Follow-up to previous comments from Reviewer3:

1. Figure 3 (supplementary Figure 1 of initial submission), could you please clarify if the top thicker blue arrows, one labeled "Random exclude N genes", the other labeld "Exclude downstream target genes",indicate the same processing step? If yes, could you please label them using the same text. If no, could you please detail the difference in the Two-class Bootstrap Simulation section (line 256)?

Those two labels represent two different processes. We have now included an explanation for the difference in the Two-class Bootstrap simulation section (Line XXX). Here we would like to test the impact of an upstream regulator on its downstream targets (N) using two simulation strategies. In the “elimination without replacement” process, we attempted to eliminate the same number (N) of irrelevant genes from the test gene list (i.e. GATA2 significant gene list in Figure 3), and then continued with the following SEM modeling steps. On the other hand, in the “elimination with replacement” process, we first eliminated the regulator’s target genes from the test gene list, and then randomly selected the same number of “irrelevant genes” from the gene pool (indicated by the blue cylinder in Figure 3, All genes except tested genes). The randomly selected irrelevant genes were next put back into the shrunken list to restore to the same number of genes as the initial test gene list (i.e. GATA2 significant gene list in Figure 3) and subsequently went through the following SEM modeling steps.

~~[Thanks for the follow up question. The reason we labeled them differently was that those were two different processes, but they do end up with “N” number of genes eliminated, and we see whether the confusion comes from. The starting point is that we try to test the impact of an upstream regular on its downstream target. In the “elimination without replacement” process, we attempt to eliminate same number (N) of irrelevant genes, the continue with the following SEM modeling steps etc. On the other hand, in the “elimination with replacement” process, we firstly eliminate “actual downstream target genes (N)”, and then randomly select same number of “irrelevant genes” from the pool – indicated by the blue cylinder and put them back into the shrunken list to restore back to the same number of genes as the “GATA significant gene list” followed by the following SEM modeling steps.~~

~~The two simulation strategies have different objectives.~~

~~We have updated the manuscript for the Two-class Bootstrap section line xxx]~~

2. My earlier question about the two bootstrapping methods are not adequately addressed. In my question, I meant to ask if the two methods have different computational costs, and if they have different test power. In what situtations would you prefer one method over the other?

Sorry that we did not fully address this question in our initial response. As explained in responding to your further comment #1these two methods do have different objectives.

Q2.1 if the two methods have different computational costs

Since the “elimination with replacement” involves selecting the “same” number (N) of genes from the pool and then restore the shrunken list back to its original size, it does take a little bit more computational resource during each bootstrap. Such extra computation time will be more prominent if we execute the bootstrap 1000 rounds.

Q2.2 if they have different test power

The two simulations serve for different purposes. The “elimination without replacement” process helps to evaluate whether eliminating the “downstream target” will change the downstream SEM analysis results, so that we can deduce the function/impact of the upstream regulator. The “elimination with replacement” process helps to rule out the possibility of changed SEM analysis results is caused by “artificial effect” due to the shrunken significant gene set. There is no direct comparison on their test power per se.

Q2.3. what would you prefer one method over the other?

We suggest using both so that not only can we identify any candidate upstream regulator with significant impact on its downstream targets, we can also rule out any possibility that such observation is caused by “artificial effect” due to the shrunken significant gene set.

~~[Sorry that we did not fully address this question in our initial response. As explained in responding to your further comment #1these two methods do have different objectives.~~

~~The “elimination without replacement” process was designed to test whether random elimination of the irrelevant “downstream genes” render the same SEM testing result as those from the “real downstream target genes”. From such 1000 bootstrap simulation, we shall have an empirical distribution of “statistics” we are interested (deleted?), and a significance result will be obtained if our actual measure falls in any of the significant tails.~~

~~For the “elimination with replacement” process, since we stored the “original gene list” i.e. “GATA significant gene list” by supplying the same number of genes that we eliminated from the initial set. Thus, this will further provide us the information on whether what we observe is from “eliminating N number of genes” or from “eliminating the target genes”. This procedure does take a little bit more computational resource, but it helps us to rule out the unnecessary question and it is necessary.~~

Additional minor comments:

1. Line 36, move "two-sided t-statistic" to line 30, where "T-score" is first mentioned.

~~Revised accordingly.~~ It is corrected

2. Line 57, suggest to remove "unbias[ed]ly", unless the authors would clarify what kind of bias is of concern and how the gene signature constructed this way are un biased.

~~Revised accordingly.~~ It is corrected

3. Line 227, typo "uwe"

~~Revised accordingly.~~ It is corrected

4. Figure 3, "Random[ly] exclude N genes", "Random[ly] draw (N) genes"

It is corrected