hARACNe: improving the accuracy of regulatory model reverse engineering via higher-order data processing inequality tests

In Sock Jang[1], Adam Margolin[1] and Andrea Califano[2]

[1]Sage Bionetworks, 1100 Fairview Avenue North, Seattle, WA 98109, USA
[2]Department of Systems Biology, Biochemistry and Molecular Biophysics, Biomedical Informatics, and Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, NY 10032, USA

A key goal of systems biology is to elucidate molecular mechanisms associated with physiologic and pathologic phenotypes based on the systematic and genome-wide understanding of cell context-specific molecular interaction models. To this end, reverse engineering approaches have been used to systematically dissect regulatory interactions in a specific tissue, based on the availability of large molecular profile datasets, thus improving our mechanistic understanding of complex diseases, such as cancer. In this paper, we introduce high-order Algorithm for the Reconstruction of Accurate Cellular Network (hARACNe), an extension of the ARACNe algorithm for the dissection of transcriptional regulatory networks. ARACNe uses the data processing inequality (DPI), from information theory, to detect and prune indirect interactions that are unlikely to be mediated by an actual physical interaction. Whereas ARACNe considers only first-order indirect interactions, i.e. those mediated by only one extra regulator, hARACNe considers a generalized form of indirect interactions via two, three or more other regulators. We show that use of higher-order DPI resulted in significantly improved performance, based on transcription factor (TF)-specific ChIP-chip data, as well as on gene expression profile following RNAi-mediated TF silencing.

1. Introduction

Cellular phenotypes are determined by a complex web of physical interactions between gene products [1]. Modelling these relationships helps to organize the list of parts encoded in the genome into functional genetic networks, a crucial step towards the understanding of mechanisms contributing to normal cell physiology as well as of their dysregulation in disease.

With the advent of high-throughput technologies, a large amount of molecular profile data have been generated from large numbers of samples associated with a variety of diseases. Through use of reverse engineering algorithms, these data have shown great promise in the dissection of transcriptional regulatory networks on a genomic scale [2–10]. While the caveats associated with the use of gene expression data for transcriptional network inference have been well documented and studied throughout the past decade [11], a preponderance of studies and high-impact discoveries in systems biology have established such approaches as widely accepted tools in a systems biologist's arsenal [6,12–18]. Reverse engineering approaches have been developed using the mathematical frameworks established in disciplines such as Bayesian networks [4,5,19], dynamical systems [8] and information theory [2,6,7,11–13,16,17,20]. For instance, the relevance network approach [2,21], one of the earliest proposed, assumes that genes co-expressed above statistical significance are more likely to represent regulatory interactions. Unfortunately, owing to long chains of regulatory events,

a large number of gene pairs may be co-expressed without necessarily implicating a physical interaction. Thus, relevance networks, whether based on mutual information or Spearman correlation, typically generate a large number of false positives.

A survey of such approaches and their relative merits is reviewed in [11] and are outside the scope of this article. Rather, the current study is intended to build on and improve the body of research associated with the information theoretic algorithm ARACNe (Algorithm for the Reconstruction of Accurate Cellular Network), which has emerged as a widely referenced approach that has been experimentally validated in numerous applications, leading to key biological discoveries [6,12–14].

ARACNe was developed to maintain the simplicity of relevance networks, while using rigorous information theoretic principles to eliminate the vast majority of indirect interactions (i.e. false positives). By limiting interactions only to pairs where at least one of the genes is a transcription factor (TF), ARACNe also addressed the issue of interaction directionality, since a TF can regulate a non-TF but the opposite is not true. Thus, the only undirected interactions in ARACNe networks are those between two TFs, where directionality cannot be disambiguated.

To eliminate indirect interactions, ARACNe uses the data processing inequality (DPI) theorem, from information theory, stating that information transferred directly (i.e. through a physical interaction) is always larger than information transferred indirectly, i.e. via an intermediary. Here, information is formally computed by the mutual information, $I(x;y) = S(x) + S(y) − S(x, y)$, where $S(u)$ represents the entropy of the variable u. We note that the DPI does not apply for other measures, such as Spearman correlation, unless interactions are linear, which is clearly not the case in biological systems. Experimental validation shows that the DPI removes a vast majority of false positive interactions, leading to highly accurate regulatory models [6,7,20,22].

Given its simplicity and the fact that it could consider each interaction independently, ARACNe was the first reverse engineering algorithm to successfully scale up to the complex regulatory networks of mammalian cells. For instance, 90 per cent of ARACNe-inferred and experimentally tested MYC targets in human B cells were validated as directly regulated by MYC [6], validation of targets in high-grade human glioma confirmed 40 out of 50 tested interactions for the transcription factors C/EBPβ, Stat3, RUNX1 and BHLHB2 [12], and the algorithm could correctly dissect a large number of transcriptional interactions involved in the synergistic control of germinal centre B cell proliferative programmes by FOXM1 and MYB [14,23,24]. More importantly, interrogation of ARACNe-inferred regulatory networks has allowed elucidation of key drivers of normal and disease-related phenotypes [6,12,13,15].

In this study, we explore improvements based on the nested iterative application of higher-order DPI tests [25,26]. The resulting algorithm, hARACNe (higher-order ARACNe), is thus designed to identify and remove a significant number of false positive interactions that could not be identified by first-order DPI (DPI[1]) analysis. Use of higher-order DPI analysis does not affect the generality of the method and is applicable to any network analysis, including those using static, time-course and even post-translationally predicted data [7,27].

Our analysis shows that hARACNe can systematically eliminate false positive interactions that were missed by DPI[1] logic of ARACNe, thus significantly improving inferred

TF–target interaction accuracy, based on MYC-binding data from ChIP-chip data [6,28] as well as gene expression profile (GEP) analysis following RNAi-mediated silencing of BCL6 [16]. Thus, our data show that hARACNe constitutes an advance in the identification of bona fide TF–target interactions of biological relevance.

2. Material and methods

2.1. Data sources

For this study, 254 previously published [6,17,20] GEPs were used, representing 17 distinct normal and tumour-related B-cell phenotypes from primary patient biopsies and tumour-derived cell lines. In addition, we used a set of 226 previously published [16] GEPs from human B-cell lymphoma, including normal samples from naive, memory and germinal centre B cells isolated from human tonsils and patient derived tumour samples, including diffused large B-cell lymphoma (DLBCL), follicular lymphoma and chronic lymphocytic leukaemia (CLL). The first dataset was profiled using the HG-U95Av2 GeneChip platform, whereas the second set was profiled using the HG-U133 Plus2.0 GeneChip (Affymetrix) [6,7,14,16,17,20]. All profiles discussed in this research are accessible from the Gene Expression Omnibus (GEO; National Center for Biotechnology Information), through GEO series accession nos. GSE2350 and GSE12195. The list of phenotypes is found in table 1.

Ramos and Mutu (human Burkitt's lymphoma) cell lines were analysed in ChIP-chip assays to identify the genes whose proximal promoter is bound the MYC protein. ChIP-chip significance analysis (CSA) was applied, where p-values were first derived for each probe from three replicate experiments, as described in [29]. Values of p were then integrated across a 500-base region surrounding the transcription start site of the gene, using a gamma cumulative distribution function. Each promoter was associated with the highest 500-base MYC-localization segment, and the false discovery rate (FDR) was computed using the Benjamini Hochberg procedure, as a function of gene rank. More precise procedures were described in [28].

In addition, DLBCL cell lines, including LY7, Pfeiffer and VAL, were profiled following lentivirus-mediated shRNA silencing of the BCL6 TF. The experimental procedures and conditions used to perform these experiments are described in [16]. Following BCL6 silencing, differentially expressed genes were identified (FDR < 0.05). ARACNE- and hARACNe-inferred BCL6 targets (BCL6 regulons) were then compared with differentially expressed genes.

For both the ChIP-chip and shRNA experiments, the procedures described earlier were used to infer a list of 'positive' TF–target interactions, which were compared with the predictions made by hARACNe and ARACNe to compute the number of true positives versus false positives removed by the additional pruning steps:

$$\text{TPgain} = \text{TP(hARACNe)} − \text{TP(ARACNe)}$$
$$\text{FPgain} = \text{FP(ARACNe)} − \text{FP(hARACNe)},$$

For comparison with ChIP-chip experiments, TP and FP are the number of ARACNe- or hARACNe-inferred MYC targets that were identified as bound and not bound by MYC, respectively. For comparison with shRNA experiments, TP and FP are the number of ARACNe- or hARACNe-inferred BCL6 targets that were identified as differentially regulated or not differentially regulated, respectively, following BCL6 inhibition. We note that the decrease in true positives due to removing interactions is equivalent to the increase in false negatives.

As with most 'gold standards' used in assessment of systems biology applications, we note that both the ChIP-chip and

Table 1. Human B-cell phenotypes in the two GEPs.

HG-U95 Av2		HG-U133 Plus2.0	
cell type	**no.**	**cell type**	**no.**
diffuse large B-cell lymphoma	68	naive B cell	5
Burkitt's lymphoma	33	memory B cell	5
follicular lymphoma	14	germinal centre B cell	11
mantle cell lymphoma	8	B-cell chronic lymphocytic leukaemia	16
B-cell chronic lymphocytic leukaemia	34	follicular lymphoma	38
hairy cell leukaemia	16	diffuse large B-cell lymphoma	128
multiple myeloma	4	undefined	23
Hodgkin's lymphoma	4		
primary effusion lymphoma	9		
splenic lymphoma with villous lymphocytes	12		
large cell lymphomas	5		
Burkitt's lymphoma type III	3		
undefined	2		
germinal centre B cells	17		
naive B cell	5		
memory B cells	5		
cord blood	5		
total	254	total	226

shRNA experiments are imperfect tests to identify true TF–target regulatory interactions. In particular, each assay only measures one (necessary but not sufficient) criterion of a regulatory interaction, and therefore may over-estimate the number of interactions. Moreover, each assay is performed in only a subset of tumour sub-types that were profiled in the microarray dataset, and therefore may underestimate the number of interactions, as only those active in the tested sub-types will be detected. Thus, although the exact number of true and false positives reported for each method are inexact, the TFgain and FPgain statistics should represent fair tests of the relative performance of each method. Moreover, the ChIP-chip and shRNA experiments measure different aspects of TF–target interactions and were performed on different TFs; therefore, consistent performance increases in comparison with these two distinct experiments lend strong support to the benefit of a method.

2.2. hARACNe algorithm design

The DPI is a simple but powerful theorem that starts from the simple axiom that as information is transferred through a lossy network, it can only be reduced and never increased. A Markov chain is at the heart of the DPI concept. A Markov chain is a linear sequence of states such that knowledge of the state at position i_0 makes all states at position $i > i_0$ independent of states at position $i < i_0$. In other words, any state provides all the necessary information to infer downstream states in the chain [26,30]. In a gene regulatory network, a Markov chain may represent a sequence of regulatory interactions $R_1 \rightarrow R_2 \rightarrow \cdots \rightarrow R_N$. In this case, DPI[1] analysis can remove indirect interactions of the type $R_i \rightarrow R_{i+2}$, where a distinct Markov state R_{i+1} exists that makes R_{i+2} independent of R_i (figure 1a). In this case, the DPI states that, if any information

is lost through interactions (an obvious true statement for any biological regulatory cascade), then $I(R_i; R_{i+2})$ is strictly smaller than both $I(R_i; R_{i+1})$ and $I(R_{i+1}; R_{i+2})$, i.e.

$$I(R_i; R_{i+2}) < \min(I(R_i; R_{i+1}), I(R_{i+1}; R_{i+2}))$$
$$\equiv \mathrm{DPI}(R_i \overset{R_{i+1}}{\Rightarrow} R_{i+2}). \qquad (2.1)$$

This means that the $R_i \rightarrow R_{i+2}$ interaction is removed because it is indirect through R_{i+1}.

To extend DPI[1] to higher-order DPIs, consider a Markovian quadruplet $R_i \rightarrow R_{i+1} \rightarrow R_{i+2} \rightarrow R_{i+3}$. Then, the second-order DPI (DPI[2]) could be expressed as follows:

$$I(R_i; R_{i+3}) < \min(I(R_i; R_{i+1}), I(R_i; R_{i+2}), I(R_{i+2}; R_{i+3}))$$
$$\text{with } \mathrm{DPI}(R_i \overset{R_{i+1}}{\Rightarrow} R_{i+2}) \text{ and } \mathrm{DPI}(R_{i+1} \overset{R_{i+2}}{\Rightarrow} R_{i+3})$$
$$\equiv \mathrm{DPI}(R_i \overset{R_{i+1}, R_{i+2}}{\Rightarrow} R_{i+3}). \qquad (2.2)$$

Or, in other words, the $R_i \rightarrow R_{i+3}$ interaction is removed as indirect through R_{i+1} and R_{i+2}.

Proof. It follows from the Markov property satisfying conditional independency among a Markov triplet as $I(R_i; R_{i+3}|R_{i+1}) = I(R_{i+1}; R_{i+3}|R_{i+2}) = 0$. Therefore, by repeatedly using the chain rule, $I(R_i; R_{i+1}, R_{i+2}) = I(R_i; R_{i+1}) + I(R_i; R_{i+2}|R_{i+1})$, we obtain this inequality as follows:

$$I(R_i; R_{i+3}) \leq I(R_i, R_{i+1}; R_{i+3})$$
$$= I(R_{i+1}; R_{i+3}) + I(R_i; R_{i+3}|R_{i+1}) = I(R_{i+1}; R_{i+3})$$
$$\leq I(R_{i+1}; R_{i+2}, R_{i+3})$$
$$= I(R_{i+1}; R_{i+2}) + I(R_{i+1}; R_{i+3}|R_{i+2}) = I(R_{i+1}; R_{i+2}).$$

Figure 1b shows graphically how the DPI[2] can be nested iteratively applied to identify additional false positive interactions within a Markov quadruplet.
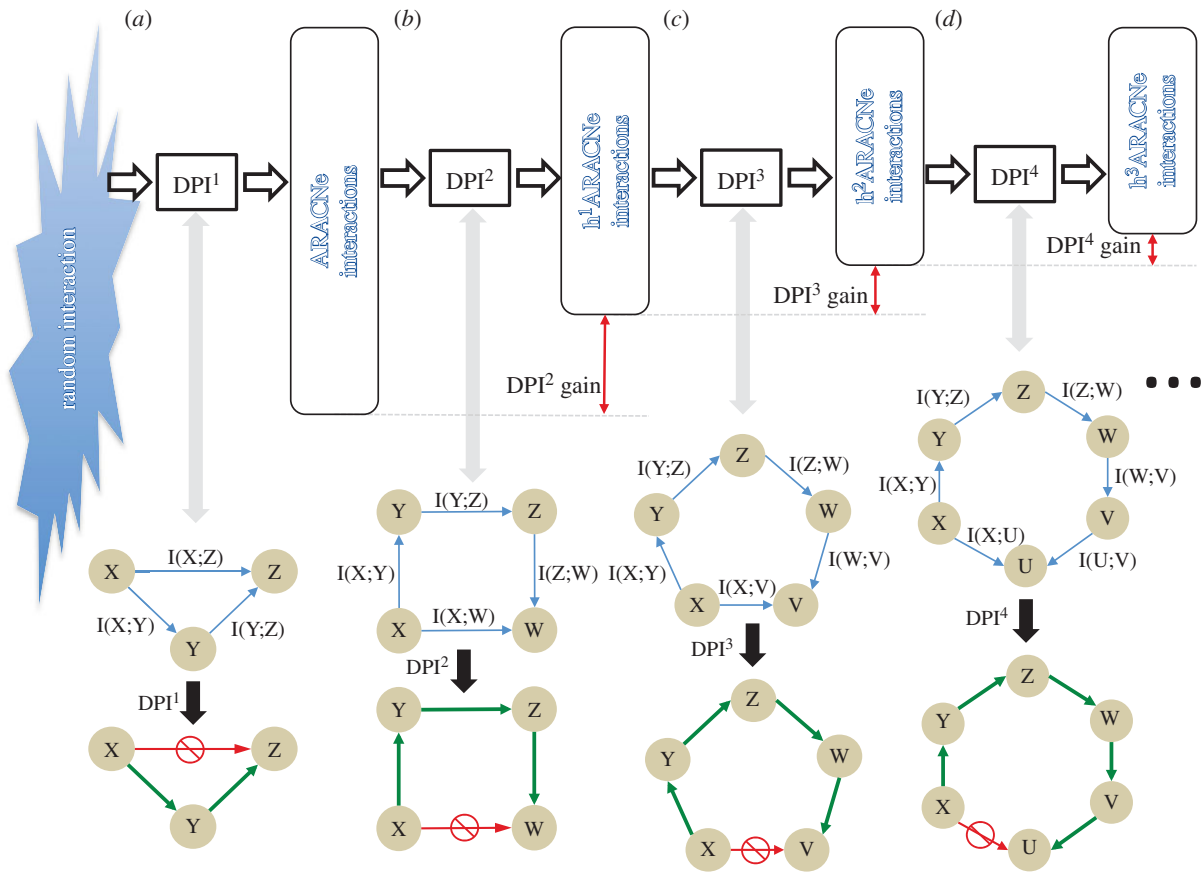
Figure 1. (a) Pruning process by DPI in triplet, (b) DPI^2 nested iterative procedure to detect indirect interaction in quadruplet, (c) DPI^3 sequential procedure to detect indirect interaction in quintuplet, and (d) DPI^4 sequential procedure to detect indirect interaction in sextuplet.

Next, consider a Markovian chain of any length $n + 2$, $R_i \rightarrow R_{i+1} \rightarrow \cdots \rightarrow R_{i+n} \rightarrow R_{i+n+1}$. We can apply the nth-order DPI (DPI^n) with conditional independence among any sub-chain as follows:

$$I(R_i; R_{i+n+1}) \leq \min(I(R_i; R_{i+1}), \ldots, I(R_{i+n}; R_{i+n+1}))$$
with {1st DPI},{2nd DPI},...,{$(n-1)$th DPI}

$$\equiv DPI(R_i \overset{R_{i+1}, \ldots, R_{i+n}}{\Rightarrow} R_{i+n+1}), \qquad (2.3)$$

where {kth DPI} means the super set of all possible combinations of kth-order DPI. Figure 1c,d, for instance, show examples for DPI^3 and DPI^4.

Higher-order DPI pruning process is always sequentially conducted after lower-order DPI is processed in advance, and pairwise mutual information between TFs and TGs are once computed in original ARACNe (DPI^1), and we iteratively prune indirect interactions from lower DPI to higher DPI. Thus, kth-order DPI does not need to revisit lower-order DPI application. This sequential pruning with higher-order DPIs may be applicable to any modified versions of ARACNe, which uses DPI^1 to prune indirect interactions.

3. Results

3.1. Reverse engineered transcriptional networks comparing ARACNe and hARACNe

The number of transcriptional interactions in the reconstructed networks was compared, and the distribution of TF targets is shown in table 2. In the human B-cell data profiled with the Affymetrix U95 platform, there were 12 600 probe IDs, of which 1225 corresponded to TFs, representing

Table 2. The number of interactions: comparisons with the adaptive partitioning method with nested iterative application of higher-order DPIs. The number of interactions inferred by ARACNe and hARACNe. Mutual information values were computed using the adaptive partitioning method [31]. See http://wiki.c2b2.columbia.edu/califanolab/index.php/Software/ARACNE.

adaptive partitioning	HG-U95 Av2	HG-U133 Plus2.0
ARACNe	155 526	198 766
h^1ARACNe	134 452	142 037
h^2ARACNe	132 018	138 697
h^3ARACNe	131 738	137 164

848 unique genes. ARACNe identified 155 526 transcriptional interactions at the probe ID level. After applying DPI^2, 134 452 interactions were left and 21 074 were removed as indirect interactions. When DPI^3 was applied, 132 018 interactions were left and an additional 2434 were removed as indirect. Analysis with fourth- and fifth-order DPI removed only 280 and no interaction, respectively. Thus, it appears that high-order DPI reaches saturation rapidly and that only DPI^2 and DPI^3 provide substantial false positive filtering power. Yet, high-order DPI, DPI^4 and above, have relatively high computational cost compared with efficiency of filtering power. With the second B-cell dataset, profiled using the Affymetrix U133 platform, there were 14 090 probe IDs, of which 1290 were TFs representing 1209 unique genes. ARACNe identified 198 766 transcriptional interactions at the probe ID level. DPI^2 produced 142 037 interactions, with
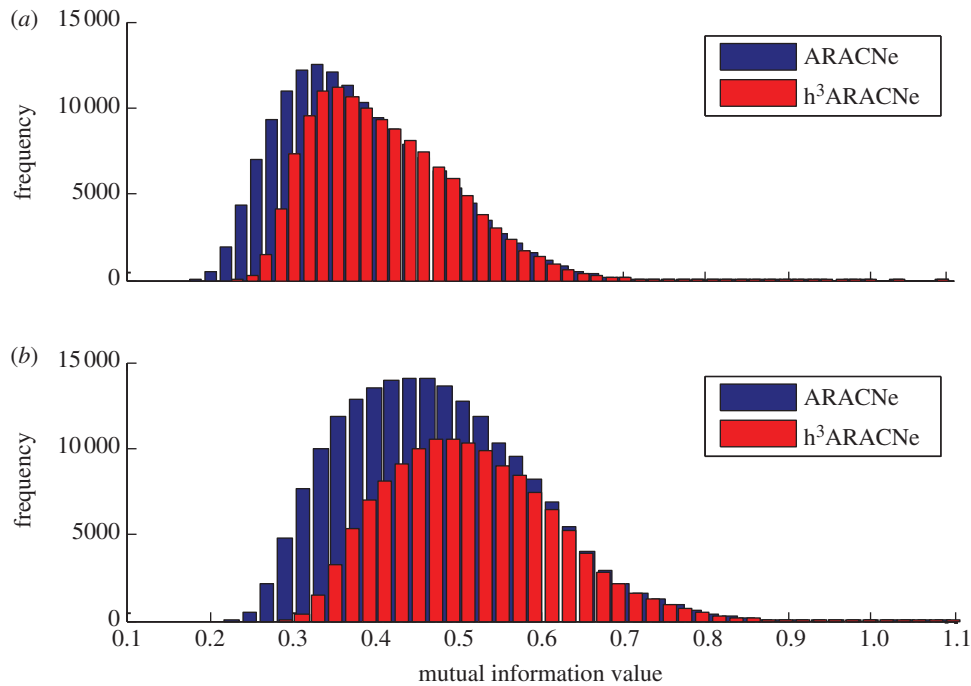
Figure 2. Overall mutual information comparison between original ARACNe and h³ARACNe networks. (*a*) U95 Av2 distribution change and (*b*) U133 Plus2 distribution change.

56 729 removed as indirect interactions. DPI³ produced 138 697 interactions, with 3340 additional ones removed as indirect interactions. Finally, DPI⁴ removed an additional 1533 interactions as indirect. We did not apply DPI of order greater than four.

In our comparative analyses, every TF had a different number of ARACNe- and hARACNe-inferred transcriptional targets. The distributions of transcriptional targets for ARACNe and hARACNe in the human B-cell analysis are shown in figure 2. For instance, from the U95 dataset, the ELK1 TF has 233 ARACNe-inferred targets, and we randomly selected ELK1 among the TFs which have more than 200 targets in ARACNe. Of these, 97 were removed by DPI², leaving only 136 targets (figure 3). However, DPI³ removed only seven additional targets and no additional targets were removed by DPI⁴. An identical pattern was observed in the U133 human B-cell dataset analysis. Among all TFs, ZNF267 had the largest number of targets removed by higher-order DPI analysis. Of 223 ARACNe-inferred targets, 121 were removed by DPI², and 30 and 5 additional targets were removed by DPI³ and DPI⁴, respectively, leaving only 86 targets.

As described, hARACNe removed a significant percentage of interactions inferred by ARACNe. This may improve the accuracy of inferred networks if the filtered interactions were enriched in false positives (i.e. indirect interactions). To test this hypothesis, we compared the ARACNe and hARACNe predictions against two datasets providing orthogonal evidence of TF–target interactions (see figure 2 for more details).

The first dataset was generated through ChIP-chip experiments designed to detect direct biochemical interactions between the MYC proto-oncogene and genome-wide promoter regions. The second dataset used RNAi experiments to detect genes differentially regulated upon inhibition of BCL6. We used data related to MYC and BCL6 based on availability of datasets and based on their central role as proto-oncogene B-cell leukaemias and lymphomas.

3.2. ChIP-chip functional validation of MYC targets

We first compared ARACNe- and hARACNe-inferred targets of the MYC TF against genes with experimentally assessed interactions between the MYC protein and their proximal promoter regions. The assessment was based on existing ChIP-chip assays in Ramos and Mutu (human Burkitt's lymphoma) cell lines [29]. Based on the CSA algorithm [28], 5307 and 3310 putative MYC-binding targets could be predicted in Ramos and Mutu cell lines, respectively. Comparing ARACNe- and hARACNe-inferred MYC regulons, from the U95Av2 data, 18 interactions were removed by hARACNe. Based on Ramos cell assays, 13 of these 18 interactions (hypergeometric test: p-value = 0.0178) were identified as false positives (i.e. no detected interaction between MYC and the gene's promoter region) and only five were identified as true positives (table 3). Based on Mutu cell assays, 14 of the 18 interactions were identified as false positives, and only four of 18 were identified as true positives (hypergeometric test: p-value = 2.8713×10^{-7}).

As a result, it appears that hARACNe was three times more likely to remove false interactions not supported by a corresponding MYC-binding site in the proximal promoter region than to remove interactions with target whose promoter was bound by MYC.

3.3. Differentially expressed genes following shRNA-mediated BCL6 silencing

To further check whether the additional pruning produced by higher-order DPI improved the accuracy of TF–target prediction, we performed lentivirus-mediated shRNA silencing of the BCL6 gene and tested the enrichment of differentially expressed genes in ARACNe- versus hARACNe-inferred BCL6 targets. GEPs were measured using the Affymetrix HG-U95A GeneChip platform with DLBCL and CLL cell lines
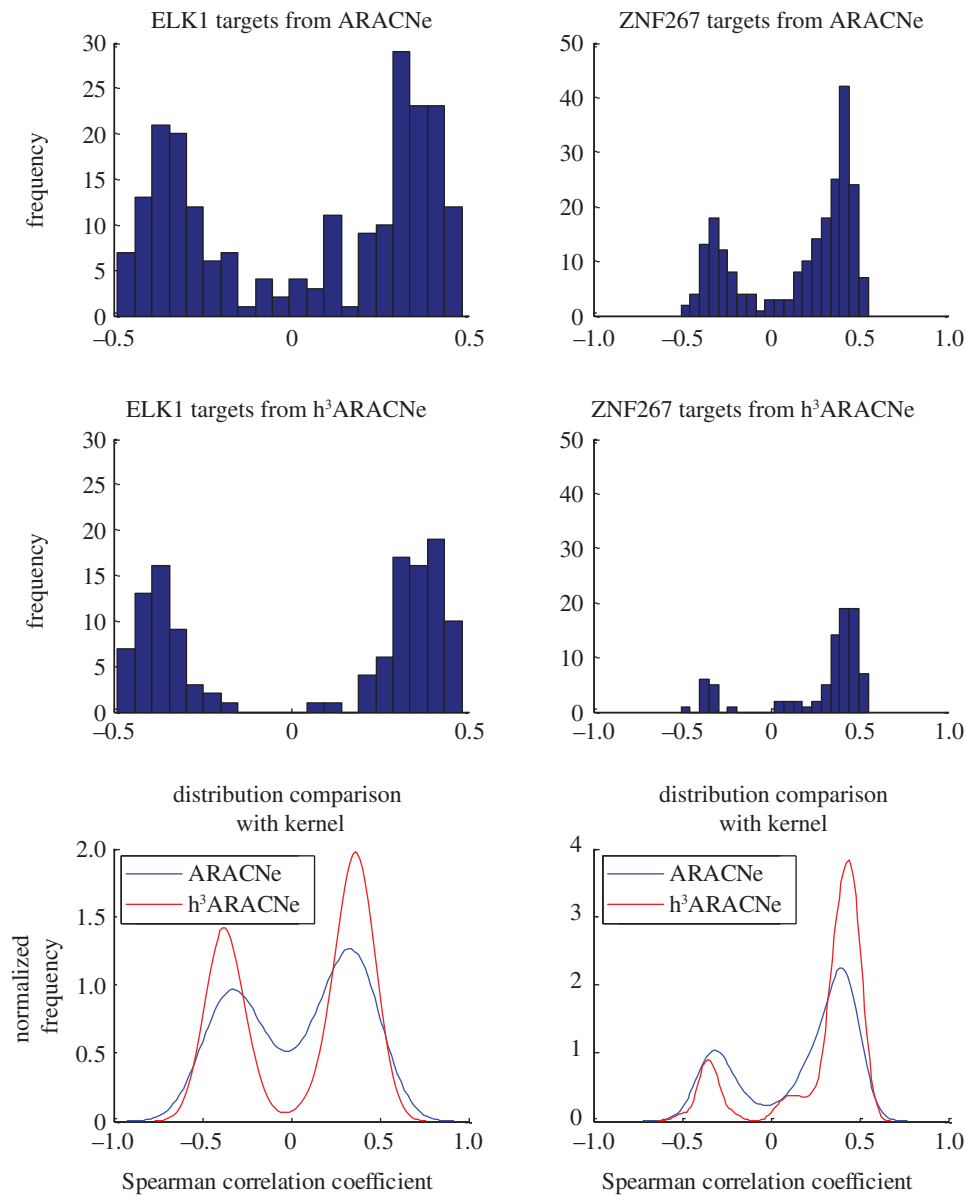
Figure 3. Correlation distribution of the most differential TF targets between original ARACNe and h³ARACNe networks.

(LY7, Pfeiffer and VAL cell lines), where cells were infected with control non-target shRNA or validated shRNA targeting BCL6. Differentially expressed genes were identified using fold-change criteria and a 0.05 FDR threshold using a non-parametric U-test. In LY7, Pfeiffer and VAL cell lines, 1507, 3706 and 3199 differentially expressed genes were identified, respectively. From LY7 data, out of 334 ARACNe-inferred BCL6 targets, 318 were identified by sequential application of DPI^2, DPI^3 and DPI^4. Twenty gains were achieved by hARACNe. Eighteen out of 20 were not differentially expressed following BCL6 silencing and may thus be considered false positives. Conversely, two BCL6 targets were newly identified due to consensus scoring analysis followed by 100 bootstrappings (hypergeometric test: p-value $= 5.5955 \times 10^{-14}$), which were differentially expressed following BCL6 knockdown and may have been considered true positives (table 4). In Pfeiffer cells, all 16 interactions removed by hARACNe(nested iterative procedures of DPI^2, DPI^3 and DPI^4) were identified as false positives and none were identified as true positives (hypergeometric test: p-value $= 8.6441 \times 10^{-7}$). Finally, in VAL cells, all 16 interactions removed by hARACNe(nested iterative procedures of DPI^2, DPI^3 and DPI^4) were also identified as false positives

and none were identified as true positives (hypergeometric test: p-value $= 2.5294 \times 10^{-8}$). Tables 3 and 4 also include the statistical analysis (Fisher's exact test) with TPs and FPs from experimentally validated targets with each step of hARACNe-inferred targets in order to show how significantly targets were identified by either ARACNe or each cumulative order of hARACNe.

4. Discussion

The goal of this work was to improve the widely used ARACNe algorithm by further reducing false positive interactions, thus leading to more accurate inference of interaction networks. We note that our work is not intended to address other foundational issues of reverse engineering approaches or of the ARACNe algorithm, such as the assumptions inherent to the use of mRNA data. Rather, we use the widely accepted ARACNe assumptions as a starting point and assess the ability of higher-order DPI analysis to improve prediction accuracy based on theoretical arguments and comparison with independent 'gold standard' datasets. Within this limited context, we

Table 3. (a) ChIP-chip experimental validation with Ramos and Mutu cell lines for MYC targets and (b) hypergeometric test in order to check if the gain is significantly derived from FPs.

(a)

MYC: ChIP-chip	Ramos			Mutu		
	TP	FP	Fisher's exact test p-value of inferred interactions	TP	FP	Fisher's exact test p-value of inferred interactions
HG-U95Av2						
DPI^1	214	258	0.1428	170	302	3.09×10^{-6}
DPI^2	210	249	0.1077	167	292	1.85×10^{-6}
DPI^3	209	246	0.09444	166	289	1.63×10^{-6}
DPI^4	209	245	0.08747	166	288	1.40×10^{-6}
gain	5	13	n.a.	4	14	n.a.

(b)

HG-U95Av2	Ramos			Mutu		
	TP	FP	hypergeometric test p-value	TP	FP	hypergeometric test p-value
	5	13	0.01779333	4	14	2.87×10^{-4}
total gain between DPI^1 and DPI^4 (nested)	18			18		

Table 4. (*a*) Gain and loss BCL6 targets comparison between ARACNe and hARACNe by shRNA library and (*b*) hypergeometric test in order to check if the gain is significantly derived from FPs.

(a)

BCL6: shRNA library	LY7			Pfeiffer			VAL		
	TP	FP	Fisher's exact test p-value of inferred interactions	TP	FP	Fisher's exact test p-value of inferred interactions	TP	FP	Fisher's exact test p-value of inferred interactions
HG-U95Av2									
DPI^1	54	280	0.0136	144	190	1.06×10^{-7}	117	217	6.97×10^{-5}
DPI^2	57	265	0.001614	146	176	1.60×10^{-9}	118	204	6.32×10^{-6}
DPI^3	56	262	0.001988	144	174	2.28×10^{-9}	117	201	5.54×10^{-6}
DPI^4	56	262	0.001988	144	174	2.28×10^{-9}	117	201	5.54×10^{-6}
gain	2	18	n.a.	0	16	n.a.	0	16	n.a.

(b)

	LY7				Pfeiffer				VAL			
	total gain between DPI^1 and DPI^4(nested)	TP	FP	hypergeometric test p-value	total gain between DPI^1 and DPI^4(nested)	TP	FP	hypergeometric test p-value	total gain between DPI^1 and DPI^4(nested)	TP	FP	hypergeometric test p-value
HG-U95Av2	20	2	18	5.60×10^{-14}	16	0	16	8.64×10^{-7}	16	0	16	2.53×10^{-8}

believe that this work represents a useful contribution to the field, based on the widespread use of ARACNe and the benefit of an extension improving its accuracy.

From a theoretical standpoint, we note that higher-order, indirect Markov chain interactions detected by hARACNe would also be eliminated by repeated application of the DPI[1], under certain assumptions. Specifically, these would require that (i) the network have a tree (or locally tree-like) structure, (ii) the network contain only pairwise interactions, and (iii) mutual information be measured without errors. Given that each of these assumptions may be individually violated and given the large number of potential pairwise interactions, higher-order DPIs provide additional filters to increase the global Markov chain stringency in ARACNe-inferred networks, thus eliminating indirect interactions that may have been missed by DPI[1]. In this context, the experimental data we have provided, demonstrating a higher accuracy of hARACNe networks, represents the most meaningful test of the method.

Yet, there are a few computational limitations in hARACNe. First, since hARACNe starts from an ARACNe-inferred network, on which it nested iteratively applies higher-order DPI analyses, it must consider a large number of candidate Markov chain paths, traversing up to four interactions, and is thus computationally intensive. Our results, however, show that DPI[2] provides the greatest pruning effect while higher-order DPIs have a significantly lower detection rate. Thus, one may want to consider whether to apply only DPI[2] or up to DPI[3] to obtain the highest increase in accuracy at the lowest computational cost.

Our experimental design for evaluating hARACNe relied on two datasets providing orthogonal evidence of direct TF–target interactions. Specifically, we used ChIP-chip assays and shRNA-mediated silencing to show that hARACNe preferentially removes false positive interactions. In BCL6 silencing experiments, interactions removed by hARACNe had significantly more false positives than true positives. Similarly, when considering ChIP-chip experimental data on MYC-binding sites, a similar improvement in the false positive predictions was introduced by hARACNe. Overall, the higher-order DPI analysis inferred a more reliable network, compared with the DPI[1] procedure of the original ARACNe algorithm.

References

1. Barabasi AL, Oltvai ZN. 2004 Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* **5**, 101–113. (doi:10.1038/nrg1272)

2. Butte AJ, Kohane IS. 2000 Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. *Pac. Symp. Biocomput.* **4**, 418–429.

3. D'Haeseleer P, Liang S, Somogyi R. 2000 Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics* **16**, 707–726. (doi:10.1093/bioinformatics/16.8.707)

4. Hartemink AJ, Gifford DK, Jaakkola TS, Young RA. 2002 Combining location and expression data for principled discovery of genetic regulatory network models. *Pac. Symp. Biocomput.* **2002**, 437–449.

5. Imoto S, Higuchi T, Goto T, Tashiro K, Kuhara S, Miyano S. 2003 Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks. *Proc. IEEE Comput. Soc. Bioinform. Conf.* **2**, 104–113.

6. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. 2005 Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **37**, 382–390. (doi:10.1038/ng1532)

7. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera R, Califano A. 2006 ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinform.* **7**(Suppl. 1), S7. (doi:10.1186/1471-2105-7-S1-S7)

8. Gardner TS, Bernardo D, Lorenz D, Collins JJ. 2003 Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* **301**, 102–105. (doi:10.1126/science.1081900)

9. di Bernardo D, Thompson MJ, Gardner TS, Chobot SE, Eastwood EL, Wojtovich AP, Elliott SJ, Schaus SE, Collins JJ. 2005 Chemogenomic profiling on a genomewide scale using reverse-engineered gene networks. *Nat. Biotechnol.* **23**, 377–383. (doi:10.1038/nbt1075)

10. Friedman N et al. 2000 Using Bayesian networks to analyze expression data. *J. Comput. Biol.* **7**, 601–620. (doi:10.1089/106652700750050961)

11. Margolin AA, Califano A. 2007 Theory and limitations of genetic network inference from microarray data. *Ann. NY Acad. Sci.* **1115**, 51–72. (doi:10.1196/annals.1407.019)

12. Carro MS et al. 2010 The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318–325. (doi:10.1038/nature08712)

13. Della Gatta G et al. 2012 Reverse engineering of TLX oncogenic transcriptional networks identifies RUNX1 as tumor suppressor in T-ALL. *Nat. Med.* **18**, 436–440. (doi:10.1038/nm.2610)

14. Lefebvre C et al. 2010 A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. *Mol. Syst. Biol.* **6**, 377. (doi:10.1038/msb.2010.31)

15. Zhao X et al. 2009 The N-Myc-DLL3 cascade is suppressed by the ubiquitin ligase Huwe1 to inhibit proliferation and promote neurogenesis in the developing brain. *Dev. Cell* **17**, 210–221. (doi:10.1016/j.devcel.2009.07.009)

16. Basso K et al. 2010 Integrated biochemical and computational approach identifies BCL6 direct target genes controlling multiple pathways in normal germinal center B cells. *Blood* **115**, 975–984. (doi:10.1182/blood-2009-06-227017)

17. Mani KM, Lefebvre C, Wang K, Lim WK, Basso K, Dalla-Favera R, Califano A. 2008 A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas. *Mol. Syst. Biol.* **4**, 169. (doi:10.1038/msb.2008.2)

18. Palomero T, Ferrando AA. 2009 Genomic tools for dissecting oncogenic transcriptional networks in human leukemia. *Leukemia* **23**, 1236–1242. (doi:10.1038/leu.2008.394)

19. Friedman N. 2004 Inferring cellular networks using probabilistic graphical models. *Science* **303**, 799–805. (doi:10.1126/science.1094068)

20. Wang K et al. 2009 Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nat. Biotechnol.* **27**, 829–839. (doi:10.1038/nbt.1563)

21. Rice JJ, Tu Y, Stolovitzky G. 2005 Reconstructing biological networks using conditional correlation analysis. *Bioinformatics* **21**, 765–773. (doi:10.1093/bioinformatics/bti064)

22. Wang K, Alvarez MJ, Bisikirska BC, Linding R, Basso K, Dalla Favera R, Califano A. 2009 Dissecting the interface between signaling and transcriptional regulation in human B cells. *Pac. Symp. Biocomput.* **20**, 264–275.

23. Carrera J, Rodrigo G, Jaramillo A. 2009 Model-based redesign of global transcription regulation. *Nucleic Acids Res.* **37**, e38. (doi:10.1093/nar/gkp022)

24. Lim WK, Lyashenko E, Califano A. 2009 Master regulators used as breast cancer metastasis classifier. *Pac. Symp. Biocomput.* **2009**, 504–515.

25. Han Y, Qiu J, Guo Z, Gao H, Song Y, Zhou D, Yang R, 2007 Comparative transcriptomics in *Yersinia pestis*: a global view of environmental modulation of gene expression. *BMC Microbiol.* **7**, 96. (doi:10.1186/1471-2180-7-96)

26. Cover TMT, Joy A. 2006 *Elements of information theory*, 2nd edn. New York, NY: Wiley-Interscience.

27. Zoppoli P, Morganella S, Ceccarelli M. 2010 TimeDelay-ARACNE: reverse engineering of gene networks from time-course data by an information theoretic approach. *BMC Bioinform.* **11**, 154. (doi:10.1186/1471-2105-11-154)

28. Margolin AA, Palomero T, Sumazin P, Califano A, Ferrando AA, Stolovitzky G. 2009 ChIP-on-chip significance analysis reveals large-scale binding and regulation by human transcription factor oncogenes. *Proc. Natl Acad. Sci. USA* **106**, 244–249. (doi:10.1073/pnas.0806445106)

29. Hudson ME, Snyder M. 2006 High-throughput methods of regulatory element discovery. *Biotechniques* **41**, 673–681.

30. Han TS, Kobayashi K. 2007 *Mathematics of information and coding*, 1st edn. Providence, RI: American Mathematical Society.

31. Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A. 2006 Reverse engineering cellular networks. *Nat. Protoc.* **1**, 662–671. (doi:10.1038/nprot.2006.106)