I find this manuscript very poorly done and that, in its current form, it does not meet the quality expected for GSE. A major revision is required that takes all my comments into careful consideration.  
   
I have many comments on the attached annotated .doc file of the manuscript and have provided a summary of my concerns below:  
   
This is a very poorly crafted manuscript. The structure should be much improved, especially the part where the software is described, the writing should be a lot more precise and less repetitive, and the example given is so poor that it severely distracts from the benefit of the proposed software. Earlier reviews suggested to remove that example. I agree that an example could be useful but please use a more sensible one, as otherwise you are missing a great opportunity to showcase your software.

Answer: We thank the reviewers and the editors for the comments. We regret to see that the manuscript still is not in the line with GSE standards but we worked to improve it in line with your comments. We significantly improved the description and demonstration of the software. First, we improved the description of the functions to make it more understandable by readers not profoundly familiar with R. Second, we provided a new more illustrative example of the software use.

It seems sometimes that the authors do not fully grasp the genetic gain and structure of breeding programs (which is of course the essential topic here). Genetic gain is the sum of selection over different selection paths. For one generation, it is the result of selection differentials created in males and females. So if in the example the same males are used in the nucleus and multiplier (very unrealistic, why have a nucleus if not using the best males only there) and the selection intensity in the females differs between N and M (as index selection versus T1 selection) that would simply explain differences in means. However the manuscript explains it in a more complicated way in lines 261-264, and this seems to me incorrect.

Now partitioning can be interesting since the long-term effects of selection paths can differ from the initial contribution, but this is only relevant if the paths compete with each other, in other words, there is truncation selection across paths. This is not the case in the example, as it is not an open nucleus I was not sure whether there was truncation selection across N and M sires in the second scenario, I think not.

Also, although the authors talk about a gene flow effect,  this is not relevant if the same sires are used for both tiers in the same generation and the sires geneflow is the only geneflow between tiers)

The manuscript would benefit so much if a more illustrative example, with more contrasting and illustrative scenarios gave partitioning results that are not totally predicable form the design. If such an example cannot be generated, I suggest indeed to leave the example out and just describe the software. In its current form the example seems naïve

Answer: We thank the reviewer for the comment. However, we do not agree that the stylized breeding example is predictable and non-illustrative. We took this example since it stems from a real life question, since pig breeding organisations observe animals with higher genetic value and higher genetic gain in the multiplier than in the nucleus. The explanation for this is not straightforward, since multiplier has genetic lag. By simulating although stylized breeding example and partitioning the breeding values and genetic trends we were able to show, that the performance of the multiplier is a sensitive interplay of the heritability of the trait, accuracy of the estimated breeding values and the percentage of import. We were able to show that the advantage of the import depends on the accuracy of the multiplier breeding values. If they are accurate and we truly have the power to select the superior animals, then reducing the import from the nucleus benefits the multiplier. Since we do not know the true breeding values in real life (and hence it is difficult to estimated the true accuracy of selection), partitioning the results can provide a powerful tool to select the breeding strategy.

Furthermore, the word stylized is correct, but somewhat unusual, and it does not have to be repeated so many times.  
   
How can T2 almost have the same dG as T1? The h2 is much lower. Simple index theory tells me it should be about 3 times lower with equal weighting and no correlation (and BLUP selection)  
   
I found the implementation part poorly written, although this is the key part, describing the function and usage of the package. It might be understood by the person that developed the package and is intimately familiar with R. However, the general reader will get confused at many places. The R-prints have an ugly format and could be more formally shown as tables  or boxes. E.g at lines 110-115:  
   
I have quite a few other clarifying questions, in lines 123-143 (see annotated .doc file)  
   
In summary, I suggest another major revision of this manuscript for which a much better example should be used and in which the description of the software's functionality is much improved and more palatable for the GSE readers.  
  
  
Indeed Referee 1 also thinks that the description of the R package contains some specific phrases that are only understandable for R programmers, and the authors should rewrite it for a wider audience and not so much tailored for R programmers.