Differential Network Analysis of Anti-sense Regulation

Marc Legeay[1,2], Béatrice Duval[1], and Jean-Pierre Renou[2]

[1] LERIA - UNIV Angers - Université Bretagne Loire
2 bd Lavoisier 49045 Angers FRANCE
marc.legeay@univ-angers.fr,beatrice.duval@univ-angers.fr,
[2] IRHS, INRA, AGROCAMPUS-Ouest, Université d'Angers, SFR 4207 QUASAV
42 rue Georges Morel 49071 Beaucouzé cedex, FRANCE
jean-pierre.renou@inra.fr

Abstract. A challenging task in systems biology is to decipher cell regulation mechanisms. By comparing networks observed in two different situations, the differential network analysis approach enables to highlight interaction differences that reveal specific cellular responses. The aim of our work is to study the role of natural anti-sense transcription on cellular regulation mechanisms. Our proposal is to build and compare networks obtained from two different sets of actors: the "usual" sense actors on one hand and the sense and anti-sense actors on the other hand. Our study only considers the most significant interactions, called an Extended Core Network; therefore our differential analysis identifies important interactions that are impacted by anti-sense transcription. This paper first introduces our inference method of an Extended Core Network; this method is inspired by C3NET, but whereas C3NET only computes one interaction per gene, we propose to consider the most significant interactions for each gene. Secondly, we define the differential network analysis of two extended core networks inferred with and without anti-sense actors. On a local view, this analysis relies on *change motifs* that describe which genes have their most important interactions modified when the anti-sense transcripts are considered; they are called AS-impacted genes. Then from a more global view, we consider how the relationships between these AS-impacted genes are rewired in the network with anti-sense actors. Our analysis is performed by computing Steiner trees that represent minimal subnetworks connecting the AS-impacted genes. We show that the visualisation of these results help the biologists to identify interesting parts of the networks.

1 Introduction

Gene regulation is a key issue in bioinformatics. As micro-array produce large-scale expression datasets, gene network inference is a useful approach to study gene interactions, and a lot of methods have been proposed in the literature for this reverse engineering task [1–4]. Differential network analysis [5–7] proposes to study the cellular response to different situations. In medicine, differential

network analysis is used to compare healthy tissues and diseased tissues in order to reveal network rewiring induced by the disease [8]. In these approaches, the comparison is made by using the same set of genes to construct the gene networks.

Our work has the particularity to study gene networks with sense and anti-sense transcripts. Anti-sense RNAs are endogenous RNA molecules whose partial or entire sequences exhibit complementarity to other transcripts. Their different functional roles are not completely known but several studies suggest that they play an important role in stress response mechanisms [9]. For instance, a significant effect is the post-transcriptional gene silencing: the self-regulatory circuit where the anti-sense transcript hybridizes with the sense transcript to form a double strand RNA (dsRNA) that is degraded in small interfering RNAs (siRNA). Previous studies on *Arabidopsis Thaliana* showed that sense and anti-sense transcripts for a defense gene (RPP5) form dsRNA and generate siRNA which presumably contributes to the sense transcript degradation in the absence of pathogen infection [10].

Our work focuses on apple because a recent study [11] suggests that a large majority of protein coding genes of this organism are actually concerned by anti-sense transcription. The authors have combined microarray analysis with a dedicated chip and high-throughput sequencing of small RNAs to study anti-sense transcription in eight different organs (seed, flower, fruit, ...) of apple (*Malus* × *domestica*). Their atlas of expression shows several interesting points. Firstly, the percentage of anti-sense expression is higher than that reported in other studies, since they identify anti-sense transcription for 65% of the sense transcripts expressed in at least one organ, while it is about 30% in previous *Arabidopsis Thaliana* studies. Secondly, the anti-sense transcript expression is correlated with the presence of short interfering RNAs. Thirdly, anti-sense expression levels vary depending on both organs and Gene Ontology (GO) categories. It is higher for genes belonging to the "defense" GO category and on fruits and seeds.

The work described in this paper proposes a large-scale analysis of apple transcriptomic data, with measures of anti-sense transcripts in the context of fruit ripening. The fruit ripening is a stress-related condition involving "defense" genes. To highlight the impact of anti-sense transcription, we propose to compare context-specific gene networks that involve two kinds of actors, on one hand the sense transcripts that are usually used in gene networks and on the other hand the sense and anti-sense transcripts. However gene network inference methods generally find many false positive interactions, and some authors have proposed to study the core part of a gene network [2], by only computing for each gene the most significant interaction with another gene. This constraint seems too restrictive and we developed a method where we extend the core networkby considering for each gene a small number of significant interactions. We call our gene network inference method the *Extended Core Network*. Then we use this inference method in order to discover which interactions of the core network are modified when we integrate the anti-sense transcripts. We define the notion of *AS-impacted genes* to describe sense genes whose interactions are

highly impacted by the integration of anti-sense transcripts in the data. It is also interesting to study the relationships between these AS-impacted genes and we explain how the rewiring around AS-impacted genes can be computed by Steiner trees.

In section 2, we present the Extended Core Network Inference method and we evaluate it on artificial datasets. In section 3, we present our workflow to compare two core gene networks built on different sets of actors and we define the change motifs that highlight significant differences between the interactions. For each step of our workflow, we provide results obtained for the apple data.

2 Extended Core Network inference method

2.1 Motivations

Several methods have been proposed to infer gene networks from transcriptomic data. Reviews of these reverse engineering methods can be found in [1,3,4]. Some inference methods reconstruct pairwise gene interaction networks by measuring with a statistical criterion whether two genes are co-expressed or co-regulated. This statistical measure can be Spearman or Pearson correlation [12], or mutual information [2, 13], that enables the detection of non-linear relationships. These methods need a step of thresholding to decide which values of the statistical measure are significant. One major drawback of these methods is that many of the predicted interactions are false positives. We can differentiate two types of false positive interactions: an interaction that does not biologically exist, and an indirect interaction. If two genes g_2 and g_3 are regulated by g_1, then mutal information (as well as correlation) between g_2 and g_3 is high and an indirect interaction is put in the inferred network. Indirect interactions lead in large gene networks difficult to interpret by biologists and they must be pruned from the output networks [13, 14]. To avoid false positives, the Conservative Causal Core Network (C3NET) [2] proposes to compute the core of a gene network, by selecting for each gene a unique interaction defined by the maximal mutual information value. Our aim is to compare two inferred networks to identify significant changes in the interactions when we take into account the anti-sense actors. Therefore only considering the maximal interaction for each gene is too restrictive, since several mutual information values may be very close to the maximum, and a strict comparison of the maximal values in two situations is not relevant to compare two networks. So we propose a gene network inference method, based on C3NET and named *Extended Core Network (ECN)*, where for each gene the most significant interactions are put in the inferred network.

2.2 Preprocessing

The microarray data we use are intensity values of each gene in each sample. We first normalize the data using quantile normalization [15, 16]. We also copula-transform the data [17, 18] before estimating the mutual information for each pair of genes.

Extended Core Network uses the mutual information in order to estimate the connections between genes. Given the expression vectors I and J of genes i and j respectively, we estimate the mutual information $M[i,j]$ with the same estimator used in C3NET:

$$M[i,j] = \frac{1}{2}\log(\frac{\sigma_I^2\sigma_J^2}{|C|})$$

where σ_I^2 and σ_J^2 are the variance of I and J respectively, and $|C|$ is the determinant of the covariance matrix.

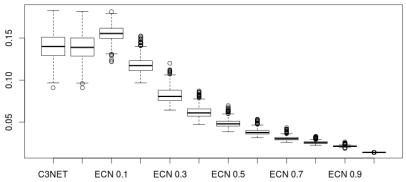
As C3NET or ARACNE [13], we test the statistical significance of pairwise mutual information by re-sampling methods. All non-significant values are set to 0 before applying the inference algorithm.

2.3 Algorithm

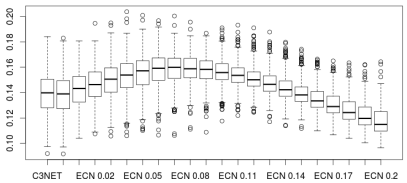
The Extended Core Network (ECN) algorithm computes an adjacency matrix network for a set of genes G. We first initialize the matrix by considering that there is no connection at all. Then, for each gene, we determine its neighbours. The neighbours of a gene g are the ones with the best mutual information values with g. We use an accepting rate r in order to identify the threshold value that determine if a gene g' has one of the best mutual information with g. The threshold of g is fixed by the maximal mutual information of g with other genes, and the accepting rate r. The accepting rate must be between 0 and 1; 0 means that only the best neighbour will be selected and 1 means that all significant interactions will be selected. When the accepting rate is 0, it is almost the same as in C3NET: if two interactions share the best mutual information, both of them will be selected in ECN whereas only one will be selected in C3NET. When the accepting rate is 1, the interactions of the output network are all the significant mutual information values. Non-significant values are set to 0 in the preprocessing step, so that only non-null values are selected as interactions in the output network. Mutual information is a symmetrical measure, that is why the final step of C3NET is to transform the asymmetrical adjacency matrix into a symmetrical one: the result is thus an undirected graph. Because we want to compare two networks and see which interactions are modified, our algorithm provides a directed network so that we can identify which significant changes occur in the connexions of a gene when we integrate anti-sense actors in the algorithm.

2.4 Evaluation on artificial datasets

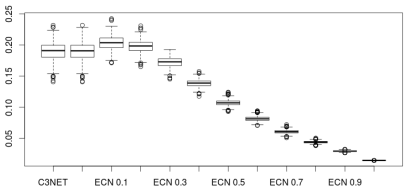
We now evaluate our Extended Core Network algorithm on simulated data, and compare it with C3NET. We use the same protocol to simulate the artificial data as in [2]. We use biological networks from *E. coli* [19, 20] and *S. cerevisiae* [21]. We generate our simulated data with sub-networks of these biological networks thanks to SynTREN [22]. SynTREN allows us to simulate the activity of genes



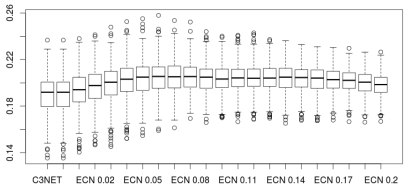
(a) *E. coli*, accepting rates from ECN_0 (0%) to ECN_1 (100%).



(b) *E. coli*, accepting rates from ECN_0 (0%) to ECN_0.2 (20%).



(c) Yeast, accepting rates from ECN_0 (0%) to ECN_1 (100%).



(d) Yeast, accepting rates from ECN_0 (0%) to ECN_0.2 (20%).

Fig. 1: Box plots of F_1 scores for C3NET (first left) and ECN with different accepting rates. The number following ECN indicates the accepting rates. The box plot next to the C3NET one is ECN_0.

from the selected network with noise. We simulate two different expression data: one with *E. coli*, the other with the yeast. SynTREN gives us two expression datasets of 100 samples and 200 genes.

For each simulated expression data, we performed 500 simulations. For each simulation, we selected randomly the samples used to infer the gene networks. To evaluate the error rate of each inference method, we use the F_1 score[3].

Figure 1 shows the box plots of the F_1 scores obtained by the simulations. We use ECN that provides undirected network, as C3NET, to compare them. The ECN_0 method corresponds to the Extended Core Network with a 0% accepting rate. It differs from C3NET if there are several genes g' which share the maximal mutual information with the gene g. ECN has the same complexity as C3NET, that is $\mathcal{O}(n^2)$, where n is the number of genes. We can see that ECN is better than C3NET when the accepting rate is low on *E. coli* (Figure 1a and b) and yeast (Figure 1c and d). The score of ECN depends on the accepting rate used. We can see that first, rising the accepting rate improves the inference method, but quickly it degrades it. When we look closer between 0 and 20% of accepting rate (Figure 1b and d), we can see that an accepting rate of 7% (ECN_0.07) is acceptable on simulated data. We have tested our method on other simulated data and we observed that an accepting rate between 5 and 10% is the best compromise.

[3] $F_1 = 2 \cdot \frac{\text{presicion} \cdot \text{recall}}{\text{precision} + \text{recall}}$

3 Differential Network Analysis

We now describe the differential analysis that we perform on two extended core networks to explore the role of anti-sense transcription. We first describe the sense and anti-sense data obtained in experiments about apple ripening. Then we define the notions of *change motifs* and *AS-impacted genes* revealed by the comparison of the two networks and we report the results for the apple data. To complete this analysis, we study how the interactions between AS-impacted genes are rewired.

3.1 Biological material

In order to study the impact of anti-sense transcripts, we use data of apple fruit during fruit ripening. We analyse RNA extracted from apple fruits thanks to the chip AryANE v1.0 containing 63011 predicted sense genes and 63011 complementary anti-sense sequences. This chip allows us to study the role of anti-sense transcripts at the genome-wide level by supplying transcriptional expression on both sense and anti-sense transcripts. We study the fruit ripening process described by two conditions: harvest (`H`) and 60 days after harvest (`60DAH`), and for each condition, 22 samples of apple fruit have been analysed. We first identify transcripts displaying significant differences between the two conditions (p-val<1%). With a further threshold of 1 log change between the two conditions, we found 931 sense and 694 anti-sense transcripts differentially expressed, with among them, 200 transcripts for which both sense and anti-sense fulfil the condition. We use these 1625 transcripts into our analysis. A differential functional analysis has been performed in [23] with those transcripts. The differential functional analysis reveals ontological terms that are significantly represented only when anti-sense transcripts are considered in conjunction with the sense transcripts. The fruit ripening is a stress-related condition involving "defense" genes, and our analysis revealed terms such as `hyperosmotic response` and `response to cold`, which are ontological terms related to a stress response. So this functional approach has shown the importance of taking into account the anti-sense information.

3.2 Change motifs and AS-impacted genes

Our differential network analysis compares the extended core network inferred from the sense data (S) with the one inferred from the sense and anti-sense data (SAS). We compare these two networks thanks to *AS-impacted genes* and *change motifs*. We define an *AS-impacted gene* as a sense gene s that is linked to one or many other sense actors in the S network, but these interactions are no longer present in the SAS network. In other words, an AS-impacted gene is a sense node s which has no outer link common to the S and the SAS networks. This change of interactions occurs because in the SAS core network, the most significant interactions (greatest values of the mutual information criterion) for this gene s are interactions with anti-sense actors. By merging the S and SAS networks,

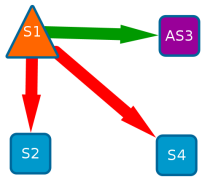


Fig. 2: Illustrations of *change motifs* with Extended Core Network. A sense node is represented in blue, an anti-sense node is represented in purple. The orange triangle-shaped node is an *AS-impacted gene*. A red link is a link only present in the S network. A green link is a link only present in the SAS network.

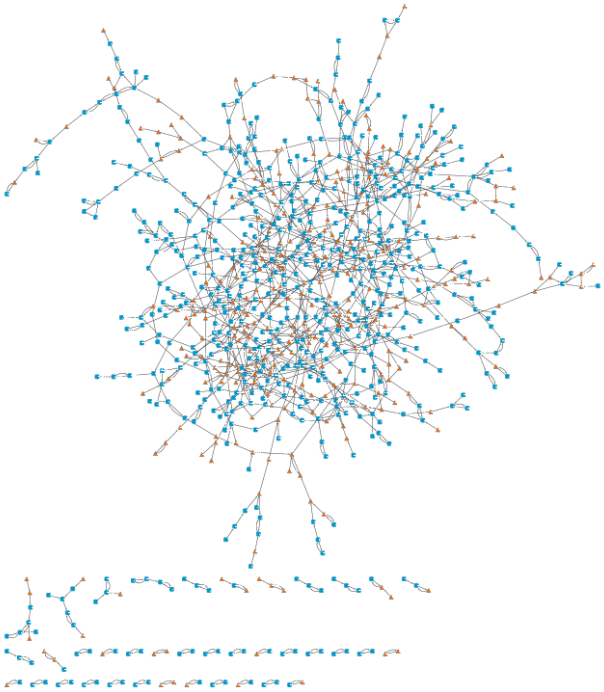


Fig. 3: Extended Core Network with a 5% accepting rate for sense-only `60DAH` experiment. Orange triangle-shaped nodes represent AS-impacted genes: they are connected to one or several sense nodes in this graph, but in the SAS network, they only have anti-sense neighbours.

we define a *change motif* [23, 24] as an AS-impacted gene surrounded by all of its neighbours from the S and SAS networks. The change motifs are identified by comparing the two adjacency matrices of S and SAS.Change motifs allow us to identify interactions that are "modified" by adding anti-sense actors into the network inference. The information provided by these change motifs would be omitted by a classical gene network inference relying only on sense data. Figure 2 illustrates what we define as a change motif. The S and SAS network are both drawn in the same graph. A sense node is blue and an anti-sense node is purple. A red link is a connection present only in S network, a green link is a connection present only in SAS network and a gray link is a connection present in both networks. Figure 2 represents a change motif because the node $S1$ has no outer link common to the S network ($S1 \rightarrow S2$ and $S1 \rightarrow S4$) and the SAS network ($S1 \rightarrow AS3$). It means that $S1$ shares a greater mutual information with $AS3$ than with $S2$ and $S4$.

The accepting rate of Extended Core Network determines the number of neighbours a node has, and consequently the number of change motifs. With an

accepting rate of 5%, we identified 300 change motifs in the `H` experiment and 308 motifs in the `60DAH` experiment. The number of change motifs shows that anti-sense transcripts have a real impact because about 30% of sense actors are AS-impacted genes.

Change motifs allow us to identify areas in the sense core network where anti-sense transcription has a deep impact, and we can identify anti-sense actors that modify the core network connections. The genes involved into a change motif will be studied to analyse the impact of anti-sense transcription.

Figure 3 shows the graphical result of the differential network analysis for the `60DAH` experiment. The graph was drawn using Cytoscape [25]. We can see that AS-impacted genes (orange triangles) are spread all over the network.

3.3 Steiner trees to compute the rewiring of AS-impacted sub-graphs

Figure 3 shows that it is difficult to extract relevant information from the graph visualisation. Therefore it is interesting to focus on some restricted parts of the network to help the study of biological processes. We see on Figure 3 that many AS-impacted genes are connected together and thus form a part of the network on which the anti-sense transcription has a great impact. Therefore a second stage of our differential analysis proposes a more thorough study of the sub-networks containing many AS-impacted genes. This analysis relies on the Steiner tree problem that we explain below, before presenting our approach.

Given an undirected graph $G = (V, E)$ with a set of vertices V and a set of edges E, the classical Steiner tree problem aims to find, for a set of vertices S called terminal nodes, a sub-graph G' of G containing S such that there exists a path between every pair of vertices of S and with a minimal number of edges. The nodes of $V - S$ necessary in the Steiner tree G' to obtain a connected graph are called Steiner vertices. When dealing with weighted edges, the general Steiner tree problem aims to find a minimal cost sub-graph spanning the terminal nodes. In bioinformatics, the Steiner tree problem can be used to extract information from the large databases of assessed molecular interactions. For example, if a biologist wants to study the implication of a set of proteins in the interactome, it can extract from a dedicated database a Steiner tree with these proteins as terminal nodes. As this combinatorial optimization problem is known to be NP-complete [26], several heuristic methods have been proposed to deal with large graphs. A recent experimental study on a large human protein-protein interaction network can be found in [27].

Our proposal to study the impact of AS-transcription on the gene interactions is the following. Let us recall that we have computed two extended core networks: the network S contains sense actors whereas the network SAS contains sense and anti-sense actors. The comparison of these two networks leads to identify AS-impacted genes, that are nodes of the network S strongly connected to anti-sense transcripts in the network SAS. It is thus interesting to study the impact of anti-sense transcription on groups of AS-impacted genes. We now define an *AS-impacted sub-graph* as a connected sub-graph of the sense network

S restricted to AS-impacted genes and containing at least two such genes. If SG is an AS-impacted sub-graph, that means that its nodes are connected in the sense network S but as these nodes are AS-impacted, these interactions no longer appear in the SAS network. Thus it is interesting to wonder what are the relationships between these nodes in the SAS network. One way to answer this question is to consider the Steiner tree problem. For each AS-impacted sub-graph of S, we look for a Steiner tree in the SAS network connecting its AS-impacted genes. If it can be found, this Steiner tree shows how the interactions present in S between these AS-impacted genes are rewired in SAS. Different ways can be proposed to the biologists to exploit this information. A visualisation of the rewiring of the interactions can help to focus on interesting interactions. A functional analysis of the AS-impacted genes can be realized to decipher the main functions impacted by anti-sense transcription.

To compute the minimum Steiner tree, we use a heuristic approach called the shortest paths based approximation [27]. The methods constructs the Steiner tree ST by successive steps. The first step selects one terminal node as the first node of the Steiner tree ST. Then each following step connects to ST, by a shortest path, one terminal node n that is not yet in the current ST. The node n is chosen as the closest node to all the nodes of the current ST. The algorithm stops when all the terminal nodes are in ST and the resulting tree is pruned by a final step that computes a minimum spanning tree for ST.

As shown experimentally in [27], the shortest paths heuristic is a good compromise between solution quality and computational speed for large graphs such as the ones addressed in bioinformatics networks. SteinerNet is a R package that implements four different approximate and one exact algorithms to solve the minimum Steiner tree problem described in [27]. The SteinerNet package has not been updated since 2013 but its dependencies have changed; so the SteinerNet package has been deleted from CRAN repository. We updated the package in order to use it with R version 3.2.0. We will be pleased to share this updated package upon request.

We use SteinerNet in order to compute Steiner trees of AS-impacted subgraphs from the `H` and `60DAH` experiments. The Steiner tree from Figure 4 is composed by 26 terminal nodes (orange) and 41 Steiner nodes (blue and purple). This tree is the rewiring of the 26 AS-impacted genes from one AS-impacted subgraph of the `60DAH` experiment. The 26 terminal nodes are the 26 AS-impacted genes, meaning that all the genes from the AS-impacted graph were rewired. The red links from Figure 4 are the connections from the S graph, that is the connections from the AS-impacted sub-graph. This representation allows us to see if the Steiner nodes connecting terminal nodes "replace" an existing connection or "create" a new one between two AS-impacted genes.

Steiner trees allow biologists to identify the rewiring among AS-impacted genes and thus the genes that are highly impacted by anti-sense transcripts. The next question is to find if Steiner nodes that allow this rewiring share biological functions with the AS-impacted genes. We identify GO terms associated with each gene from the Steiner tree and we look for common terms between terms

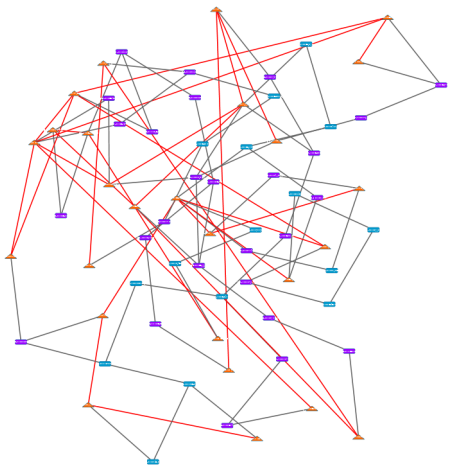


Fig. 4: Steiner tree of an AS-impacted sub-graph from the `60DAH` experiment. Orange nodes are AS-impacted genes, blue nodes are Steiner sense nodes, purple nodes are Steiner anti-sense nodes. Gray links are connections from the SAS graph, red links are connections from the S graph.

associated with terminal and Steiner nodes. Then we can give to the biologists a list of genes which are associated with the common terms.

From the Steiner tree of Figure 4, we identified 57 different terms associated to the terminal nodes, and 143 associated to the Steiner nodes, with only 16 terms in common. Because the fruit ripening studied with the `60DAH` experiment is a stress for the fruit, the biologists want to study the response of stress. We identify 4 common terms related to a stress response. Our method permits to identify a set of genes which appear in the Steiner trees with a GO term related to our biological function, the stress response. From the 29 AS-impacted sub-graphs of the `60DAH` experiment, we computed 35 Steiner trees and created a set of 10 genes related to the response of stress.

4 Conclusion

In order to study on a large scale the impact of anti-sense transcription, we propose to infer gene networks, with the particularity to integrate anti-sense transcripts in the process. We extend the recent ideas introduced in the field of differential network analysis, because in our framework, the main idea is to compare a network inferred from sense data and a network inferred from sense and anti-sense data. To achieve such a comparison, we first propose the Extended Core Network inference method. Extending the idea proposed by C3NET, this method builds a co-expression gene network where only the core of the network is considered by selecting the most important interactions for each gene. Secondly we extend the differential network analysis that we performed in [23] by using

the Extended Core Network method instead of C3NET. The comparison of two networks inferred on different sets of actors leads to define change motifs and AS-impacted genes, that highlight the impact of integrating anti-sense transcription in the gene network. We also propose to complete the differential analysis by studying the rewiring of the AS-impacted sub-graphs.The Steiner trees computed by our method provide a more global view to study the impact of anti-sense transcription on biological functions.

Acknowledgement

The authors would like to thank the GRIOTE project, funded by the Pays-de-la-Loire Region. We thank the Institut de Recherche en Horticulture et Semences teams for providing micro-array data on apple with sense and anti-sense probes.

References

1. Bansal, M., Belcastro, V., Ambesi-Impiombato, A., di Bernardo, D.: How to infer gene networks from expression profiles. Molecular Systems Biology **3**(1) (February 2007) 78
2. Altay, G., Emmert-Streib, F.: Inferring the conservative causal core of gene regulatory networks. BMC Systems Biology **4**(1) (September 2010) 132
3. Marbach, D., Costello, J.C., Küffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., The DREAM5 Consortium, Kellis, M., Collins, J.J., Stolovitzky, G.: Wisdom of crowds for robust gene network inference. Nature Methods **9**(8) (August 2012) 796–804
4. Friedel, S., Usadel, B., von Wiren, N., Sreenivasulu, N.: Reverse Engineering: A Key Component of Systems Biology to Unravel Global Abiotic Stress Cross-Talk. Frontiers in Plant Science **3** (December 2012) 294
5. Sharan, R., Ideker, T.: Modeling cellular machinery through biological network comparison. Nature Biotechnology **24**(4) (April 2006) 427–433
6. Altay, G., Asim, M., Markowetz, F., Neal, D.E.: Differential C3net reveals disease networks of direct physical interactions. BMC Bioinformatics **12**(1) (July 2011) 296
7. Ideker, T., Krogan, N.J.: Differential network biology. Molecular Systems Biology **8** (January 2012) 565
8. Barabási, A.L., Gulbahce, N., Loscalzo, J.: Network Medicine: A Network-based Approach to Human Disease. Nature reviews. Genetics **12**(1) (January 2011) 56–68
9. Pelechano, V., Steinmetz, L.M.: Gene regulation by antisense transcription. Nature Reviews Genetics **14**(12) (December 2013) 880–893
10. Yi, H., Richards, E.J.: A Cluster of Disease Resistance Genes in Arabidopsis Is Coordinately Regulated by Transcriptional Activation and RNA Silencing. The Plant Cell **19**(9) (January 2007) 2929–2939
11. Celton, J.M., Gaillard, S., Bruneau, M., Pelletier, S., Aubourg, S., Martin-Magniette, M.L., Navarro, L., Laurens, F., Renou, J.P.: Widespread anti-sense transcription in apple is correlated with siRNA production and indicates a large potential for transcriptional and/or post-transcriptional control. New Phytologist **203**(1) (March 2014) 287–299

12. Langfelder, P., Horvath, S.: WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics **9**(1) (December 2008) 559
13. Margolin, A.A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Favera, R.D., Califano, A.: ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. BMC Bioinformatics **7**(Suppl 1) (March 2006) S7
14. Zhang, X., Liu, K., Liu, Z.P., Duval, B., Richer, J.M., Zhao, X.M., Hao, J.K., Chen, L.: NARROMI: a noise and redundancy reduction technique improves accuracy of gene regulatory network inference. Bioinformatics **29**(1) (January 2013) 106–113
15. Bullard, J.H., Purdom, E., Hansen, K.D., Dudoit, S.: Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinformatics **11** (2010) 94
16. Qiu, X., Wu, H., Hu, R.: The impact of quantile and rank normalization procedures on the testing power of gene differential expression analysis. BMC Bioinformatics **14** (2013) 124
17. Steuer, R., Kurths, J., Daub, C.O., Weise, J., Selbig, J.: The mutual information: Detecting and evaluating dependencies between variables. Bioinformatics **18**(suppl 2) (January 2002) S231–S240
18. Kurt, Z., Aydin, N., Altay, G.: A comprehensive comparison of association estimators for gene network inference algorithms. Bioinformatics **30**(15) (2014) 2142–2149
19. Shen-Orr, S.S., Milo, R., Mangan, S., Alon, U.: Network motifs in the transcriptional regulation network of Escherichia coli. Nature Genetics **31**(1) (May 2002) 64–68
20. Ma, H.W., Kumar, B., Ditges, U., Gunzer, F., Buer, J., Zeng, A.P.: An extended transcriptional regulatory network of Escherichia coli and analysis of its hierarchical structure and network motifs. Nucleic Acids Research **32**(22) (2004) 6643–6649
21. Guelzim, N., Bottani, S., Bourgine, P., Képès, F.: Topological and causal structure of the yeast transcriptional regulatory network. Nature Genetics **31**(1) (May 2002) 60–63
22. Bulcke, T.V.d., Leemput, K.V., Naudts, B., Remortel, P.v., Ma, H., Verschoren, A., Moor, B.D., Marchal, K.: SynTReN: a generator of synthetic gene expression data for design and analysis of structure learning algorithms. BMC Bioinformatics **7**(1) (January 2006) 43
23. Legeay, M., Duval, B., Renou, J.P.: Differential functional analysis and change motifs in gene networks to explore the role of anti-sense transcription. In: International Symposium on Bioinformatics Research and Applications, Springer (2016) 117–126
24. Legeay, M., Duval, B., Renou, J.: Inference and differential analysis of extended core networks: A way to study anti-sense regulation. In: IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2016, Shenzhen, China, December 15-18, 2016. (2016) 284–287
25. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T.: Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research **13**(11) (November 2003) 2498–2504
26. Karp, R.M.: Reducibility among Combinatorial Problems. In: Complexity of Computer Computations. The IBM Research Symposia Series. Springer US (1972) 85–103
27. Sadeghi, A., Fröhlich, H.: Steiner tree methods for optimal sub-network identification: an empirical study. BMC Bioinformatics **14** (2013) 144