-------------------------------------------------

Associate editor comments 06.08.20

-------------------------------------------------

The manuscript has improved and there are only few minor comments from Reviewer 2. Although these comments are about the breeding example (and not about the program), I still think that the authors should check if they can adjust the manuscript accordingly.

We thank the editor for another round of opportunity to improve our manuscript. We have addressed some concerns of the second reviewer in relation to the breeding example. We would like to emphasise here that this is only an example and not a full-fledged pig breeding simulation with a large number of scenarios. We simulated scenarios that we found relevant based on real life cases. We provide the AlphaPart and complete simulation code as open source for further studies.

-------------------

Reviewer #1

-------------------

The authors made substantial changes that improved the paper.

We thank the reviewer for another read of the manuscript!

-------------------

Reviewer #2

-------------------

The manuscript has improved, and is close to being acceptable. However, I am a little confused by the simulation. Given that selection of multiplier males seems to be an important part of the story (L222), I was surprised that the multiplier females (and in particular multiplier males for MaleFlow 20) do not contribute more to the genetic trend, given the statement about Multiplier males being very good. I presume this is because the Nucleus is closed, and this I presume is common in pig and poultry situations because of bio-security limitations.

Yes, nuclei are generally closed in pig and poultry situations due to bio-security. Hence the flow of genetic gain is nucleus  multiplier. Once we noticed larger breeding values in some multipliers for some traits (but lower for others) compared to the nucleus, we were confused too!

Prior to the analysis our view was that multiplier animals can not contribute to the multiplier genetic trend, but we have not appreciated the effect of multiplier selection. In the absence of multiplier selection, the expected contribution of multiplier animals to the multiplier genetic trend is zero (the multiplier trend follows the nucleus trend, possibly with a lag). With multiplier selection the expected contribution of multiplier animals is small because most of the genetic gain still comes from nucleus (direct impact of nucleus males) and selection in multiplier uses ½ of parent average that pertains to the nucleus males (indirect impact of nucleus males). The multiplier males (and partly females) are an important part of our story because they cause a non-zero, yet small, contribution. The partitioning method gives us a way to account for all these contributions jointly.

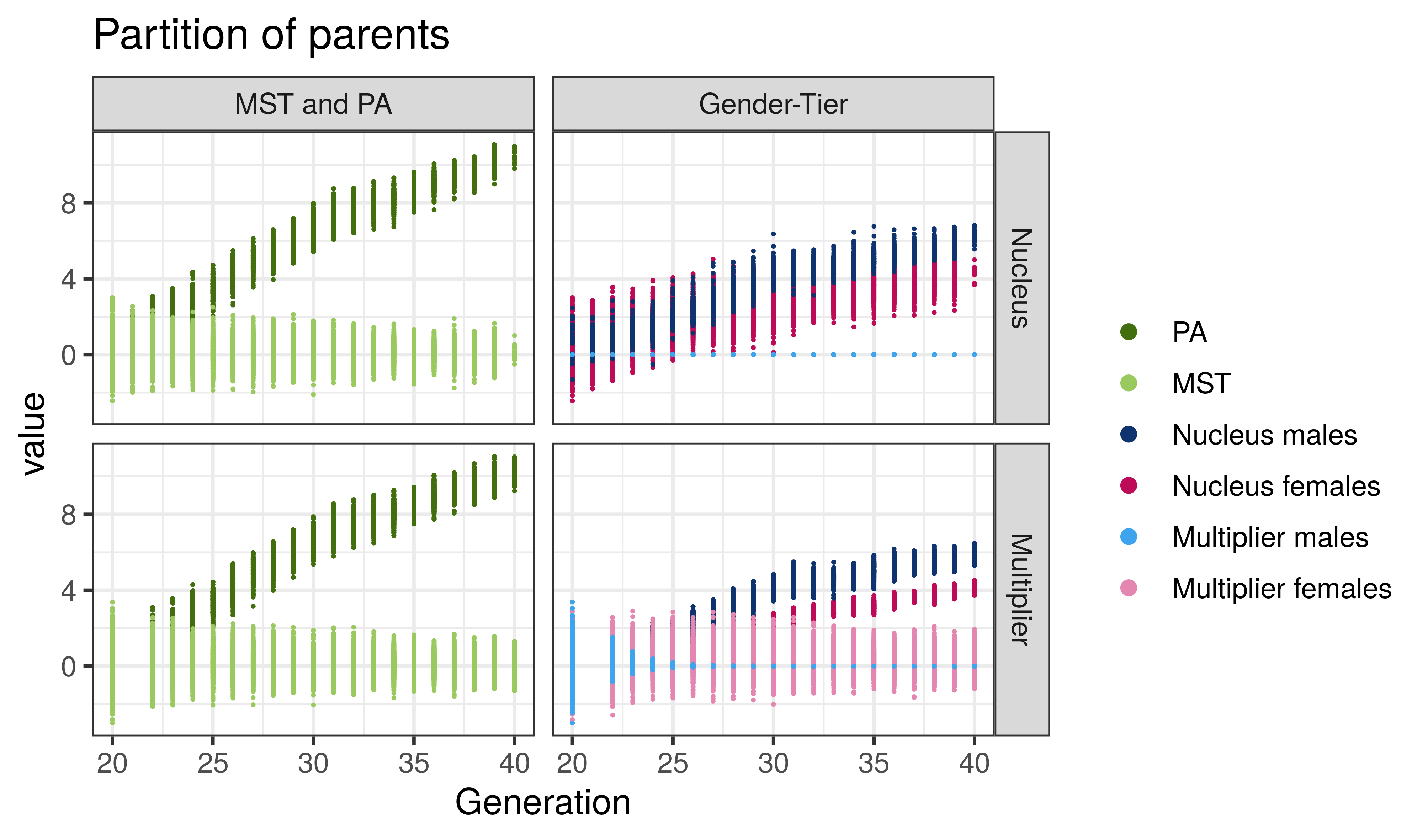
In the results and discussion, there is no mention of genetic lag, which is the main driver of the normal expectation of why multiplier animals are inferior to nucleus animals. In this simplistic example, the genetic lag is very small, and so it is relatively easy for superior recording in the multiplier to offset the lag effect, and allow high merit multiplier females (and males) to be identified, relative to what is in the nucleus. I am not fully convinced that the partitioning method is helpful in explaining this tradeoff between lag versus higher accuracy of evaluation in multiplier candidates. This doubt is compounded by the relatively low contribution of multiplier selections to the genetic trend as shown in Figures 4 and 5.

Progressive pig breeding programmes limit lag between nucleus and multiplier by using nucleus semen in nucleus and multiplier as we have simulated in our study. We did not embark on simulating all possible scenarios in relation to flow and lag in pig breeding as this was not the aim of this study. We merely wanted to communicate an intriguing example. We highlight now in discussion that our results need to be viewed in this context of a limited time-lag.

I would also like to see an explanation of why generational responses are declining over time. And why do the multiplier contributions start off high and fade over time?

We believe this is due to cumulative (recursive) property of the partitioning. We partitioned the genetic values from generations 20 to 40, hence the animals in year 20 are the founders. Their entire breeding values is thus assigned to the tier of origin. As the programme progresses, the multiplier animals have many generations of nucleus ancestors. The majority of their parent average is then assigned to nucleus selection and only their Mendelian sampling term is assigned to multiplier selection. Also, when the accuracy is low, the selection acts on parent average, which further increases the contribution of nucleus selection. Below we show a figure with partition of nucleus and multiplier parents’ breeding values.

We also added a clarification of this phenomena in the paper: “In both scenarios we see a trend of decreasing contributions of multiplier selection over the generations. Since we partitioned the genetic values from generation 20 onward, the entire breeding values of generation 20 multiplier animals is assigned to multiplier selection. As the selection progresses, they have more generations of nucleus ancestors and an increasing proportion of of their breeding value is assigned to nucleus selection.”



L37 compared to a closed nucleus

We have added this clarification. Thanks!

L64 suggest add sentence "This is to avoid counting the imported sire's contribution to the domestically born animal, when aggregating genetic trend contributions arising from the domestic selection effort."

Thanks! We have added a slightly different version: “The partitioning by origin allows for quantifying impact of import versus domestic selection.” Please note we cite a number of publications that use the partitioning method for this type of analysis in the following paragraph.

L174 supply a much large number of breeding animals (usually females) for commercial purposes.

Thanks! We have added a slightly different version as there are different multiplier business “models”: “a large number of breeding animals for commercial purposes”.

L199 I do not find the names MaleFlow20 and MaleFlow100 very informative or helpful. I presume because only 20% of the males for multiplier breeding come from the nucleus. However, why are there so many males required in MaleFlow20 (125 males versus 25 for MaleFlow100?). It would seem to make more sense to select only an additional 25 from the Nucleus, and so show the impact of multiplier selection even more clearly?

The 20 and 100 numbers indeed denote the percentage of immediate “gene-flow” into multiplier from the nucleus males. We now state this in the text. A larger number of males in MaleFlow20 is a consequence of a larger number of multiplier males (the number of nucleus males stays the same as in MaleFlow100) to resemble the lower intensity of multiplier selection.

We have used these numbers after real-life cases where larger or lower estimated breeding values were observed for some multiplier animals compared to the best nucleus animals. These real-life cases are based on several multipliers with different regimes. We designed the two scenarios to evaluate what is driving these unexpected results. We do not claim anything in particular with these scenarios – we are merely showcasing the application of AlphaPart.

L213 delete "the"

Done, thanks!

L245 replace "had" with "made a"

Done! Thanks!

L283 "a" freely available

Done! Thanks!

L297 delete the first occurrence of "the"

Done! Thanks!

L315 numberS

Done! Thanks!

L373 reword. Poor English

Thanks! We changed the text to: “The plot shows the overall genetic trend and its partition into the contributions of domestic selection and import by generation.”

L384 remover "r: from malerflow

Done! Thanks!

Figure 5. I was surprised to see such a low contribution of the multiplier to the genetic trend

We actually expected low, even zero, contribution of the multiplier to the multiplier genetic trend before we embarked on this study! On average the contribution is not large, but it is non-zero. Further, for individuals there can be quite a bit of variation behind this mean and this causes some multiplier males to outcompete nucleus males for traits measured in the multiplier.

As described in the text we show two things: i) distribution of breeding values for one generation of one replicate in Fig 3 and ii) partitioning of mean breeding values by year summarised over many replicates (to ensure this is not just sampling noise) in Fig 4 (MaleFlow100) and Fig 5 (MaleFlow20). In the previous version of the manuscript we showed distributions for partitions for one generation of one replicate, but that seemed to cause confusion, so we omitted them.

We now emphasise this in the text and provide a supplementary plot of distributions for the partitions for one generation of one replicate (Additional files 1 and 2).