Dear Gemma,

First, we are delighted to be part of this special issue and are happy to be labeled as an opinion article. This article is meant to be a review but is also our perspective as practitioners of network inference. Second, the reviewers made several excellent suggestions which we address below and in the manuscript. Third, we provide editable versions of both text and figures as well as a pdf in the other attachments.

Warm Regards,

Dennis and Gloria

Reviewer 1:

1. The topic of the manuscript needs to be better defined. The specific meaning of "gene regulatory network" used in the manuscript needs to be given. For example, is it a network of genes, connected by transcriptional regulatory interactions? Or is it a network of gene products that includes interactions among transcription factors and genes?

*As stated by the reviewer we indeed clarified this point. GRNs are now defined page x, line x and refer to genes connected by transcriptional regulatory interactions.*

second part. The "causal relationships in gene regulatory networks" will similarly need to be defined. Is it restricted to putative transcriptional regulatory interactions? This is important because it affects what work to review. For example, there is recent work that uses gene expression information to infers causal relationships between hormones (e.g. ABA) and genes which is not mentioned in the review.

*We agree that this sentence could have been misleading. It is now clarified. We restricted the notion of GRNs and the related causal relationships to transcriptional networks (TF🡪Target relationships) excluding Metabolism interaction with transcriptional network.*

2. I understand the logic of showing first successful case studies, but the methods used will need to be intuitively explained in order for the reader to be able to interpret the success of the method. For example, it's not enough to say "correlation networks have been used extensively to study GRNs in plants."

*We now evoke the technique when it is first used in the “successful case studies” part. These methods are then classified in the second paragraph. We agree that this may improve the readability of the paper.*

3. Some of the statements in the Strong Prior approaches section need to be revised/clarified. These models do not only determine the strength of the edges, they determine or predict the behavior of the network in time. The models are used to explore the system's behavior (steady state, oscillation) inuntested conditions and to determine the elements that are key to this behavior. One clear way of separating the two types of approaches was used in the review article "Network Inference, Analysis, and Modeling in Systems

Biology" (Plant Cell 2007). The examples given as "Strong Prior approaches" are dynamic models of gene regulatory networks. They start from the network and a description of how each node's state depends on the state of its regulators.

Their output is the state (behavior) of the system in time. The examples given in "Weak Prior approaches" are examples of network inference. They start from gene expression timecourses (state information) and their output is a putative gene regulatory network.

*Thank you for bringing this excellent reference to our attention. We have incorporated it into our discussion.*

#

4. The sentence on page 4, "This gene regulatory

network.... as an output." is too long and unclear. What exactly is "buffering

capacity"? What is the "in planta behavior" in question?

*We agree with the reviewer remark that this part was not very clear as stated. We re-worded/clarified. As stated in the manuscript the “buffering capacity” of the auxin network is the fact that auxin-induced genes are stably activated even if auxin treatments/inputs/local concentrations may be very variables.*

5. On top of page5, steady states are examples of attractors, not basins of attraction. The statement that plants arrive at these steady states independently of the initial gene expression values implies that there is a single steady state. If there are several possible steady states (as stated), each has a basin of attraction. The states inside each basin are indeed equivalent, but it does matter if the initial state is in one basin or the other.

*We have adopted the reviewer terminology in the manuscript.*

6. On page7, more detail is needed on what exactly "modeled with transcriptional regulators" means. An inspection of reference 31 reveals that in the case of

SPL9 over-expression multiple target genes had a different expression timecourse than in the control. But this does not necessarily mean that the over-expression "modified the regulation of the predicted target genes", not in the sense that these genes acquired new regulators or lost regulators.

Certainly, the strength of the regulation can change because it depends on the state of the regulators, and an overexpression of a regulator will directly or indirectly change the state of multiple nodes. But "modified the regulation" is misinterpretable, as is the conclusion that "network relationships adapt to genetic perturbations"

*We agree with this remark. This also has been re-worded.*

7. The explanation of what "causal link" means needs to be given earlier than page 7. Several aspects of this explanation need to be clarified or strengthened. Since the explanation invokes "object A" and "object B", it is incorrect to say "removing some B". One cannot remove some of a single object. Rather, a population or ensemble of objects A, B needs to be used. When explaining linear influence, it is important to specify that each element contributes additively (that is, the elements' contributions are added).

*Thank you. We have clarified our discussion of causal link. We have chosen to leave the concept of causality informal earlier in the paper, because we believe the intuitive notion of causality are sufficient for the explanations that precede the causal link discussion.*

8. Correlation does not usually imply causation; that is, the source and target of a pairwise correlation cannot be determined. If Graphical

Gaussian models or partial correlations can in fact infer causality, it needs to be explained in what way. This explanation should be jargon-free. The same with mutual information. Since the topic of the review is learning causality, methods that do not yield causality should not be included.

*This has been clarified. We continue to mention correlation networks because, together with other data types, they can contribute hypotheses to causality.*

Minor comments: In the abstract, "such" at the end of the fifth line and

"cycle" at the end of the ninth line should be deleted. 2. On page 9,

"change in Gene A concentration" should be "rate of change in Gene A

concentration". 3. On page 12, "biological phenomenon" should be

"biological phenomena".

*Thank you. We have corrected these.*

Reviewer 2 Comments:

1. The authors offer a high-level view of gene network approaches in plants. Overall the review is well written. They classify the methods in 'strong prior' and 'weak prior' approaches. This may be a useful distinction in particular in relationship to machine learning, Bayesian inference, information content and the iterative nature of modelling but often the approaches have different goals and this seems to get lost in the current presentation. I wonder if a problem and data focussed classification in terms of which steps to perform may be more accessible.

*A paper that focusses on methods is a good idea. In previous work, we followed the data type approach and related the discussion to machine learning. So we have added a citation to a book (Network Inference in Molecular Biology -- a hands-on framework Jesse Lingeman and Dennis Shasha,*

*Springer Verlag, 2012, 109 pages, ISBN 978-1461431121). However that approach sounded to be too much of a computer-scientist oriented way of thinking. Since the Genome Biology readers tend to be biologists , we decided to highlight the biological insights in this review. The value to the biologist will be to understand that certain techniques work in certain settings.*

2. Some examples are listed but the level of detail in terms of the underlying biological question, the chosen approach, or the conclusions is not very informative. Further information on the underlying principles of each method would be helpful as well as their advantages, disadvantages, etc.

Also a discussion of the type of data (in particular how this relates to plant genomics) would add to this review and make the choice of methods easier to follow.

*Thank you. We have elaborated.*

3. I was a bit puzzled by what the authors mean with 'information

richness'. How does this relate to how much of the variance each method can

capture and the number of parameters? Why are differential equations less

information rich than Boolean networks in Table 1?

*We agreed and we have elaborated with an example: “High information richness would for example allow the inference of the dynamic behavior of a network, whereas low information richness would give some idea of the connectivity of a network.”*

4. I think more emphasis needs to be given to goal of modelling in the examples (which is of course problem specific). Also, a clearer discussion and distinction between trying to find correlations in order to identify connections in a graph vs dynamic simulation of an existing network.

*Thank you. We have reflected this in various places.*

5. This is a rather high-level review that only touches lightly on any advances in field and at such a level that the reader learns little more than that the work was undertaken. I missed any kind of new insight, critical review, or a novel perspective. The only thing that appears to be new is the distinction based on prior knowledge, which in connection to a Bayesian view may have been interesting but in isolation I didn't find helpful.

*We agree with this. We decided to provide a high-level review in order to help working biologists to gain a sense of the effectiveness of different techniques rather than exhaustively review each of them. Operationally, this will help those biologists point computational colleagues in the right directions.*