for fitness). The decrease in the number of selected plastic genes was thus a part of the simulation design. This fact is now mentioned in the discussion:
"The network was less plastic after domestication, which was a consequence of a modelling choice (domestication was associated with a drop in the number of genes expected to respond to the environmental cue)"

Figures:
- Fig 2: The switch of the color meaning makes it very hard to interpret figures. Would be better to use a different set of colors for a and b

The reviewer is right, we used the same color code (stable in blue, plastic in red, unselected in black), but in panel A colors stand for selection before domestication, and in panel B for selection after domestication. This was arguably confusing. The color code has been simplified (and gene-type specific trends have been removed for clarity). Note: former figure 2 is now figure 3.

- Fig. 4B: maybe colored lines might help to separate the groups of genes.

Done. This clearly improves the figure readability for the "Before domestication" part; the pattern for "Present" is less clear because the selection pattern on genes was shuffled at the onset of domestication. (Note: Figure 5B now).

Supplement

Fig. S1: for which scenario is Ne shown here? Shouldn't it be calculated for all three scenarios individually?

This was for the "default" scenario only. Ne for all scenarios are now provided (in Figures 5, S4, S5, and S9).
Fig S3. The values in "stable" don't add up.

Thanks for noticing it, it was a typo in the table. Row 'Stable' x column 'Non selected' should be 2 instead of 4. Now fixed in Fig S2.