Dynamic Model Inference of Gene Regulatory Network based on Hybrid Parallel Genetic Algorithm and Threshold Qualification Method

Xiaomei Ding, Huaibao Ding*, Fei Zhou

School of Computer Engineering
Anhui Wenda University of Information Engineering
Hefei 231201, Anhui, China
E-mails: dingxiaomei@mail.wenda.edu.cn,
dinghuaibao@mail.wenda.edu.cn
zhoufei@mail.wenda.edu.cn

*Corresponding author

Received: February 05, 2023 **Accepted: February 13, 2024**

Published: March 31, 2024

Abstract: *Gene regulation is the process by which various substances in cells regulate the behaviour of gene expression, thereby controlling almost all cellular activities. Therefore, studying gene regulation not only helps to uncover the internal laws governing life processes but also plays a crucial role in predicting, diagnosing, treating, and designing drugs for genetic diseases. By utilizing multi-source biological information such as gene expression profiles, transcription factor information, and protein interaction data, a network model can be developed to depict the regulatory relationships between genes, facilitating further research. To address the limitations of traditional gene regulatory network construction methods, a novel dynamic model has been created by combining hybrid genetics and threshold restriction. This model comprises two parts: solution space reduction and parameter fitting. During solution space reduction, singular value decomposition is employed to define a mathematically feasible gene regulatory network, reducing unnecessary calculations. Subsequently, the control genes of each gene are constrained within a certain range using threshold limitation, enhancing computational efficiency while adhering to bioinformatics principles. In the parameter fitting phase, parallel genetic algorithms are utilized to expediently optimize the entire solution space. The mountain climbing method is then applied to solve problems meticulously within a limited scope, improving calculation accuracy. In this study, this approach was applied to establish genetic regulatory systems for complex skin melanoma and type 2 diabetes. Through comparison with actual networks, the validity of the method was confirmed. Compared to traditional genetic and particle swarm optimization methods, the effectiveness of the proposed method was verified. This paper models the intricate mechanism of gene regulation and elucidates the regulatory process involving genes, proteins, and small biological molecules in greater detail than other models, aligning more closely with intracellular dynamics laws.*

Keywords: *Hybrid parallel genetic algorithm, Threshold limit method, Singular value decomposition, Gene regulatory network.*

Introduction

With the advancement and widespread use of second and third-generation DNA and protein sequencing technologies, as well as improvements in mRNA (a class of single-stranded RNA that is transcribed from a strand of DNA as a template and carries genetic information that guides protein synthesis) hybridization microarray data determination techniques, researchers have enhanced their ability to study life at the molecular and genomic levels [18]. The emergence of big data, enabled by gene-chip data, has made it possible to compute systems

biology in ways previously unattainable [2]. Large-scale gene regulatory networks (GRNs, a combination of transcription factors and their target genes that determine and maintain cell type or stable state) are constructed using various biological data, particularly gene expression matrix data. Complex human diseases such as cancer and coronary heart disease are caused by multiple gene abnormalities [5, 19]. While each gene may have little effect on disease occurrence and development, the non-linear accumulation and combination of small gene effects can exacerbate disease formation. Consequently, the construction and analysis algorithm of gene regulatory networks has become a focal point and challenge in bioinformatics research, with profound implications for human understanding of oneself and complex disease prevention and treatment [6]. Studying the topological structure of GRNs and the relationship between genes has become an extremely challenging task with significant theoretical and practical implications.

The construction algorithm of GRNs primarily involves utilizing various known biological data, particularly gene expression data, through a specific network construction model and corresponding algorithm to infer the interactions between genes or gene products. This process ultimately results in the formation of a network. The biological macromolecules involved in these networks encompass genes, mRNA, proteins, and more. Various models are employed, including Boolean network models, Bayesian network models, neural network models, and differential equation models (DEM), among others [8]. Each type of network model has its own application domain, advantages, and disadvantages. Boolean network models focus on basic global networks rather than quantitative biochemical models and are suitable for large but imprecise networks. In contrast, Bayesian network models and neural network models are well-suited for constructing statistically-based medium-sized networks. DEM regards the regulatory relationship between genes as differential equations, overcoming the uncertainty factor inherent in statistical models [11]. This approach is particularly suitable for constructing small and accurate networks. However, with the continuous improvement of hardware computing power, DEM has gradually become the mainstream method and can now build larger-scale networks, even genome-scale structures. In this paper, the mathematical model utilized is the differential equation model.

A network graph with N nodes has $N2$ edges. Gene networks with many nodes face challenges [17]. However, GRNs have sparse edges. We used two control methods: eliminating weak edges with threshold control and reducing the dimension of differential equation datasets using singular value decomposition (SVD) [3]. This aligns with sparse matrices and reduces calculations. The model and dimensionality affect algorithm complexity [10]. Parameter fitting can use maximum likelihood or expectation maximization, but large datasets require heuristic search. Various algorithms exist, like the hybrid genetic algorithm (GA) - firefly algorithm (FA) (GFA) [18], simulated annealing (SA), and particle swarm optimization (PSO) [10]. GFA excels at multi-peak optimization, but SA has long runtimes. The hybrid parallel GA offers improved efficiency [12].

In GA research, the parallel gene method is used for fast global solution finding, followed by the climbing method for precise search [15]. The hybrid parallel GA (HGA) combines the broad optimization of the GA with the precise optimization of the climbing method.

Overview of gene regulatory networks

Background on gene regulatory processes

A living system is a complex system made of genes, RNA, proteins, and their interactions [21]. The goal of molecular biology is to understand this system to control cell behaviour.

Genome sequencing projects have deepened our understanding of genes, but genes alone do not explain cell behaviour. Cells with the same DNA show different properties and functions in different tissues [13, 14]. Gene expression is not constant and is influenced by cell environment factors. Understanding these factors' effects on gene behaviour could lead to better predictions, diagnoses, and treatments of gene-related diseases [25, 29].

The study of the mechanism belongs to bioinformatics, which relies on rigorous reasoning and quantitative calculation to infer rules and mechanisms from phenomena and states. Although promising achievements have been made, the problem remains complex and mysterious [26].

Due to the development of next-generation sequencing and high-volume gene expression data, along with various biological test data sources (including transcription factor info, protein interaction info, protein complex info, etc.), it is possible to construct more detailed, multi-level gene regulation dynamic models [24]. This allows for the exploration of cell interactions among various substances (DNA, RNA, proteins, etc.), offering a deeper understanding of regulatory reaction processes. However, while many promising results have been achieved integrating multi-source data into network construction, most current networks only roughly reflect gene-to-gene regulatory relationships [23, 28]. The full utilization of multi-source data still needs further development to fully explore its information content.

Basic principles of gene regulatory networks

The GRN essence is the mathematical representation of gene expression, which involves the synthesis of active protein molecules through chemical reactions following transcription and translation of DNA's stored genetic information [27]. Transcription, the core step of gene expression, copies single-stranded RNA with the same sequence as DNA. Translation uses mRNA as a template and tRNA (RNA composed of 76 to 90 nucleotides, and its 3'end can attach specific types of amino acids under the catalysis of aminoacyl tRNA synthetase) as a carrier, under the influence of enzymes and cofactors, to express genetic information and often results in the synthesis of active proteins. Translation is much more complex than transcription in cell life activities [7].

Gene expression occurs in multiple stages, each controlled by other genes. Microarray gene data has made the translation stage of gene expression a hotspot, revealing more hidden biological info. Gene expression affects other genes or itself, creating the cell's life activities through gene interactions [9, 16].

There are several clustering methods like *k*-mean clustering, spectral clustering, hierarchical clustering, artificial neural networks, etc. used to segment large gene regulatory networks and build models for the segmented networks [1, 4, 22]. These methods are commonly used for dimensionality reduction of gene regulatory data. The algorithm's performance depends on the situation and needs to be theoretically discussed. Currently, the main method is to use the "reverse technique" to infer possible hidden relationships in the network topology through gene expression data. There are few successful cases of mining hidden information in the original gene network through experiments, and this technique has certain requirements for parameter estimation and model selection. The appropriate algorithm has a significant impact on network inference results [20].

Related algorithm models

SVD determines the solution space

Mathematically, DEM can be described as the derivative of the gene *i*'s expression concentration value at time *t*, which has differential equation features as per Eq. (1):

$$\frac{\partial x_i(t)}{\partial t} = \sum_{j=1}^{n} \omega_{ij} x_j(t) + b_i, \quad (1)$$

where ω_{ij} represents the effect of gene *j* on *i* in the network, while the owner value is the weighting matrix $\boldsymbol{W}$; $\boldsymbol{W}$ is $N \times N$ matrix, N refers to the number of genetic nodes in the genetic network; $\omega_{ij} \neq \omega_{ji}$ because of the directed network; $x_i(t)$ represents the expression of gene *i* at time *t*; b_i represents the background level of gene *i*, that is, the expression of gene *i* itself, which is not influenced by external factors.

Eq. (1) is written as a matrix, as shown in Eq. (2):

$$\frac{d\boldsymbol{X}_{N\times T}}{dt} = \boldsymbol{W}_{N\times N}\boldsymbol{X}_{N\times T} + \boldsymbol{B}_{N\times T}. \quad (2)$$

The singular value decomposition of the expression value matrix $\boldsymbol{X}_{N\times T}$ in Eq. (2) under the definition of the generalized inverse matrix is carried out, and Eq. (3) is obtained:

$$\boldsymbol{X}_{N\times T} = \boldsymbol{U}_{N\times T}\boldsymbol{A}_{T\times T}(\boldsymbol{V}^T)_{T\times T}, \quad (3)$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are both orthogonal matrices that satisfy the constraints of orthogonal matrices, as defined in Eq. (4):

$$\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{U} = \boldsymbol{V}^{-T}\boldsymbol{V} = \boldsymbol{V}^{-1}\boldsymbol{V} = \boldsymbol{E}, \quad (4)$$

where $\boldsymbol{E}$ is an identity matrix.

Eq. (4) implies that $\boldsymbol{U}$ and $\boldsymbol{V}$'s inverse matrices are equal to their respective transposes. After deducing this formula, Eq. (5) shows the matrix form of weight matrix $\boldsymbol{W}$:

$$\boldsymbol{W}_{N\times T} = (d\boldsymbol{X}_{N\times T}/\,dt - \boldsymbol{B}_{N\times T})\boldsymbol{V}^{-T}\boldsymbol{A}^{-1}\boldsymbol{U}^{-1}, \quad (5)$$

where $\boldsymbol{W}$ is a specific solution of a network from the singular value decomposition of the differential equation. This matrix may not be the optimal solution for the whole space, but its general optimal solution form can be deduced using linear algebra rules. Eq. (5) further defines the range of $\boldsymbol{W}$, as per Eq. (6):

$$\boldsymbol{W}_{N\times T} = (d\boldsymbol{X}_{N\times T}/\,dt - \boldsymbol{B}_{N\times T})\boldsymbol{V}^{-T}\boldsymbol{A}^{-1}\boldsymbol{U}^{-1} + \boldsymbol{C}\boldsymbol{U}^{-1}, \quad (6)$$

where $\boldsymbol{C}$ is an arbitrary constant independent.

Eq. (6) represents the general solution structure under the condition that Eq. (5) is a particular solution, i.e., the whole solution space. Using this formula can reduce the solution space of differential equations and minimize unnecessary calculations [30].

Deleting redundant edges by threshold restriction method

SVD reduces understanding scope but still involves many unnecessary operations. Difference equations define how genetic factors vary, but gene regulation networks are sparse matrices

where gene penetration follows a power law distribution. This means genes are only affected by a limited number of genes (usually 1-3), not all of them. The threshold method limits the number and influence of genes, reducing unnecessary calculations and facilitating large-scale network construction.

The threshold limitation method is mainly divided into two steps. The first step is the deletion operation under threshold control, which mainly refers to the direct deletion of the weight ω_{ij} less than the threshold value. After deletion, it must meet the singular value decomposition limit shown in Eq. (6). The second step is the weight ω_{ij} larger than the threshold value. The roulette strategy is adopted to select the preset number of control nodes c, after which the singular value decomposition condition shown in Eq. (6) must be satisfied. The specific steps are as follows:

Step 1. Set the threshold (frame). In this paper, 0.3 is selected as the test object. In this experiment, up to 5 controlled gene ***C*** could be obtained. In the weighted matrix ***W***, one element is selected so that ω_{ij} is smaller than *s*.

Step 2. Directly delete the weight ω_{ij}, and put the matrix ***W*** into Eq. (6) after deleting it. If the corresponding ***C*** value can be obtained, it indicates that ***W*** is still reasonable after deleting ω_{ij}, otherwise, cancel the deletion.

Step 3. *i* and *j* pairs are selected again so that ω_{ij} is less than *s*, and the deletion decision as shown in Step 2 is carried out again. When all ω_{ij} decisions are finished, the algorithm enters Step 4.

Step 4. For each gene *i*, a random number $m \in [1, c]$ is generated, and weights larger than s or unsuccessful deletion. A roulette strategy selects *m* genes (including gene *i*) as control genes, deleting other weights.

Step 5. After *m* control genes are chosen for gene *i* in Step 4, ***W*** is used in Eq. (6). If a corresponding ***C*** value is obtainable, it shows gene *i*'s operation is reasonable. Otherwise, gene *i*'s roulette strategy is re-selected until Eq. (6) is met or max replacements are reached.

Step 6. Select the gene controlled by the next threshold, and conduct the judgment as shown in Steps 4 and 5 again until all genes are selected as control genes, and the algorithm ends.

The above algorithm flow is shown in Fig. 1.

Hybrid parallel genetic algorithm to optimize GRN parameters

SVD narrows the space and removes redundant control genes with threshold restrictions. The remaining ***W*** weight matrix is solved by an HGA. The global solution space uses a parallel GA for fast search, and after locking the general position, the mountain climbing method is used for a detailed search to obtain the optimal weight matrix ***W***. To evaluate the weighted matrix ***W*** in DEM, a suitable range of applications is proposed, along with evaluation criteria to assess its advantages and disadvantages. In this paper, the least square method is used as the evaluation function, expressed by Eq. (7):

$$c' = \sum_i^n \sum_k^t r_{ik}^2, \tag{7}$$

where c' is the difference between the observation and the formula, and r_{ik} is the residual, which can be expressed by Eq. (8):

$$r_{ik} = x_i'(t) - \sum_j^n \omega_{ij} x_j'(t) - b_i, \tag{8}$$

where $x_i'(t)$ represents the derivative of the value of gene i expression at time point t, that is, the rate of change, as shown in Eq. (9):

$$x_i'(t) = x_i(t) - x_i(t-1). \tag{9}$$

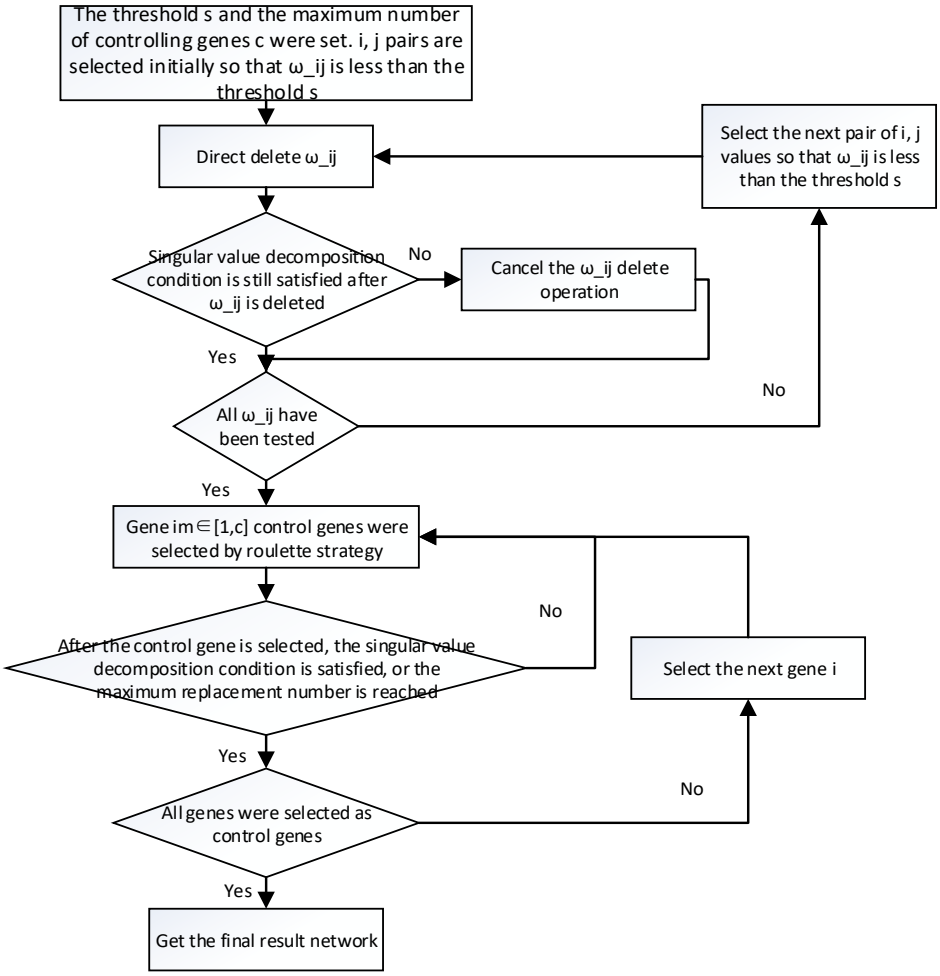

Fig. 1 Flow chart of threshold limitation method

Coding is an important factor affecting the performance of GA. A new HGA is proposed to solve the minimum weight matrix $\boldsymbol{W}$ in Eq. (7). Taking the whole $\boldsymbol{W}$ as a chromosome code is an intuitive and operable choice. According to the existing bioinformatics knowledge, the ω_{ij} control intensity of any weight matrix element should not be too large, and its value range is defined as [-5, 5].

The whole process of the HGA is shown as follows:

Step 1. Initialization: the population number of the parallel GA is $\boldsymbol{P}$, and each independent population initializes N chromosomes. Each chromosome is the weight matrix $\boldsymbol{W}$ as shown in the code, and N is the population size. Set the crossover probability p_c and mutation probability p_m.

Step 2. Reduce solution space: decompose weight matrix $\boldsymbol{W}$ using singular value. If Eq. (6)'s $\boldsymbol{C}$ value is obtainable, $\boldsymbol{W}$ is valid; otherwise, re-initialize or recalculate $\boldsymbol{W}$.

Step 3. Threshold limitation: the weight matrix is chosen through Fig. 1's process to determine the number of control genes for each gene within the range [1, *c*] (*c*'s value is in *Section 1.2*).

Step 4. Selection, crossover and variation: in each population of the parallel GA, chromosomes with small fitness values (see Eq. (7)) are selected according to the roulette strategy to enter the crossover operation. The smaller the fitness value, the higher the probability of being selected. With the selection probability p_c, the chromosomes entered into the crossover operation through the selection operation were crossed at a single point by two individuals. The mutation probability p_m was used to mutate a certain locus of the chromosome after crossing, and the weight ω_{ij} was changed.

Step 5. Exchange: after GA operation (Step 4), *P* populations tend to reach local optima. The individual with the smallest fitness from each population is selected and swapped with the largest population's best individual. This preserves species diversity, and the GA operation (Step 4) is repeated.

Step 6. End judgment of GA: when the optimal value difference among populations is small or max iterations are reached, the GA part ends.

Step 7. Hill-climbing method: following parallel GA operation, the whole weight matrix $\boldsymbol{W}$ is optimized using hill-climbing, searching its neighbourhood for the optimal solution. Sequentially adjusting each non-zero weight to find its dimension's extreme value, until the whole $\boldsymbol{W}$'s fitness function reaches a minimum or max iterations are reached.

Step 8. Legitimacy test: the optimal solution is plugged into Eq. (6) to satisfy SVD conditions. If not, a sub-optimal solution is chosen. If Eq. (6) is not met, the mixed parallel GA (Steps 4-7) is repeated until satisfied or max iterations are reached.

The above HGA process is shown in Fig. 2.

Screening methods for regulatory genes

In small gene regulatory networks, selecting one gene as the target gene allows all other genes to be made regulatory genes. Parameters of each gene can be solved quickly using mixed parallel algorithms and threshold limiting methods. However, in large gene regulatory networks, directly solving parameters takes a long time and the parameters are not unique. Therefore, it is necessary to screen out possible regulatory genes from all genes to reduce their number.

After dividing the large GRN into multiple subnetworks using the algorithm, genes within each subnetwork are interconnected, and subnetworks are connected through representative genes. When a gene within the same subnetwork is unrepresentative, the regulated gene can be screened out. When a gene is typical, genetic information from the same subnetwork must be selected, and these representative heritages are adjusted. However, due to the large number of genes, there are still many genes in each subnetwork, making effective parameter estimation challenging. Therefore, it is necessary to consider gene relationships and select regulatory genes with strong relationships to target genes from all genes and representative genes in the subnetwork. While correlation indices measure genetic connection strength, they measure straight-line connections. This system's nonlinear relationship between target and regulatory genes makes selecting regulatory genes inappropriate.

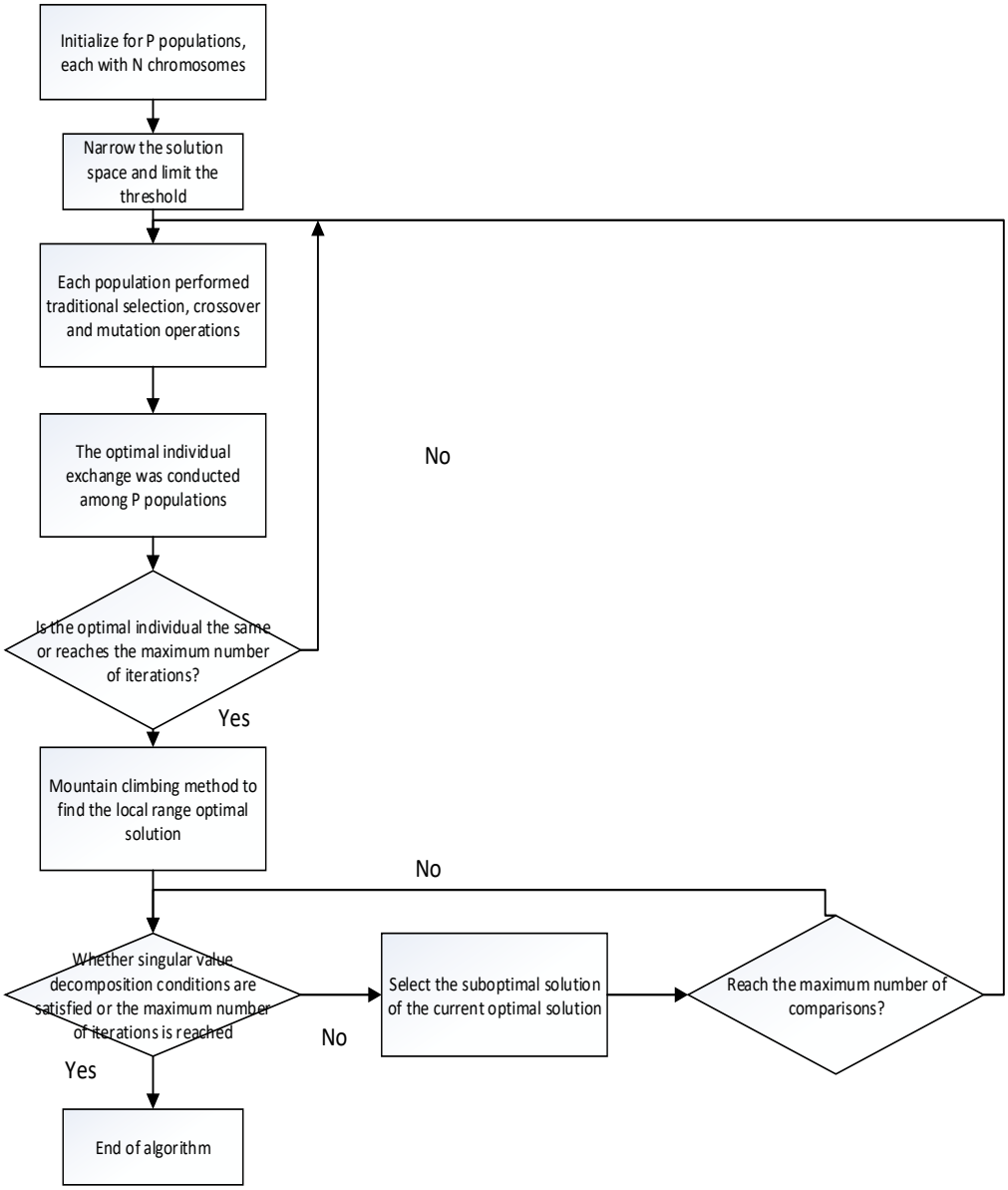

Fig. 2 Flowchart of the HGA

Given that gene expression profiles within the same subnetwork are similar, genes with expression data closer to the target gene are more likely to regulate it. Therefore, in this paper, we use gene expression data distance to screen regulatory genes. When the target gene is atypical, neighbouring genes in the same subnetwork are selected as candidates. When the target gene is typical, neighbouring genes and representative genes are selected as candidates. The HGA is then used to solve the target and candidate genes. Genes with an element value less than the threshold s are eliminated and the remaining genes are used as regulatory genes for the target gene.

Experimental test and comparison

The proposed algorithm is designed to build gene regulatory networks for complex human diseases, specifically melanoma and diabetes. It utilizes three gene databases: OMIM for disease-related gene classes, GO for genes with similar annotations, and GEO for gene expression data, for which it is assumed:

1) OMIM: Genes associated with the same disease interact.
2) GO: Genes with related or similar annotations interact.
3) GEO: Genes in the microarray equation modal correlation interact.

The algorithm uses data fusion: First, OMIM finds disease-related genes, and then GO filters them by annotation similarity. GEO provides expression values for the remaining genes, which are used to construct the gene regulatory network. The computational efficiency of genetic control networks is compared to two other methods without thresholding or singular value decomposition. GA and PSO are used as optimization algorithms, and their results are compared to the proposed algorithm's results.

The evaluation criteria of the algorithm can be calculated by using Bansal's positive predictive value (*PPV*) and sensitivity (*Se*) functions, as shown in Eq. (10):

$$\begin{cases} PPV = \frac{TP}{TP+FP} \\ Se = \frac{TP}{TP+FN} \end{cases}, \tag{10}$$

where *TP* represents true positives, where both the calculated and molecular biological values are true; *FP* stands for false positives, where the algorithm falsely identifies true molecular biological values; *FN* is false negatives, where the algorithm correctly identifies false molecular biological values.

Eq. (10) shows that this algorithm aims to minimize *FP* and *FN*. Lower values of these metrics lead to higher *PPV* and *Se*, indicating greater accuracy. The validation of *TP*, *FP*, and *FN* was done using the KEGG database, which contains known results from most human GRN experiments.

Melanoma gene experiment

Melanoma, a skin tumour formed by melanocytes, often appears and changes in a distinctive pattern over weeks or years. By analyzing OMIM genes, we identified 117 associated melanoma genes. Using the GO database, we narrowed it down to 64 relevant genes. The GDS4989 melanoma gene sequences from the GEO database matched 64 probe expressions. In cytoscape plotting, the completed genetic control network was designed using a circular layout, the results of which are shown in Fig. 3.

The mapk8 and mapk14 series have shown positive effects. These substances are exported in large quantities and play a key role in tumour formation and development. However, our research also indicates that NPHB2 and NOS2 have high carcinogenicity, which warrants further biological testing.

Table 1 compares our method with traditional GA and PSO. All three methods are implemented in Matlab. When considering *PPV* and *Se*, calculation speed is also crucial. By combining GA and PSO techniques using the genetic optimization method described in this paper, we have established a new genetic control network for melanoma.

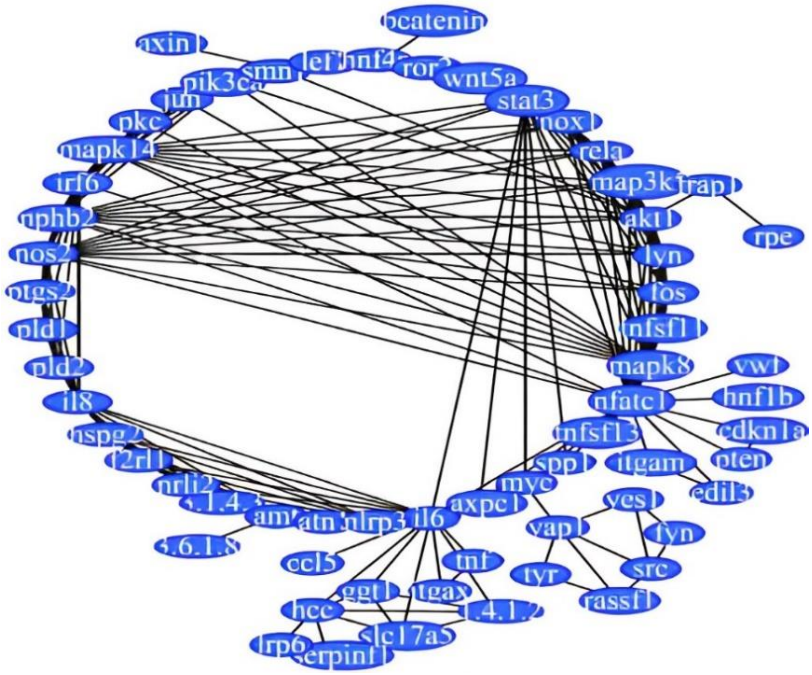

Fig. 3 The GRN for melanoma was calculated using GA

Table 1. Comparison table of calculation results of GDS4989 by three algorithms

Algorithm	*PPV*	*Se*	Operation time, s
HGA	0.588	0.799	4.54
GA	0.358	0.585	4.99
PSO	0.362	0.607	4.58

In both *PPV* and *Se* accuracy measures, the HGA demonstrates clear advantages. It utilizes singularity analysis and thresholding to eliminate uncertain values. This approach is followed by an HGA, incorporating a climbing method to enhance search speed and reduce local optimization risk. Computational speed: The proposed method offers only marginally improved computational speed compared to GA and PSO. This is because singular value decomposition and thresholding consume time, but reduce computation, balancing out the extra time from parallel genetic algorithm and climbing method compared to traditional genetic algorithm.

In conclusion, the proposed algorithm offers improved accuracy in skin melanoma calculations compared to GA and PSO, with minimal differences between GA and PSO. Its computational time is similar to that of GA and PSO.

Type 2 diabetes gene experiment

This paper's focus is on the establishment of a control network for the pathogenesis of type 2 glycosuria. Using OMIM analysis, 31 DM genes related to type 2 were identified. GO analysis was conducted on these genes, resulting in the selection of 23 annotated genes. The cytoscape mapping program was then used to analyze genes related to GDS3681 in the GEO database through the orthogonal arrangement. The result of the display is shown in Fig. 4. It can be clearly seen from this figure that the foxo1 factor is the most important therapeutic factor of type 2 diabetes and the key protein of type 2 diabetes, which is consistent with previous studies.

The proposed HGA algorithm was compared with GA and PSO algorithms, and *PPV*, *Se* and running time were compared for melanoma. The statistical results are shown in Table 2.

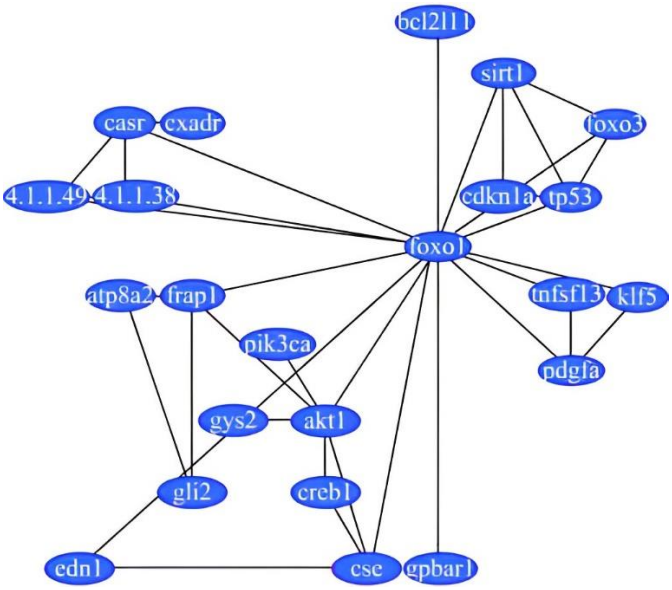

Fig. 4 Gene regulatory network of type 2 diabetes calculated by HGA

Table 2. Comparison table of calculation results of GDS3681 by three algorithms

Algorithm	*PPV*	*Se*	Operation time, s
HGA	0.419	0.814	3.88
GA	0.232	0.635	4.02
PSO	0.225	0.599	3.96

Table 2's statistical results show that the proposed algorithm significantly improves *PPV* and *Se* accuracy measures. GA is marginally more effective than PSO. The HGA performs similarly to PSO, while GA takes a slightly longer calculation time.

In summary, the proposed HGA offers good performance in constructing gene regulatory networks for cutaneous melanoma and type 2 diabetes mellitus with more accurate results.

Conclusion

GRN reflects the relationship between genes that directly or indirectly influence their expression. GRN research offers insights into life processes and guides gene-disease diagnosis, treatment, and drug design. With the advancement of biological techniques and subject knowledge, genetic control system research based on genetic information is emerging. This includes Boolean networks, probabilistic Boolean networks (reflecting uncertain regulation), Bayesian networks (handling noise), dynamic Bayesian networks (showing gene regulation feedback), and differential equations (showing more detailed gene regulation phenomena).In this paper, a new hybrid parallel genetic method based on the threshold method is employed to enhance accuracy. Initially, the solution domains of each genetic control network are derived using SVD. Then, the sparse matrix with greater biological value is obtained via thresholding. This matrix uses a hybrid parallel genetic method to fit the network's weight. This method combines global solution optimization using a parallel GA with local approximation using mountain climbing.

The proposed gene regulation dynamics model directly represents regulatory relationships, theoretically offering a deeper explanation of gene regulation mechanisms. It directly captures

detailed reactions among substances involved in gene regulation. In model research, prior networks' regulation relations are sometimes hidden and models optimized to fit the complete network space are searched to identify these hidden relations. However, the model's high-dimensional solution space and limited time-series gene expression profile data lead to overfitting. Therefore, further research is needed to address this issue using more experimental data.

Acknowledgements

This work was supported by the Natural Science Foundation of Anhui Province: Research on Routing Optimization Strategy in Cloud and Fog Computing Network Architecture Based on SDN (No. 2022AH052854); The Human Face Emotion Recognition Algorithm Research based on the Integrated Depth Neural Network (No. 2022AH052852), the Anhui Province Social Science Foundation; and the 3D Special Effect Technology in the Game Scene Virtual Interactive Application Research (No. 2022AH052847).

References

1. Akman O., T. Comar, A. L. Harris, D. Hrozencik, et al. (2018). Dynamics of Gene Regulatory Networks with Stochastic Propensities, International Journal of Biomathematics, 11(3), https://doi.org/10.1142/S1793524518500328.
2. Babu M. M. (2011). Structure, Evolution and Dynamics of Gene Regulatory Networks, Abstracts of Papers of the American Chemical Society, 241, 1155.
3. Dhillon-Richardson R. M., A. K. Haugan, M. L. Martik (2022). Fishing for Developmental Regulatory Regions: Zebrafish Tissue-specific ATAC-seq, In DNA-Protein Interactions: Methods and Protocols, 271-282.
4. García-Calvo R., J. L. Guisado, F. Diaz-del-Rio, A. Córdoba, et al. (2018). Graphics Processing Unit-enhanced Genetic Algorithms for Solving the Temporal Dynamics of Gene Regulatory Networks, Evolutionary Bioinformatics, 14, https://doi.org/10.1177/1176934318767889.
5. George R. M., B. A. Firulli, R. Podicheti, D. B. Rusch, et al. (2023). Single Cell Evaluation of Endocardial Hand2 Gene Regulatory Networks Reveals HAND2-dependent Pathways that Impact Cardiac Morphogenesis, Development, 150(3), dev201341, https://doi.org/10.1242/dev.201341.
6. Greenham K., C. R. McClung (2018). Time to Build on Good Design: Resolving the Temporal Dynamics of Gene Regulatory Networks, Proceedings of the National Academy of Sciences of the United States of America, 115(25), 6325-6327.
7. Huang J., B. Tang, R. Ren, M. Wu, et al. (2022). Understanding the Potential Gene Regulatory Network of Starch Biosynthesis in Tartary Buckwheat by RNA-Seq, International Journal of Molecular Sciences, 23(24), 15774, https://doi.org/10.3390/ijms232415774.
8. Huang L., B. Clauss, M. Lu (2022). What Makes a Functional Gene Regulatory Network? A Circuit Motif Analysis, The Journal of Physical Chemistry B, 126(49), 10374-10383.
9. Ironi L., D. X. Tran (2016). Nonlinear and Temporal Multiscale Dynamics of Gene Regulatory Networks: A Qualitative Simulator, Mathematics and Computers in Simulation, 125, 15-37, https://doi.org/10.1016/j.matcom.2015.11.007.
10. Kamimoto K., M. Adil, K. Jindal, C. Hoffmann, et al. (2023). Gene Regulatory Network Reconfiguration in Direct Lineage Reprogramming, Stem Cell Reports, 18(1), 97-112.
11. Keyl P., P. Bischoff, G. Dernbach, M. Bockmayr, et al. (2023). Single-cell Gene Regulatory Network Prediction by Explainable AI, Nucleic Acids Research, 51(4), e20, https://doi.org/10.1093/nar/gkac1212.

12. Khullar S., D. Wang (2022). Predicting Gene Regulatory Networks from Multi-omics to Link Genetic Risk Variants and Neuroimmunology to Alzheimer's Disease Phenotypes, Alzheimer's & Dementia, 18, e068648, https://doi.org/10.1002/alz.068648.
13. Liu C., L. Hao, J. Lei (2021). Macroscopic Dynamics of Gene Regulatory Networks Revealed by Individual Entropy, International Journal of Modern Physics B, 35(32), https://doi.org/10.1142/S0217979221503318.
14. Liu J., S. Liu, Z. Yu, X. Qui, et al. (2022). Uncovering the Gene Regulatory Network of Type 2 Diabetes through Multi-omic Data Integration, Journal of Translational Medicine, 20(1), 604, https://doi.org/10.1186/s12967-022-03826-5.
15. Luo X., Y. Wei, Y. Zheng, L. Wei, et al. (2023). Analysis of Co-expression and Gene Regulatory Networks Associated with Sterile Lemma Development in Rice, BMC Plant Biology, 23(1), 11, https://doi.org/10.1186/s12870-022-04012-x.
16. Núñez-Lillo G., W. Pérez-Reyes, A. Riveros, V. Lillo-Carmona, et al. (2022). Transcriptome and Gene Regulatory Network Analyses Reveal New Transcription Factors in Mature Fruit Associated with Harvest Date in *Prunus persica*, Plants, 11(24), 3473, https://doi.org/10.3390/plants11243473.
17. Shi K., B. Wang, L. Dou, S. Wang, et al. (2022). Integrated Bioinformatics Analysis of the Transcription Factor-mediated Gene Regulatory Networks in the Formation of Spermatogonial Stem Cells, Frontiers in Physiology, 13, https://doi.org/10.3389/fphys.2022.949486.
18. Spanos A. C., S. T. Ponis, I. P. Tatsiopoulos, I. T. Christou, et al. (2014). A New Hybrid Parallel Genetic Algorithm for the Job-shop Scheduling Problem, International Transactions in Operational Research, 21(3), 479-499.
19. Sun Z., J. Yang, J. Zhou, J. Zhou, et al. (2023). Tissue-specific Oxysterols as Predictors of Antidepressant (Escitalopram) Treatment Response in Patients with Major Depressive Disorder, Biological Psychiatry Global Open Science, 3(4), 663-672.
20. Tang X. L., X. R. Wang (2011). Text Summarization Using Hybrid Parallel Genetic Algorithm, Advanced Materials Research, 271, 154-157.
21. Wang J., Y. Chen, Q. Zou (2023). Inferring Gene Regulatory Network from Single-cell Transcriptomes with Graph Autoencoder Model, PLoS Genetics, 19(9), e1010942, https://doi.org/10.1371/journal.pgen.1010942.
22. Wang P., H. Wang (2020). Research on the Construction of the Gene Regulatory Network Based on the Hybrid Parallel Genetic Algorithms, Basic and Clinical Pharmacology and Toxicology, 126, 150-151.
23. Wang P., J. Wang, X. Sun, X. Yang, et al. (2022). Construction of a Hierarchical Gene Regulatory Network to Reveal the Drought Tolerance Mechanism of Shanxin Poplar, International Journal of Molecular Sciences, 24(1), 384, https://doi.org/10.3390/ijms24010384.
24. White M., J. Arif-Pardy, K. L. Connor (2023). Identification of Novel Nutrient-sensitive Gene Regulatory Networks in Amniocytes from Fetuses with *Spina Bifida*, Reproductive Toxicology, 116, 108333, https://doi.org/10.1016/j.reprotox.2022.12.010.
25. Williams T. M., M. Rebeiz (2023). JEZB Special Issue on Animal Gene Regulatory Network Evolution, Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 340(2), 89-91.
26. Xie Q., Y. Wang, D. Wang, J. Li, et al. (2023). The Multilayered Hierarchical Gene Regulatory Network Reveals Interaction of Transcription Factors in Response to Cadmium in *Tamarix hispida* Roots, Tree Physiology, 43(4), 630-642.
27. Xu Q., S. Huang, G. Guo, C. Yang, et al. (2022). Inferring Regulatory Element Landscapes and Gene Regulatory Networks from Integrated Analysis in Eight Hulless Barley Varieties

Under Abiotic Stress, BMC Genomics, 23(1), 843, https://doi.org/10.1186/s12864-022-09070-x.
28. Xu W., Y. Wang, J. Xie, S. Tan, et al. (2023). Growth-regulating Factor 15-mediated Gene Regulatory Network Enhances Salt Tolerance in Poplar, Plant Physiology, 191(4), 2367-2384.
29. Yang C. H., S. V. Scarpino (2023). The Ensemble of Gene Regulatory Networks at Mutation-selection Balance, Journal of the Royal Society Interface, 20(198), 20220075, https://doi.org/10.1098/rsif.2022.0075.
30. Zheng M., S. Zhang, Y. Zhou, G. Liu (2018). Inferring Gene Regulatory Networks Based on a Hybrid Parallel Genetic Algorithm and the Threshold Restriction Method, Interdisciplinary Sciences: Computational Life Sciences, 10, 221-232.

Xiaomei Ding, M.Sc.
E-mail: dingxiaomei@mail.wenda.edu.cn

Xiaomei Ding was born in Fuyang, Anhui Province, P.R. China in 1984. She received her M.Sc. Degree from Hefei University of Technology, China. Now, Xiaomei Ding works at Anhui Wenda University of Information Engineering, China. Her research interests include algorithms, software engineering, and computer science.

Huaibao Ding, M.Sc.
E-mail: dinghuaibao@mail.wenda.edu.cn

Huaibao Ding was born in Anqing, Anhui, P.R. China, in 1985. From 2007 to 2010, she studied at Shanghai Normal University and received her M.Sc. Degree in 2010. Now, Huaibao Ding works at Anhui Wenda University of Information Engineering, China. Her research interests include algorithms, artificial intelligence, and software engineering.

Fei Zhou, M.Sc.
E-mail: zhoufei@mail.wenda.edu.cn

Fei Zhou was born in Feixi, Anhui, P.R. China, in 1982. From 2000 to 2004, he studied at Hefei University of Technology and received his B.Sc. Degree in 2004. From 2009 to 2012, he studied again at Hefei University of Technology and received his M.Sc. Degree in 2012. Since 2004, Fei Zhou has worked at Anhui Wenda University of Information Engineering. His research interests include software engineering and internet technology.