



OPEN ACCESS

EDITED BY
Mukesh Dhamala,
Georgia State University, United States

REVIEWED BY
Nguyen Quoc Khanh Le,
Taipei Medical University, Taiwan
Yuanyue Li,
Genome Center, College of Biological
Sciences, University of California, Davis,
United States

*CORRESPONDENCE
Zhi-Ping Liu,
zpliu@sdu.edu.cn

SPECIALTY SECTION
This article was submitted to Bioprocess
Engineering,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

RECEIVED 27 May 2022
ACCEPTED 15 August 2022
PUBLISHED 27 September 2022

CITATION
Chen G and Liu Z-P (2022), Inferring
causal gene regulatory network via
GreyNet: From dynamic grey
association to causation.
Front. Bioeng. Biotechnol. 10:954610.
doi: 10.3389/fbioe.2022.954610

COPYRIGHT
© 2022 Chen and Liu. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Inferring causal gene regulatory network *via* GreyNet: From dynamic grey association to causation

Guangyi Chen¹ and Zhi-Ping Liu^{1,2*}

¹Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, Jinan, Shandong, China, ²Center for Intelligent Medicine, Shandong University, Jinan, Shandong, China

Gene regulatory network (GRN) provides abundant information on gene interactions, which contributes to demonstrating pathology, predicting clinical outcomes, and identifying drug targets. Existing high-throughput experiments provide rich time-series gene expression data to reconstruct the GRN to further gain insights into the mechanism of organisms responding to external stimuli. Numerous machine-learning methods have been proposed to infer gene regulatory networks. Nevertheless, machine learning, especially deep learning, is generally a “black box,” which lacks interpretability. The causality has not been well recognized in GRN inference procedures. In this article, we introduce grey theory integrated with the adaptive sliding window technique to flexibly capture instant gene–gene interactions in the uncertain regulatory system. Then, we incorporate generalized multivariate Granger causality regression methods to transform the dynamic grey association into causation to generate directional regulatory links. We evaluate our model on the DREAM4 *in silico* benchmark dataset and real-world hepatocellular carcinoma (HCC) time-series data. We achieved competitive results on the DREAM4 compared with other state-of-the-art algorithms and gained meaningful GRN structure on HCC data respectively.

KEYWORDS

gene regulatory network inference, dynamic grey association, adaptive sliding window, causation, machine learning

1 Introduction

The gene regulatory network plays a central role in understanding the mechanisms of gene expression regulation, complex diseases, and cellular heterogeneity (Xiang et al., 2019; Fang et al., 2020; Liao et al., 2022). Compared to the genomes between human *Homo sapiens* and yeast *Saccharomyces cerevisiae*, we can conclude that the complexity in life does not result from the number of genes, but the essence and dynamics of the interactions between genes (Huynh-Thu and Sanguinetti, 2019; Freyre-González et al., 2022; Jansen et al., 2022), i.e., gene regulatory network.

Diverse methods have been proposed to infer gene regulatory networks. Recently, the emergence and rise of machine-learning methods to infer gene regulatory networks have dated back to GENIE3 (Vân et al., 2010), which won the DERAM4 (Dialog on Reverse Engineering Assessment and Methods) *in silico* multifactorial challenge. Then, Jump3 (Huynh-Thu and Sanguinetti, 2015) was proposed to learn the promoter state of the target gene from candidate regulators based on the decision tree. SWING (Finkle et al., 2018) introduced sliding windows to address heterogeneous time delays in the network structure inference. To further improve the inference accuracy, BTNET (Sungjoon et al., 2018) and BiXGBoost (Zheng et al., 2018) transformed the random forest into gradient boosting algorithms. Regularization-based regression introduced different constraints for *de novo* GRN reconstruction (Phan and Rosemary, 2018; Ghosh Roy et al., 2020; Zhang et al., 2021). BETS applied bootstrap elastic net regression based on Granger causality to infer the GRN (Lu et al., 2021). Recurrent neural network (RNN) was utilized to model gene interactions due to the superior capability of tracking complicated temporal behaviors in the real underlying regulatory system (Cheng et al., 2011; Biswas and Acharyya, 2018). Although machine-learning methods achieved great success, the internal procedures are unknown to us or they are known but hard to be understood by observers (Guidotti et al., 2018).

In this article, we propose an interpretable machine-learning method named GreyNet, i.e., dynamic grey association and regression, to infer the gene regulatory network from time-course gene expression data. We first apply dynamic grey association to model intricate underlying the regulatory system. Different from the static grey association, we assimilate the adaptive sliding window technique to conduct dynamic analysis which can better capture instant interactions over time. The dynamic grey association takes advantage of local temporal information to search for candidate regulators. Then, we embed the Granger causality framework (Arnold et al., 2007; Finkle et al., 2018; Li et al., 2020) based on regression models which can find causal and directional regulatory links. Through this hybrid strategy, GreyNet can better model gene interactions in real scenarios and is easier to be understood in network inference.

2 Materials and methods

In this section, we demonstrate GreyNet to reconstruct the directional graph of GRN $G(v, e)$ from time-series gene expression data. Time-series can effectively disclose the dynamic interactions of genes with time (Haonan et al., 2020). As for the data $D\{G \in R^{m \times n}\}$, v represents the nodes or vertexes and e represents the regulatory links between pairwise vertexes. For the edge e_{ij} , it represents gene j under the upstream regulation of gene i . The arrangement of the whole gene expression time-series is as follows:

$$G(i, j) = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} \end{bmatrix} \quad (1)$$

where each column represents an instance and each row represents the value of the selected instance at a specific timestamp. The overall framework of GreyNet is illustrated in Figure 1. We will demonstrate the major steps in the following sections.

2.1 Dynamic grey association

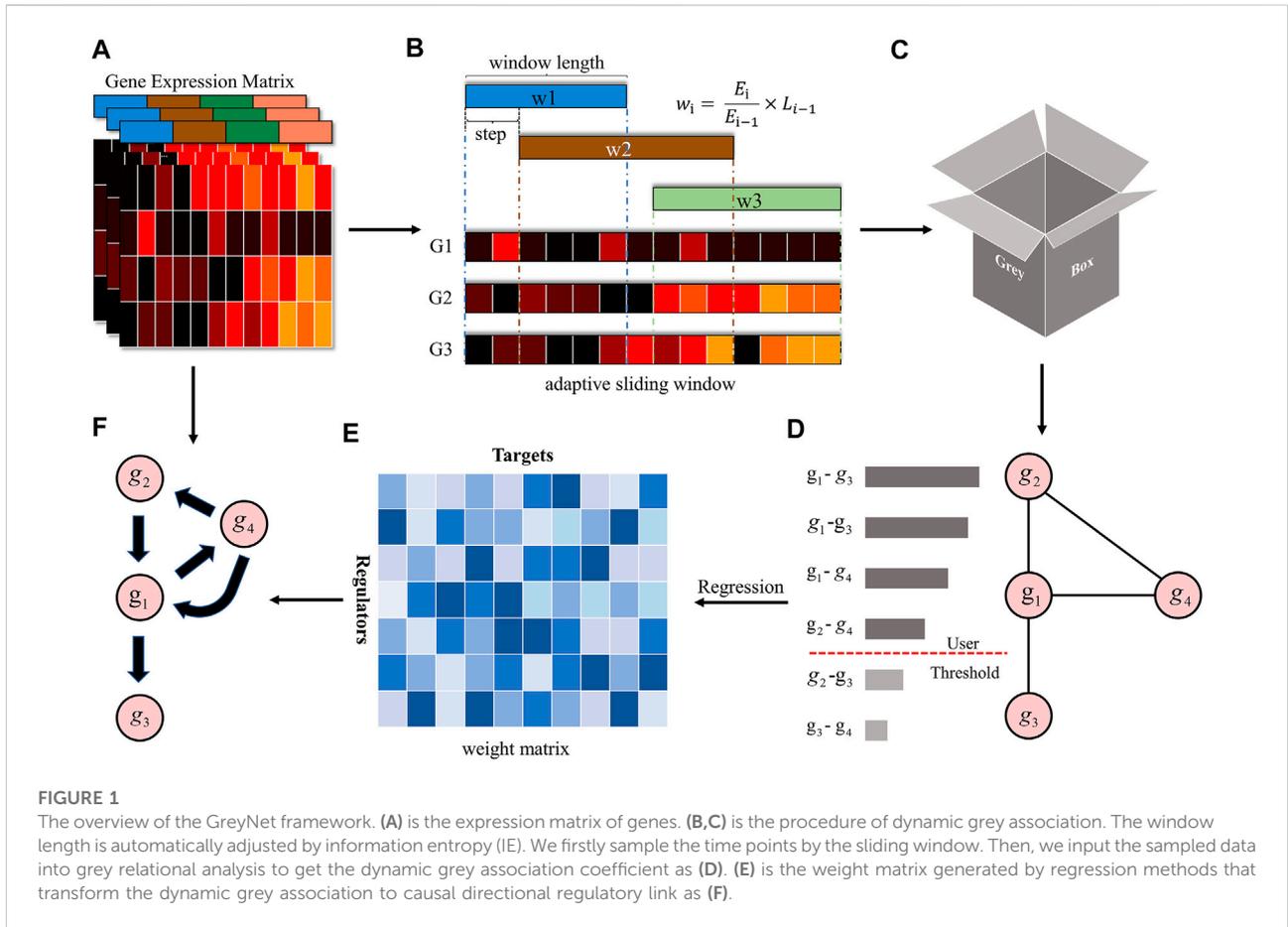
The biological system is very complicated. Although we have revealed important mechanisms of gene regulation, we are still far from being fully clear about it. In the gene regulatory system, genes generally are under the regulation of various types of regulators and most of them are still unknown or unobserved (Ming et al., 2020). With poor data and limited knowledge at present, the GRN inference works on uncertain systems, namely the grey system between black and white. In other words, reconstructing the GRN is with partially known and unknown information, but we want to draw out the valuable GRN structure from observed gene expression data. In this condition, we propose the dynamic grey association to model gene interactions. The dynamic grey association consists of grey relational analysis (Deng, 1989; Sallehuddin et al., 2008; Yuansheng et al., 2019) and the adaptive sliding window (Papadimitriou et al., 2006). Under the circumstance of the relationships between two components are usually variational over time in the biological system, we integrate grey association with the adaptive sliding window which endows the capacity of flexibly tracking instant interactions in time-series data. Therefore, the dynamic grey association is much more interpretable than the “black box” and conforms to current knowledge.

We formally introduce how to obtain the dynamic grey association coefficient here. Firstly, we demonstrate the way to design the adaptive sliding window (Figure 1B). For the gene expression time-series $D\{G \in R^{m \times n}\}$, we take the first-order difference to get the time derivative. We select one gene as a target or reference node $y = \{g_1, g_2, \dots, g_n\}$ and the rest of the genes are comparative nodes $X_k = \{G_1, \dots, G_{k-1}, G_{k+1}, \dots, G_m\}$, where G_k is a vector that represents the expression data of gene k .

$$\nabla y = \{\nabla g_1, \nabla g_2, \dots, \nabla g_n\} \quad (2)$$

$$\nabla g_n = g_n - g_{n-1} \quad (3)$$

Then, we normalize the time derivative to transform it into probability by using the *softmax* function. Information entropy is a good choice to evaluate the amount of information.



$$p_i = \text{softmax}(|\nabla y|) = \frac{\exp(|\nabla g_i|)}{\sum_{g_i \in \nabla y} \exp(|\nabla g_i|)} \quad (4)$$

$$E_i = - \sum p_i \times \log_2(p_i) \quad (5)$$

Finally, we divide the present information entropy by the previous one to obtain the window coefficient. The preceding window length multiplies the window coefficient to get the current sliding window. The window length will be adjusted by information entropy automatically.

$$L_i = \frac{E_i}{E_{i-1}} \times L_{i-1} \quad (6)$$

In every adaptive sliding window, we get the individual grey relational grade. Initially, the target node subtracts the comparative nodes' corresponding elements to get the absolute value of the first-order norm residual ∇ .

$$\nabla = |y(i) - x_k(i)| \quad (7)$$

From the maximum and minimum of the residual ∇ , the association coefficient r_i is given by:

$$\xi_k(i) = \frac{\min_k \min_i \nabla + \rho * \max_k \max_i \nabla}{\nabla + \rho * \max_k \max_i \nabla} \quad (8)$$

$$r_i = \frac{1}{L} \sum_{i=1}^L \xi_k(i) \quad (9)$$

where L is the length of the adaptive sliding window. ρ is the distinguished coefficient which is positively related to distinguishing the difference (Kuo et al., 2008). Finally, we average all the sliding windows to get the dynamic grey association score DGA :

$$DGA = \frac{1}{n} \sum_{i=1}^n r_i \quad (10)$$

DGA determines the dominant factors between the multivariable and target gene based on the geometric curve. The higher the value of DGA , the higher the association between the two variables.

2.2 Causation

GRN represents directional causal regulatory relationships among genes (Leng et al., 2019). However, the dynamic grey association cannot depict causality for pairwise genes. In this case, we incorporate the Granger causality framework to turn the dynamic grey association into causation. The Granger causality is intuitive and defined that the past values of the cause make a larger contribution to predicting the future values of the effect than auto-regression. However, it is time-consuming and ignores the possible interactions between features. Inspired by LASSO Granger (Arnold et al., 2007), we apply multivariate regression strategies to identify the subset of features on which the feature is conditionally dependent, namely, we formulate it to a problem of feature selection, given the fact that the best estimator for the target variable is the one with the least error or the maximum gain (Arnold et al., 2007; Li et al., 2020).

In this study, we mainly introduce four regressors to infer GRN. The four regressors we selected are: bagging algorithm random forest (RF) (Friedman, 2001), gradient boosting algorithm XGBoost (Chen and Guestrin, 2016), L_1 -penalty LASSO (Tibshirani, 1996), and L_2 -penalty Ridge (Hoerl and Kennard, 1970). We regress target variable y_t in terms of X^{lag} . For different regressors, the objective function of the regressor and optimizing strategy are different. With regards to RF, it is similar to GENIE3 (Van et al., 2010). The goal of the tree-based regressor is to build a fine decision tree structure in terms of splitting nodes. The criterion for measuring the quality of splitting is the mean squared error, which is equal to variance reduction. The evaluation of the objective function for the target is as:

$$I(N) = \#SVar(N) - \#S_{pt}Var(S_{pt}) - \#S_{pf}Var(S_{pf}) \quad (11)$$

where # is the number of the samples; S is the sample sets that reach the node N is a single tree; S_{pt} is the sample sets predicted true; and S_{pf} is the sample sets that predicted false. Nevertheless, the criterion of Xgboost regression is modified by:

$$L_{split} = \frac{1}{2} \left[\frac{G_{left}^2}{H_{left} + \lambda} + \frac{G_{right}^2}{H_{right} + \lambda} - \frac{G^2}{H + \lambda} \right] - \gamma \quad (12)$$

$$G_{left} = \sum_{i \in I_{left}} g_i \quad H_{left} = \sum_{j \in I_{left}} h_j$$

$$I = I_{left} \cup I_{right}$$

where g_i and h_i are the first-order and second-order gradient statistics. λ and γ are both complexity parameters. I_{left} and I_{right} are the sample sets after splitting. Regularization-based regression introduces different regularization terms to prevent overfitting and get optimal reconstruction. The objective function of regularized regression is:

$$Obj = \frac{1}{m} \sum_{i=1}^m (y_i - X^{lag}w)^2 + \lambda \sum_{j=1}^n \|w_j\|^p \quad (13)$$

where $\|w_j\|^p$ is the regularization term. λ is the regularization coefficient. When p equals one or two, it represents the LASSO regression or Ridge regression.

Based on the aforementioned regression criteria, we can obtain weight matrix w from $w^y = regressor(y, X^{lag})$ (as shown in Figure 1E). To confirm the causal regulatory direction, we summarize LASSO Granger as an example that other regression methods are similar to it. If x regulates y , $x^t \in w^y$ for some t but $y_t \notin w^x$; if y regulates x , $y^t \in w^x$ for some t but $x_t \notin w^y$; and if x and y regulate each other, then $x^t \in w^y$ and $y^t \in w^x$.

3 Results

3.1 Gene expression data

GreyNet focuses on time-course bulk gene expression data. Our method is not suitable for single-cell RNA-seq data due to dropout events. In the past decade, the DREAM challenge has been the standard benchmark dataset to evaluate the quality of the reconstructing algorithm (Marbach et al., 2009; Marbach et al., 2010). Therefore, we firstly validate GreyNet on the DREAM4 time-series dataset which contains two sizes of networks, size10 and size100, and each network includes five subnetworks. To further test the performance of our model, we evaluate it on a real-world hepatocellular carcinoma dataset (Yang et al., 2019a). HCC expression profiles are detected from 105 samples represented stepwise from pre-neoplastic lesions to HCC. 105 samples cover the nine development stages of HCC. To preprocess HCC data from NCBI GEO, we map probeset IDs to NCBI official gene symbols through the GEO annotation file. If one gene has multiple probeset mappings, the probeset with the maximum inter quartile expression range (IQR) is selected (Liu et al., 2014). Finally, we employ prior knowledge to separate the TF expression and target expression (Lambert, 2018) to conduct transcriptomic GRN inference by GreyNet. The detailed descriptions of the DREAM4 and HCC time-series expression data are shown in Table 1. The development stages of HCC from normal to hepatocellular carcinoma are shown in Table 2.

3.2 Evaluation metrics

We mainly use two common metrics, AUROC and AUPRC, to evaluate our model. AUROC is calculated from the ROC curve, showing the trade-off between true positive rate (TPR) and false positive rate (FPR) across different thresholds. AUPRC is just the area under the PR curve, where the x-axis is Precision and the y-axis is Recall. The other measurement metrics, such as Precision, Matthews correlation coefficient (MCC), and Accuracy, are shown in Supplementary Table S1 (Tng et al., 2021; Le and Ho, 2022).

TABLE 1 The description of the datasets used in the experiments.

Network	#TF	#Gene	#Timepoints	#Samples
DREAM4 <i>in-silico</i> size 10	10	10	21	5
DREAM4 <i>in-silico</i> size 100	100	100	21	10
HCC	21	258	10	105

TABLE 2 The development stages of HCC.

Development	Notation	#Samples
Normal	N	13
Chronic Hepatitis with low grade	FL	8
Chronic Hepatitis with high grade	FH	12
Cirrhosis	CS	12
Dysplastic nodules with low grade	DL	11
Dysplastic nodules with high grade	DH	11
Early hepatocellular carcinoma	eHCC	5
Hepatocellular carcinoma (TG1)	TG1	9
Hepatocellular carcinoma (TG2)	TG2	12
Hepatocellular carcinoma (TG3)	TG3	12

$$TPR = Recall = \frac{TP}{TP + FN} \quad (14)$$

$$FPR = \frac{FP}{FP + TN} \quad (15)$$

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

3.3 Performance on DREAM4 data

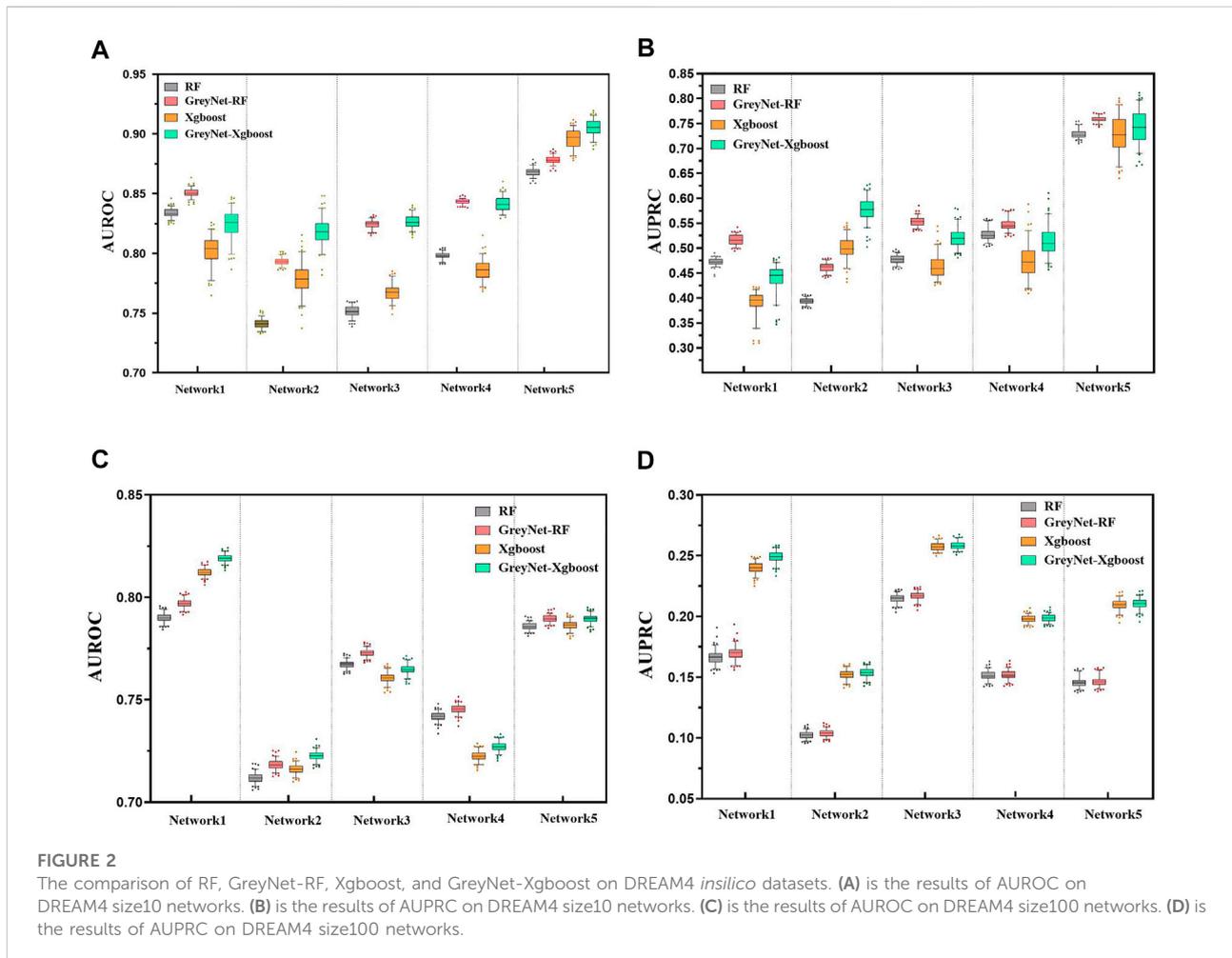
In the DREAM4 challenge, we implement four different regression strategies on the DREAM4 dataset combined with the dynamic grey association. We select one typical model from each bagging and gradient boosting algorithm, i.e., random forest (RF) and Xgboost. Analogously, we also incorporate two different regularized regression methods L_1 -norm LASSO and L_2 -norm Ridge. For each regression method, we run them 100 times to reduce randomness. To further validate the capacity of the grey technique, we compare the four different regression methods with and without the dynamic grey association. The performances of bagging and gradient boosting comparisons are shown in Figure 2. As shown, GreyNet-RF and GreyNet-Xgboost are significantly superior to the corresponding regressors without the dynamic grey association in both AUROC [$P_{RF} - value = 3.93 \times 10^{-18}$; $P_{Xgboost} - value = 6.57 \times 10^{-18}$, Wilcoxon test] and AUPRC [$P_{RF} - value = 3.96 \times 10^{-18}$; $P_{Xgboost} - value = 3.96 \times 10^{-18}$]. In the 100 trials, two

regularized regression methods have little changes in AUROC and AUPRC. The results of the two regularized regression comparisons are shown in Figure 3. From Figure 3, we can see that GreyNet-LASSO and GreyNet-Ridge significantly outperform LASSO and Ridge regressors with respect to both AUROC [$P_{LASSO} - value = 1.86 \times 10^{-18}$; $P_{Ridge} - value = 1.55 \times 10^{-23}$] and AUPRC [$P_{LASSO} - value = 2.10 \times 10^{-18}$; $P_{Ridge} - value = 1.55 \times 10^{-23}$]. Therefore, the dynamic grey association is effective and efficient to improve the GRN structure inference.

We further compare our model with other seven state-of-the-art methods, including GENIE3-lag (Huynh-Thu, 2011), Jump3 (Huynh-Thu and Sanguinetti, 2015), SWING (Finkle et al., 2018), BTNET (Sungjoo et al., 2018), BiXGBoost (Zheng et al., 2018), BETS (Lu et al., 2021), and TIGRESS (Anne-Claire et al., 2012). In the five subnetworks, the average AUROC and AUPRC of GreyNet achieve 0.854 ± 0.032 , and 0.622 ± 0.108 in size10 and 0.768 ± 0.036 , 0.222 ± 0.039 in size100, respectively. The detailed results of AUROC and AUPRC in each subnetwork are shown in Table 3. The ROC and PR curves of GreyNet are shown in Supplementary Figure S1. From Table 3, we can see that our model achieves the highest AUROC in networks 2, 3, and 4 of size10 among the comparing methods. Other than network 1, our model gets the highest AUPRC. In DREAM4 *in-silico* size100, GreyNet achieves all of the highest AUROC and AUPRC in the five networks other than AUROC in network 4. The complete performances of GreyNet with four different regression strategies are shown in Supplementary Tables S2, S3.

3.4 Performance on hepatocellular carcinoma data

Hepatocellular carcinoma (HCC) accounts for >90% of liver cancers with a five-year survival of only 18% and the fourth leading cause of cancer-related deaths (Yang et al., 2019b; Villanueva, 2019). It is estimated that more than one million individuals will be affected by liver cancer annually by 2025 and the World Health Organization (WHO) predicted that the mortality of liver cancer will also arrive at one million in 2030 (Yang et al., 2019b; Villanueva, 2019; Llovet et al., 2021). It is imperative to search for important biomarkers of molecular and immune classes to guide therapy. A GRN of HCC will significantly benefit this kind of search.



In this article, we apply GreyNet to the HCC time-course expression data. We select the top 500 regulatory links in HCC GRN by the score of the inference. The inferred HCC GRN is shown in Figure 4. From Figure 4, we can see that, TP53, PIK3CA, AXIN1, MET, APC, CTNNB1, and TERT (all aforementioned genes are protein-coding genes) are elite genes highly related to HCC. TP53, TERT (promoter), and CTNNB1 are dominant mutational driver cancer genes, which account for 21–31 %, 44–65 %, and 27–40 % of patients with HCC (Yang et al., 2019b; Llovet et al., 2021). In terms of TF genes, ARID2 is correlated with the initiation and progression of HCC (Schulze et al., 2015); NFE2L2 is involved in hepatocarcinogenesis and progression (Nault et al., 2014; Niu et al., 2016); HNF1A is related to promoting genetic liver adenomatosis occurrence and possibly further malignant transformation to HCC (Zucman-Rossi et al., 2015).

Then, we enrich the HCC GRN by NOA (Wang et al., 2011). The result of the enriched Gene Ontology (GO) biological process and documented pathways in KEGG are

shown in Table 4. From Table 4, we can see that the deregulations of the multiple signal pathways in HCC affect cell proliferation, RNA, nucleobase, nucleoside, nucleotide, nucleic acid metabolic process, and liver development. The ‘Wnt signaling pathway’ has a significant impact on cancer development and cancer mechanism evolution (Polakis, 2012; Zhan et al., 2017). The dysregulation of Wnt/ β -catenin (a key component of the Wnt pathway) brings out the aberrant activation of signaling in HCC (Waisberg and Saba, 2015). Activated β -catenin translocates to the nucleus, interacting with TCF (T cell factor) and LEF (lymphoid enhancer-binding factor), and activates the transcription of the target genes which participate in CSC maintenance and EMT (Zhan et al., 2017; Farzaneh et al., 2021). Ultimately, it will lead to cell proliferation, angiogenesis, and anti-apoptosis. The ‘‘JNK pathway’’ is implicated in multiple cancers, including the regulation of liver tumorigenesis. In the mice model, it is shown that the increased expression of p21 (a cell-cycle

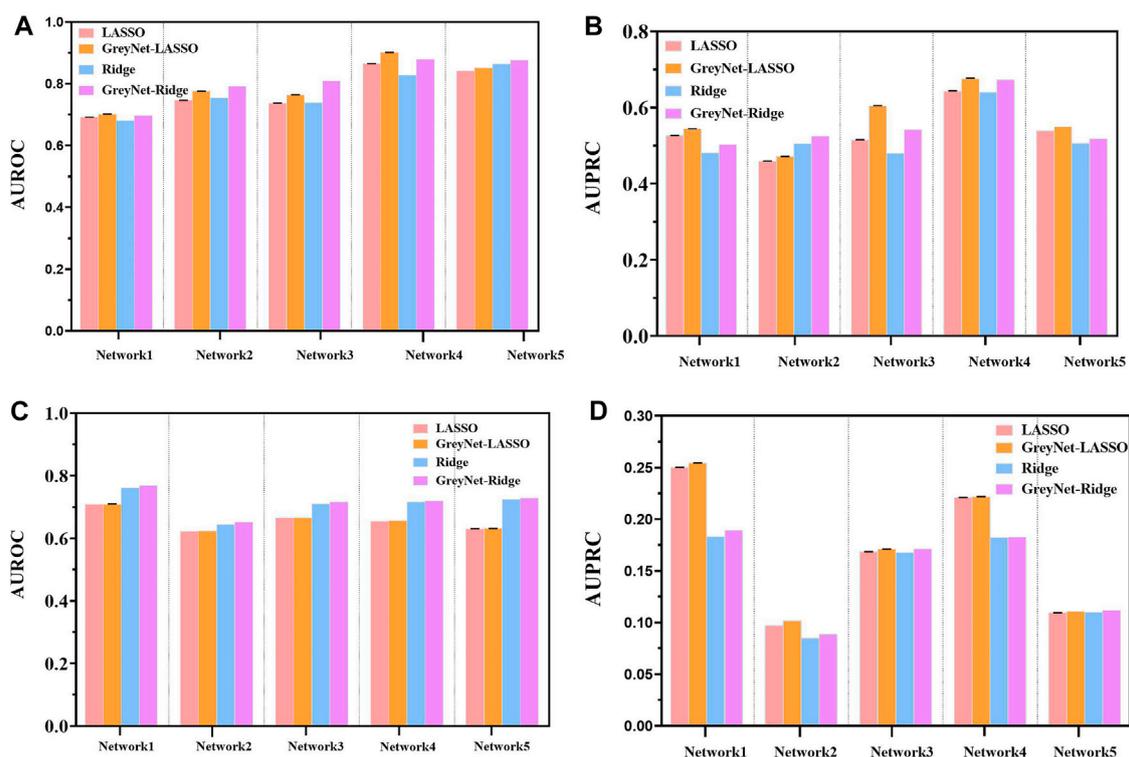


FIGURE 3

The comparison of LASSO, GreyNet-LASSO, Ridge, and GreyNet-Ridge on DREAM4 in *in silico* datasets. (A) is the results of AUROC on size10 networks. (B) is the results of AUPRC on size10 networks. (C) is the results of AUROC on size100 networks. (D) is the results of AUPRC on size100 networks.

inhibitor) can cause impaired proliferation. In human HCC, the activity of JNK can affect liver cell proliferation *via* p21 and c-Myc (a negative regulator of p21). It is found that the growth of xenografted human HCC cells can be reduced by pharmacologic inhibition of JNK (Hui et al., 2008; Dimri and Satyanarayana, 2020). The ‘receptor tyrosine kinase pathways’ implicate in activating multiple downstream signals, including the epidermal growth factor (EGF) receptor, the fibroblast growth factor (FGF) receptor, the hepatocyte growth factor (HGF/c-MET), the stem cell growth factor receptor c-KIT, the platelet-derived growth factor (PDGF) receptor, and the vascular endothelial growth factor (VEGF) receptor (Dimri and Satyanarayana, 2020). The “transforming growth factor-beta” (TGF- β) involves multiple stages of HCC development from liver injury toward fibrosis, cirrhosis, and cancer. In hepatocarcinogenesis, TGF- β performs as a suppressor factor in the early stages. However, TGF- β contributes to tumor progression latterly (Fabregat and Caballero-Díaz, 2018). The consistency between these enriched functions and the prior knowledge about HCC implies the effectiveness of GreyNet.

4 Discussion

Limited by current technology and knowledge, the underlying gene regulatory mechanisms in cells are not very clear to us. It is reasonable to assume that gene interactions behave as a grey system. Moreover, the similarity and association of the two components are variational, evolving the time in the biological systems. It is less useful to assign a single static score fraction to two variables over an entire time-series. In this condition, we turned the static coefficient to the dynamic grey association by incorporating the adaptive sliding window technique to capture the dynamic evolution, which is much more aligned with real and known gene regulations.

The dynamic grey association is not enough to mine the causality in GRN. Thus, we further introduced both linear and non-linear regression methods to search for causal links by temporal information. Causal relationships between the variables can disclose the origin of the outcome and contribute to decision making. The Granger causal regression model is easy to be understood and explained.

TABLE 3 The comparative results of models on DREAM4 data.

DREAM4 *in silico* size10

Method	Network1		Network2		Network3		Network4		Network5	
	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC	AUPRC
GreyNet	0.839	0.479	0.833	0.628	0.838	0.580	0.844	0.611	0.917	0.812
BTNET(GB)	0.834	0.516	0.698	0.362	0.682	0.473	0.822	0.560	0.934	0.774
BTNET(AB)	0.875	0.552	0.719	0.370	0.719	0.465	0.791	0.506	0.903	0.701
SWING-RF	0.832	0.508	0.779	0.525	0.815	0.546	0.728	0.441	0.925	0.753
SWING-Dionesus	0.743	0.469	0.786	0.484	0.789	0.421	0.772	0.540	0.807	0.625
BiXGBoost	0.816	0.573	0.784	0.422	0.771	0.376	0.787	0.533	0.888	0.741
GENIE3-lag	0.834	0.476	0.741	0.391	0.750	0.478	0.797	0.520	0.869	0.734
Jump3	0.700	0.442	0.698	0.308	0.717	0.401	0.784	0.486	0.841	0.619
TIGRESS	0.758	0.375	0.602	0.222	0.618	0.200	0.764	0.324	0.804	0.411

DREAM4 *in silico* size100

GreyNet	0.822	0.258	0.725	0.160	0.771	0.267	0.731	0.205	0.789	0.221
BTNET(GB)	0.776	0.186	0.694	0.113	0.759	0.235	0.723	0.143	0.758	0.165
BTNET(AB)	0.776	0.207	0.699	0.116	0.770	0.224	0.740	0.158	0.780	0.169
SWING-RF	0.793	0.192	0.723	0.116	0.759	0.214	0.742	0.193	0.775	0.160
SWING-Dionesus	0.772	0.124	0.700	0.095	0.709	0.194	0.727	0.187	0.771	0.143
BiXGBoost	0.744	0.138	0.682	0.075	0.716	0.119	0.702	0.106	0.728	0.090
GENIE3-lag	0.790	0.167	0.711	0.103	0.767	0.215	0.742	0.152	0.786	0.146
Jump3	0.724	0.099	0.623	0.057	0.696	0.077	0.662	0.072	0.696	0.074
TIGRESS	0.715	0.054	0.532	0.037	0.483	0.018	0.467	0.018	0.521	0.022

The highest AUROC and AUPR are shown in bold for each network.

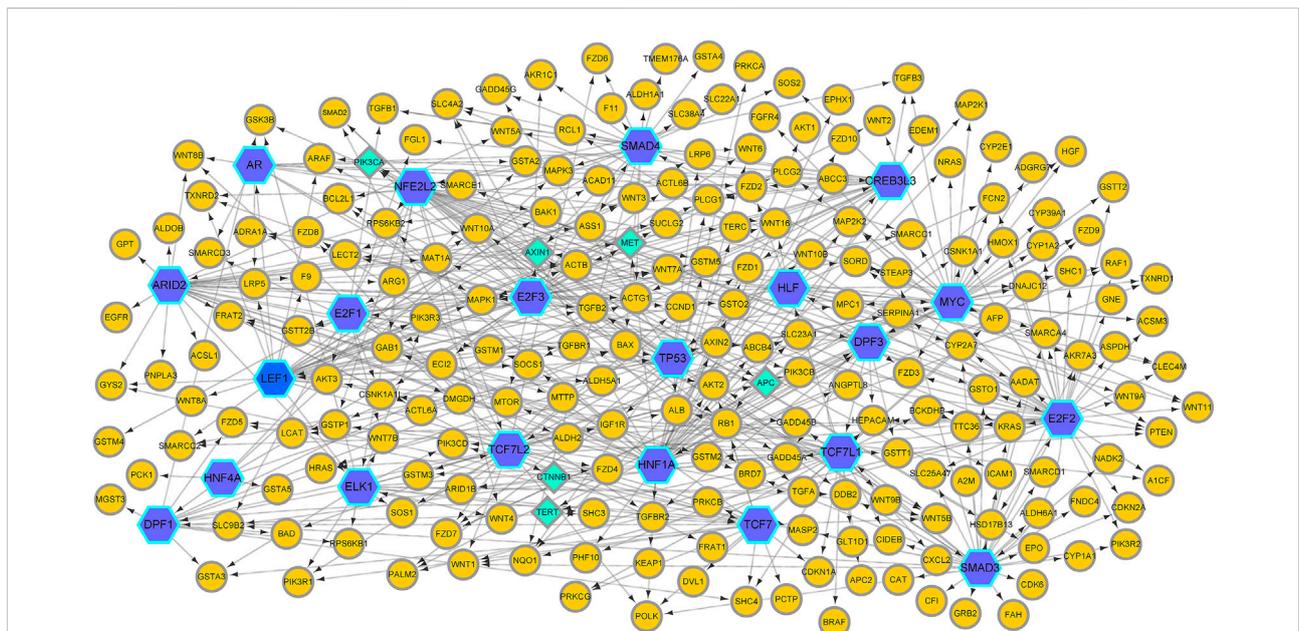


FIGURE 4 The HCC GRN reconstructed by GreyNet. In the network, the larger blue hexagon nodes represent TFs. The circle orange nodes represent the target genes. The diamond green nodes represent some elite disease genes in HCC.

TABLE 4 The enrichment of GO biological process and KEGG pathway in HCC GRN.

GO: Term	Term Name	Representative Gene	Corr. p-val
GO: 0016055	Wnt receptor signaling pathway	DVL1; WNT7B; DVL3; WNT8A; CCND1; WNT11; TCF7L2; FZD3; TCF7; LRP5; WNT9B; FZD1; WNT3; TCF7L1; FZD6; APC2; GSK3B; FRAT1; WNT4; FZD2; FZD4; APC; WNT5B; WNT7A; AXIN2; DVL2; CTNNB1; WNT1; CSNK1A1; FZD9; WNT10A; LEF1; FZD8; FRAT2; FZD5; WNT10B; AXIN1; WNT16; WNT5A	1.1E-54
GO: 0008283	Cell proliferation	WNT7B; CCND1; BAD; FZD3; SMAD4; FZD6; GSK3B; WNT4; BAK1; BCL2L1; CTNNB1; TGFB1; WNT1; FZD9; BAX; TERC; MAP2K1; HGF; WNT10B; MET; WNT5A	7.0E-17
GO: 0007169	Transmembrane receptor protein tyrosine kinase signaling pathway	AKT1; GRB2; IGF1R; PIK3R1; SOS1; SMARCC1; EGFR; PIK3R3; PTEN; FGFR4; GAB1; RAF1; HGF; MET	5.3E-10
GO: 0051252	Regulation of RNA metabolic process	ARID2; HNF1A; AKT2; TCF7L2; TCF7; SMAD3; ELK1; TCF7L1; SMAD4; SMARCC1; GSK3B; E2F3; WNT4; TGFB3; RB1; CDKN2A; WNT7A; CTNNB1; TGFB1; WNT1; E2F2; EPO; LEF1; MAP2K1; SMAD2; WNT10B; AXIN1; MET	7.6E-9
GO: 0043408	Regulation of MAPKKK cascade	AKT1; WNT7B; GRB2; AKT2; IGF1R; BRAF; APC; WNT7A; CTNNB1; GAB1; AXIN1; WNT5A	5.7E-9
GO: 0006357	Regulation of transcription from RNA polymerase II promoter	ARID2; HNF1A; AKT2; TCF7L2; SMAD3; TCF7L1; SMAD4; SMARCC1; TGFB3; RB1; CTNNB1; TGFB1; WNT1; EPO; LEF1; MAP2K1; SMAD2; WNT10B; AXIN1; MET	1.7E-7
GO: 0007265	Ras protein signal transduction	SHC3; GRB2; SOS1; CDKN1A; RB1; TP53; NRAS; MAPK3; RAF1; MAP2K1; CDKN2A; MAP2K2; KRAS; SHC2	3.3E-7
GO: 0043405	Regulation of MAP kinase activity	WNT7B; GSK3B; TGFB3; GAB1; MAP2K1; PRKCA; HGF; AXIN1; MET; WNT5A	1.0E-7
GO: 0006338	Chromatin remodeling	HNF1A; RB1; ACTL6A; ARID1A; SMARCC1; SMARCD1; ARID1B; SMARCC2; SMARCA2; SMARCB1	5.6E-6
GO: 0046328	Regulation of JNK cascade	AKT1; WNT7B; AKT2; WNT7A; GAB1; AXIN1; WNT5A	2.3E-5
GO: 0007179	Transforming growth factor beta receptor signaling pathway	SMAD4; SMAD3; TGFB1; TGFB1R1; TGFB3; TGFB2R2; TGFB2; SMAD2	9.1E-4
GO: 0006139	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	DPPF1; HNF1A; NRAS; TCF7L2; DDB2; TCF7; SMARCC3; SMAD3; ELK1; TCF7L1; SMAD4; ABCC3; SMARCC1; KEAP1; E2F3; ACTL6A; ACTL6B; ARID1A; RB1; CDKN2A; CTNNB1; E2F2; TERT; PRKCB; POLK; UGP2; LEF1; TERC; SMAD2; SMARCC2; SMARCB1; AXIN1; DPPF3	8.2E-4
GO: 0007264	Small GTPase mediated signal transduction	SHC3; GRB2; SOS1; CDKN1A; RB1; TP53; BRAF; NRAS; MAPK3; RAF1; MAP2K1; SOS2; CDKN2A; MAP2K2; HMOX1; KRAS; SHC2	4.6E-4
GO: 0043491	Protein kinase B signaling cascade	AKT1; RPS6KB1; AKT2; RPS6KB2	3.2E-4
GO: 0001889	Liver development	CCND1; HGF; SMAD3; AFP; CTNNB1; HNF1A; ALDH2; ALDOB; ASS1	2.0E-4

Reconstructing the causal gene regulatory network is a preliminary step to finding out the internal mechanism of the biological procedure and facilitating our understanding of the basic pathology of tumors and other diseases (Lesage et al., 2018; Femerling et al., 2022). However, current biological datasets generated by the facilities are usually accompanied by a low rate of signal-to-noise ratio. Simultaneously, GRN inference is an ill-posed problem with sparsity. A purely data-driven model will find it hard to accurately find real and key regulatory links (Santra, 2014). Fortunately, many databases (Liu et al., 2015; Fang et al., 2020; Zhang et al., 2020) have been established to provide adequate regulatory prior knowledge. Fusing prior knowledge in the model may be an anticipated solution to improve the quality of the GRN topology (Ideker et al., 2011; Abdelzaher et al., 2015; Dandan

et al., 2020). It is expected to investigate their effects on GreyNet in the future.

5 Conclusion

In this article, we proposed an interpretable machine-learning framework to infer the gene regulatory network from time-series expression data. We applied grey theory with the adaptive sliding window technique to model internal interactions in real regulatory procedures in the condition of limited information and knowledge. We further incorporated the Granger causality framework to search for causal regulations between genes. In DREAM4 *in silico* datasets, our model got competitive performances on AUROC and AUPRC compared with other state-of-the-art models. In the real HCC dataset,

GreyNet can find meaningful pathways in HCC development from the functional enrichment results of HCC GRN. In the future, we will provide an update to make our model applicable to single-cell RNA-seq data.

Data availability statement

The original contributions presented in the study are publicly available. The data and code can be found here: <https://github.com/zpliulab/GreyNet>.

Author contributions

GY performed the experiments, analyzed the data, and wrote the manuscript. Z-PL conceived and designed the experiments, and wrote the manuscript. Both authors read and approved the final manuscript.

Funding

National Key Research and Development Program of China under Grant Number 2020YFA0712402; National Natural Science Foundation of China (NSFC) under Grant Number 61973190; Shandong Provincial Key Research and Development Program (Major Scientific and Technological Innovation Project, 2019JZZY010423); Natural Science Foundation of Shandong Province of China under Grant Number ZR2020ZD25; the Innovation Method Fund of China (Ministry of Science and Technology of China) under Grant Number 2018IM020200; the Fundamental Research Funds for the Central Universities (No. 2022JC008); the Program of Qilu Young Scholars of

References

- Abdelzaker, A. F., Al-Musawi, A. F., Ghosh, P., Mayo, M. L., and Perkins, E. J. (2015). Transcriptional network growing models using motif-based preferential attachment. *Front. Bioeng. Biotechnol.* 3, 157. doi:10.3389/fbioe.2015.00157
- Anne-Claire, H., Fantine, M., Paola, V.-L., and Jean-Philippe, V. (2012). Tigress: Trustful inference of gene REgulation using stability selection. *BMC Syst. Biol.* 6, 145. doi:10.1186/1752-0509-6-145
- Arnold, A., Liu, Y., and Abe, N. (2007). "Temporal causal modeling with graphical granger methods," in *Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining* (New York, NY, USA: Association for Computing Machinery), 66–75. KDD '07. doi:10.1145/1281192.1281203
- Biswas, S., and Acharyya, S. (2018). A bi-objective rnn model to reconstruct gene regulatory network: A modified multi-objective simulated annealing approach. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 15, 2053–2059. doi:10.1109/TCBB.2017.2771360
- Chen, T., and Guestrin, C. (2016). "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. (New York, NY, USA: Association for Computing Machinery). doi:10.1145/2939672.2939785
- Cheng, L., Hou, Z.-G., Lin, Y., Tan, M., Zhang, W. C., and Wu, F.-X. (2011). Recurrent neural network for non-smooth convex optimization problems with application to the identification of genetic regulatory networks. *IEEE Trans. Neural Netw.* 22, 714–726. doi:10.1109/TNN.2011.2109735
- Dandan, C., Shun, G., Qingshan, J., and Lifei, C. (2020). PFBNet: A priori-fused boosting method for gene regulatory network inference. *BMC Bioinforma.* 21, 308. doi:10.1186/s12859-020-03639-7
- Deng, J. L. (1989). Introduction to grey system theory. *J. Grey Syst.* 1, 1–24.
- Dimri, M., and Satyanarayana, A. (2020). Molecular signaling pathways and therapeutic targets in hepatocellular carcinoma. *Cancers* 12, 491. doi:10.3390/cancers12020491
- Fabregat, I., and Caballero-Díaz, D. (2018). Transforming growth factor- β -induced cell plasticity in liver fibrosis and hepatocarcinogenesis. *Front. Oncol.* 8, 357. doi:10.3389/fonc.2018.00357
- Fang, L., Li, Y., Ma, L., Xu, Q., Tan, F., and Chen, G. (2020). GRNdb: Decoding the gene regulatory networks in diverse human and mouse conditions. *Nucleic Acids Res.* 49, D97–D103. doi:10.1093/nar/gkaa995

Shandong University. Publication costs are funded by NSFC. The funding bodies had no role in the design of the study, collection, the interpretation of data and in writing the manuscript.

Acknowledgments

Thanks are due to the members of our lab in Shandong University for their assistance in this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary Material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2022.954610/full#supplementary-material>

- Farzaneh, Z., Vosough, M., Agarwal, T., and Farzaneh, M. (2021). Critical signaling pathways governing hepatocellular carcinoma behavior; small molecule-based approaches. *Cancer Cell. Int.* 21, 208. doi:10.1186/s12935-021-01924-w
- Femerling, G., Gama-Castro, S., Lara, P., Ledezma-Tejeda, D., Tierrafría, V. H., Muñoz-Rascado, L., et al. (2022). Sensory systems and transcriptional regulation in *Escherichia coli*. *Front. Bioeng. Biotechnol.* 10, 823240. doi:10.3389/fbioe.2022.823240
- Finkle, J. D., Wu, J. J., and Bagheri, N. (2018). Windowed granger causal inference strategy improves discovery of gene regulatory networks. *Proc. Natl. Acad. Sci. U. S. A.* 115, 2252–2257. doi:10.1073/pnas.1710936115
- Freyre-González, J. A., Escorcía-Rodríguez, J. M., Gutiérrez-Mondragón, L. F., Martí-Vértiz, J., Torres-Franco, C. N., and Zorro-Aranda, A. (2022). System principles governing the organization, architecture, dynamics, and evolution of gene regulatory networks. *Front. Bioeng. Biotechnol.* 10, 888732. doi:10.3389/fbioe.2022.888732
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 29, 1189–1232. doi:10.1214/aos/1013203451
- Ghosh Roy, G., Geard, N., Verspoor, K., and He, S. (2020). PoLoBag: Polynomial Lasso Bagging for signed gene regulatory network inference from expression data. *Bioinformatics* 36, 5187–5193. doi:10.1093/bioinformatics/btaa651
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., and Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Comput. Surv.* 51, 1–42. doi:10.1145/3236009
- Haonan, F., Zheng, R., Wang, J., Wu, F., and Li, M. (2020). Nimce: A gene regulatory network inference approach based on multi time delays causal entropy. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 1, 1042–1049. doi:10.1109/TCBB.2020.3029846
- Hoerl, A., and Kennard, R. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12, 55–67. doi:10.1080/00401706.1970.10488634
- Hui, L., Zatloukal, K., Scheuch, H., Stepniak, E., and Wagner, E. F. (2008). Proliferation of human hcc cells and chemically induced mouse liver cancers requires jnk1-dependent p21 downregulation. *J. Clin. Invest.* 118, 3943–3953. doi:10.1172/JCI37156
- Huynh-Thu, V. A., and Sanguinetti, G. (2019). “Gene regulatory network inference: An introductory survey,” in *Gene regulatory networks: Methods and protocols*. Editors G. Sanguinetti and V. A. Huynh-Thu (New York: Springer New York), 1–23.
- Huynh-Thu, V. A. (2011). *Machine learning-based feature ranking: Statistical interpretation and gene network inference* (Liège, Belgium: University of Liège, Faculty of Applied Sciences, Department of Electrical Engineering and Computer Science). Ph.D. thesis.
- Huynh-Thu, V. A., and Sanguinetti, G. (2015). Combining tree-based and dynamical systems for the inference of gene regulatory networks. *Bioinformatics* 31, 1614–1622. doi:10.1093/bioinformatics/btu863
- Ideker, T., Dutkowski, J., and Hood, L. (2011). Boosting signal-to-noise in complex biology: Prior knowledge is power. *Cell* 144, 860–863. doi:10.1016/j.cell.2011.03.007
- Jansen, C., Paraiso, K. D., Zhou, J. J., Blitz, I. L., Fish, M. B., Charney, R. M., et al. (2022). Uncovering the mesoderm gene regulatory network through multi-omic data integration. *Cell. Rep.* 38, 110364. doi:10.1016/j.celrep.2022.110364
- Kuo, Y., Yang, T., and Huang, G.-W. (2008). The use of grey relational analysis in solving multiple attribute decision-making problems. *Comput. Industrial Eng.* 55, 80–93. doi:10.1016/j.cie.2007.12.002
- Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., et al. (2018). The human transcription factors. *Cell* 172, 650–665. doi:10.1016/j.cell.2018.01.029
- Le, N. Q. K., and Ho, Q.-T. (2022). Deep transformers and convolutional neural network in identifying dna n6-methyladenine sites in cross-species genomes. *Methods* 204, 199–206. doi:10.1016/j.jymeth.2021.12.004
- Leng, S., Xu, Z., and Ma, H. (2019). Reconstructing directional causal networks with random forest: Causality meeting machine learning. *Chaos* 29, 093130. doi:10.1063/1.5120778
- Lesage, R., Kerkhofs, J., and Geris, L. (2018). Computational modeling and reverse engineering to reveal dominant regulatory interactions controlling osteochondral differentiation: Potential for regenerative medicine. *Front. Bioeng. Biotechnol.* 6, 165. doi:10.3389/fbioe.2018.00165
- Li, L., Shangquan, W., Deng, Y., Mao, J., Pan, J., Wei, N., et al. (2020). A causal inference model based on random forests to identify the effect of soil moisture on precipitation. *J. Hydrometeorol.* 21, 1115–1131. doi:10.1175/JHM-D-19-0209.1
- Liao, J., Huang, Y., Wang, Q., Chen, S., Zhang, C., Wang, D., et al. (2022). Gene regulatory network from cranial neural crest cells to osteoblast differentiation and calvarial bone development. *Cell. Mol. Life Sci.* 79, 158. doi:10.1007/s00018-022-04208-2
- Liu, Z.-P., Wu, C., Miao, H., and Wu, H. (2015). RegNetwork: An integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. *Database* 2015, bav095. doi:10.1093/database/bav095
- Liu, Z.-P., Wu, H., Zhu, J., and Miao, H. (2014). Systematic identification of transcriptional and post-transcriptional regulations in human respiratory epithelial cells during influenza a virus infection. *BMC Bioinforma.* 15, 336. doi:10.1186/1471-2105-15-336
- Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular carcinoma. *Nat. Rev. Dis. Prim.* 7, 6. doi:10.1038/s41572-020-00240-3
- Lu, J., Dumitrascu, B., McDowell, I. C., Jo, B., Barrera, A., Hong, L. K., et al. (2021). Causal network inference from gene transcriptional time-series response to glucocorticoids. *PLoS Comput. Biol.* 17, e1008223. doi:10.1371/journal.pcbi.1008223
- Marbach, D., Prill, R. J., Schaffter, T., Mattiussi, C., Floreano, D., and Stolovitzky, G. (2010). Revealing strengths and weaknesses of methods for gene network inference. *Proc. Natl. Acad. Sci. U. S. A.* 107, 6286–6291. doi:10.1073/pnas.0913357107
- Marbach, D., Schaffter, T., Mattiussi, C., and Floreano, D. (2009). Generating realistic *in silico* gene networks for performance assessment of reverse engineering methods. *J. Comput. Biol.* 16, 229–239. doi:10.1089/cmb.2008.09tt
- Ming, S., Sheng, T., Xin-Ping, X., Ao, L., Wulin, Y., Tao, Z., et al. (2020). Globally learning gene regulatory networks based on hidden atomic regulators from transcriptomic big data. *BMC Genomics* 21, 711. doi:10.1186/s12864-020-07079-8
- Nault, J. C., Calderaro, J., Di Tommaso, L., Balabaud, C., Zafrani, E. S., Bioulac-Sage, P., et al. (2014). Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. *Hepatology* 60, 1983–1992. doi:10.1002/hep.27372
- Niu, Z. S., Niu, X. J., and Wang, W. H. (2016). Genetic alterations in hepatocellular carcinoma: An update. *World J. Gastroenterol.* 22, 9069–9095. doi:10.3748/wjg.v22.i41.9069
- Papadimitriou, S., Sun, J., and Yu, P. S. (2006). “Local correlation tracking in time series,” in Sixth International Conference on Data Mining (ICDM’06), Hong Kong, China, 18–22 December 2006, 456–465. doi:10.1109/ICDM.2006.99
- Phan, N., and Rosemary, B. (2018). Time-lagged ordered lasso for network inference. *BMC Bioinforma.* 19, 545. doi:10.1186/s12859-018-2558-7
- Polakis, P. (2012). Wnt signaling in cancer. *Cold Spring Harb. Perspect. Biol.* 4, a008052. doi:10.1101/cshperspect.a008052
- Sallehuddin, R., Shamsuddin, S. M. H., and Hashim, S. Z. M. (2008). “Application of grey relational analysis for multivariate time series,” in 2008 Eighth International Conference on Intelligent Systems Design and Applications, Kaohsiung, Taiwan, 26–28 November 2008, 432–437. doi:10.1109/ISDA.2008.181
- Santra, T. (2014). A bayesian framework that integrates heterogeneous data for inferring gene regulatory networks. *Front. Bioeng. Biotechnol.* 2, 13. doi:10.3389/fbioe.2014.00013
- Schulze, K., Imbeaud, S., Letouze, E., Alexandrov, L. B., Calderaro, J., Rebouissou, S., et al. (2015). Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat. Genet.* 47, 505–511. doi:10.1038/ng.3252
- Sungjoon, P., Jung, M. K., Wonho, S., Sung, W. H., Minji, J., Hyun, J., et al. (2018). Btnet: Boosted tree based gene regulatory network inference algorithm using time-course measurement data. *BMC Syst. Biol.* 12, 20. doi:10.1186/s12918-018-0547-0
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Methodol.* 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x
- Tng, S. S., Le, N. Q. K., Yeh, H.-Y., and Chua, M. C. H. (2021). Improved prediction model of protein lysine crotonylation sites using bidirectional recurrent neural networks. *J. Proteome Res.* 21, 265–273. doi:10.1021/acs.jproteome.1c00848
- Vân, H.-T., AnhAlexandre, I., Louis, W., and Pierre, G. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 5, e12776. doi:10.1371/journal.pone.0012776
- Villanueva, A. (2019). Hepatocellular carcinoma. *N. Engl. J. Med. Overseas. Ed.* 380, 1450–1462. doi:10.1056/nejmra1713263
- Waisberg, J., and Saba, G. T. (2015). Wnt/ β -catenin pathway signaling in human hepatocellular carcinoma. *World J. Hepatol.* 7 (26), 2631–2635. doi:10.4254/wjh.v7.i26.2631
- Wang, J., Huang, Q., Liu, Z.-P., Wang, Y., Wu, L.-Y., Chen, L., et al. (2011). Noa: A novel network ontology analysis method. *Nucleic Acids Res.* 39, e87. doi:10.1093/nar/gkr251
- Xiang, C., Min, L., Ruiqing, Z., Fang-Xiang, W., and Jianxin, W. (2019). D3GRN: A data driven dynamic network construction method to infer gene regulatory networks. *BMC Genomics* 20, 929. doi:10.1186/s12864-019-6298-5

- Yang, H. D., Kim, H. S., Kim, S. Y., Na, M. J., Yang, G., Eun, J. W., et al. (2019). Hdac6 suppresses let-7i-5p to elicit tsp1/cd47-mediated anti-tumorigenesis and phagocytosis of hepatocellular carcinoma. *Hepatology* 70, 1262–1279. doi:10.1002/hep.30657
- Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A global view of hepatocellular carcinoma: Trends, risk, prevention and management. *Nat. Rev. Gastroenterol. Hepatol.* 16, 589–604. doi:10.1038/s41575-019-0186-y
- Yuansheng, H., Lei, S., and Hui, L. (2019). Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. *J. Clean. Prod.* 209, 415–423. doi:10.1016/j.jclepro.2018.10.128
- Zhan, T., Rindtorff, N., and Boutros, M. (2017). Wnt signaling in cancer. *Oncogene* 36, 1461–1473. PMID: 27617575. doi:10.1038/onc.2016.304
- Zhang, Q., Liu, W., Zhang, H.-M., Xie, G.-Y., Miao, Y.-R., Xia, M., et al. (2020). hTFtarget: a comprehensive database for regulations of human transcription factors and their targets. *Genomics, proteomics Bioinforma.* 18, 120–128. doi:10.1016/j.gpb.2019.09.006
- Zhang, Y., Chang, X., and Liu, X. (2021). Inference of gene regulatory networks using pseudo-time series data. *Bioinformatics* 37, btab099–2431. doi:10.1093/bioinformatics/btab099
- Zheng, R., Li, M., Chen, X., Wu, F.-X., Pan, Y., and Wang, J. (2018). BiXGBoost: A scalable, flexible boosting-based method for reconstructing gene regulatory networks. *Bioinformatics* 35, 1893–1900. doi:10.1093/bioinformatics/bty908
- Zucman-Rossi, J., Villanueva, A., Nault, J. C., and Llovet, J. M. (2015). Genetic landscape and biomarkers of hepatocellular carcinoma. *Gastroenterology* 149, 1226–1239.e4. e4. doi:10.1053/j.gastro.2015.05.061