Research Paper

Identifying drug-pathway association pairs based on $L_1L_{2,1}$-integrative penalized matrix decomposition

Dong-Qin Wang[1], Ying-Lian Gao[2], Jin-Xing Liu[1], Chun-Hou Zheng[1] and Xiang-Zhen Kong[1]

[1]School of Information Science and Engineering, Qufu Normal University, Rizhao, China

[2]Library of Qufu Normal University, Qufu Normal University, Rizhao, China

Correspondence to: *Jin-Xing Liu, email: sdcavell@126.com*

Chun-Hou Zheng, email: zhengch99@126.com

ABSTRACT

The traditional methods of drug discovery follow the "one drug-one target" approach, which ignores the cellular and physiological environment of the action mechanism of drugs. However, pathway-based drug discovery methods can overcome this limitation. This kind of method, such as the Integrative Penalized Matrix Decomposition (iPaD) method, identifies the drug-pathway associations by taking the lasso-type penalty on the regularization term. Moreover, instead of imposing the L_1-norm regularization, the $L_{2,1}$-Integrative Penalized Matrix Decomposition ($L_{2,1}$-iPaD) method imposes the $L_{2,1}$-norm penalty on the regularization term. In this paper, based on the iPaD and $L_{2,1}$-iPaD methods, we propose a novel method named $L_1L_{2,1}$-iPaD ($L_1L_{2,1}$-Integrative Penalized Matrix Decomposition), which takes the sum of the L_1-norm and $L_{2,1}$-norm penalties on the regularization term. Besides, we perform permutation test to assess the significance of the identified drug-pathway association pairs and compute the P-values. Compared with the existing methods, our method can identify more drug-pathway association pairs which have been validated in the CancerResource database. In order to identify drug-pathway associations which are not validated in the CancerResource database, we retrieve published papers to prove these associations. The results on two real datasets prove that our method can achieve better enrichment for identified association pairs than the iPaD and $L_{2,1}$-iPaD methods.

INTRODUCTION

With the rapid development of data generated from genetic analyses and functional genomics, identifying drug targets has become more and more feasible [1]. The modern drug discovery has found thousands of drug targets and the clinical testing agents. These agents can produce curative effects via regulating different targets in various biological and related diseases of pathways. Traditional drug discovery methods follow the "one drug-one target" line. But many complex diseases are related to dysfunction of multiple pathways rather than individual genes [2]. These methods ignore the relationship among genes and the systemic nature of human diseases. In general, pathways are defined as the interaction of multiple genes, that is, pathways can be regarded as the target of drugs [3, 4]. Abnormal biological pathways can provide views for the aberrant imbalance underlying diseases and find targets for complex diseases intervention [5].

Compared with the "one drug-one target" methods, the systematic biology approach takes the drug effects into the global physiological environment account [6]. These existing computational methods for identifying drug targets include the Gene Set Enrichment Analysis (GSEA) method [7], the FacPad method [2], the iFad method [5] and the iPaD method [8], etc. The GSEA

method has several disadvantages. Firstly, for every paired drug-pathway association, calculation must be done once at every turn. Secondly, the genes which belong to the same pathway will have the common weights, therefore, when a subblock of genes serve as the critical interaction groups for a specific drug. In summary, the GSEA method is time-consuming and inaccurate. The FacPad method has been proposed to identify drug-pathway association pairs, and it develops a sparse Bayesian factor analysis model to deal with treatment response data that are derived from microarray platforms [2]. And the iFad method is also a sparse Bayesian factor model, which is proposed to infer the drug-pathway association using gene expression and drug sensitivity data. In literature [5], the authors apply this method on the NCI-60 data set. The gene expression and drug sensitivity data are downloaded from the CellMiner database [9] (http://discover.nci.nih.gov/cellminer). The iFad method is effective in identifying paired drug-pathway associations. Since this method uses the Markov Chain Monte Carlo (MCMC) [10] to perform statistical inferences, it is computationally expensive. Also it has many turning parameters which should be specified by users. In order to improve the algorithm speed and performance, a method named iPaD (integrative Penalized Matrix Decomposition) is proposed by Li et al.[8] to identify paired drug-pathway associations. The iPaD method applies the integrative penalized matrix decomposition method to analyze the gene expression data and the drug sensitivity data. The scalable bi-convex optimization algorithm is used to solve the objective function. Since the L_1-norm penalty can produce sparsity, the L_1-norm regularization item is added on the drug-pathway association matrix. By applying the iPaD method on the NCI-60 and Cancer Cell Line Encyclopedia (CCLE) datasets, Li et al. [8] prove that the iPaD method performs better than the iFad method in computational efficiency and identifying drug-pathway association pairs. Since the $L_{2,1}$-norm regularization can penalize each row of the matrix as a whole and can enhance the sparsity among the rows [11], based on this theory, a method named $L_{2,1}$-iPaD ($L_{2,1}$-integrative Penalized Matrix Decomposition) is proposed to identify paired drug-pathway associations [12]. The L_1-norm penalization can produce scattered and unstructured sparsity matrix, yet $L_{2,1}$-norm penalization can produce structured row sparsity matrix [11]. Moreover, the sum of the L_1-norm and $L_{2,1}$-norm penalization can produce row structure with intra-row sparsity. In this paper, for the purpose of enhancing the sparsity of the drug-pathway association matrix and improving the performance of the method, we use the sum of the L_1-norm regularization and $L_{2,1}$-norm regularization instead of the L_1-norm regularization. In this article, a novel method named "$L_1L_{2,1}$-iPaD" is proposed to identify paired drug-pathway associations. Our method has the following advantages. Firstly, for the first time, we propose the $L_1L_{2,1}$-iPaD method by replacing the L_1-norm

regularization with the sum of the L_1-norm regularization and $L_{2,1}$-norm regularization. Secondly, the $L_1L_{2,1}$-iPaD method can be used to analyze gene expression and drug sensitivity data. Thirdly, it gives an effective method to identify drug-pathway association pairs. Experimental results in two real datasets prove that the $L_1L_{2,1}$-iPaD method is effective.

The remainder of this paper is organized as follows. Firstly, we will introduce two real datasets, give out the experimental results and show the comparison of our method with the state-of-art methods. Then we will provide discussion of this article and outline the future works. And thirdly, we will introduce the notations and definitions in this paper. Finally, we will describe the related works and methodology of the $L_1L_{2,1}$-iPaD.

THE INTRODUCTION FOR DATASETS

CCLE data set

The CCLE data set is downloaded from the CCLE project, which is a competitive resource to implement a detailed genetic characterization for a large panel of human cancer cell lines. The CCLE project (http://software.broadinstitute.org/software/cprg/?q=node/11) can provide public access analysis, visualization of DNA copy number, mutation data, mRNA expression, and so on, for about 1046 cancer cell lines. The CCLE data set consists of 480 cell lines with drug sensitivity data for 22 drugs and transcription data for 1802 genes covering 58 KEGG pathways. In fact, the CCLE data set in the CCLE project contains 18988 genes and 24 chemical drugs. The detailed data processing can be found in [8]. The drug sensitivities are measured by area over the dose-response curve ('activity area'). EC_{50}, IC_{50} and maximum activity area ('A_{max}') can also measure drug sensitivities. But activity area has two advantages. Firstly, it has less missing values. Secondly, it can both reflect drug potency and efficacy. But EC_{50} and IC_{50} can only reflect drug potency, and A_{max} can only reflect its efficacy. In this paper, the density of the gene-pathway association prior knowledge matrix $\mathbf{H}^{(1)}$ is 3.95%. And for the drug-pathway association prior knowledge matrix $\mathbf{H}^{(2)}$, we set it to zero.

NCI-60 data set

The NCI-60 data set contains 57 cell lines with gene expression data for 1863 genes covering 58 KEGG [13] pathways and drug sensitivity data for 101 drugs. The NCI-60 data set is from the NCI-60 project, which provides various type of 'Omics' features of 60 cell lines with 9 different cancer types. The gene expression data and the drug sensitivity profiles are downloaded from the CellMiner database [9]. The detailed data processing can be found in [5]. The drug sensitivity data are obtained by GI_{50} values, those values can reflect the potency of drugs.

The GI_{50} is defined as the needed concentration for 50% of maximum cell growth inhibition. The drug sensitivity data values are equal to the $\log 10(GI_{50})$. Higher values mean higher drug sensitivity of the cell lines. Besides, the density of the gene-pathway association prior knowledge matrix $\mathbf{H}^{(1)}$ is 3.95%. And the density of the drug-pathway association prior knowledge matrix $\mathbf{H}^{(2)}$ is 0.51%.

RESULTS

In this Section, we evaluate the performance of our proposed $L_1L_{2,1}$-iPaD method by applying to the CCLE and NCI-60 datasets. To show the effectiveness of our proposed method, we also compare our method with the iPaD and $L_{2,1}$-iPaD methods in this Section.

The results on the CCLE data set

In this experiment, we first use the five-fold cross-validation to obtain the optimal λ_1 and λ_2 values, and then obtain the sparse drug-pathway association matrix $\mathbf{B}^{(2)}$ corresponding to the optimal parameter values. Similar to the iPaD and $L_{2,1}$-iPaD methods, our method also can assess the relative importance of the coefficients in the drug-pathway association matrix $\mathbf{B}^{(2)}$ via solving the $L_1L_{2,1}$-problem for a decreasing λ sequence. For each λ value, we record the order of the coefficients in which they become nonzero. In general, the more important coefficients ought to become nonzero earlier than the less important coefficients. However, this procedure cannot be used to assess the significance of the coefficients. Therefore, we perform permutation test to assess the significance of the coefficients in the drug-pathway association matrix $\mathbf{B}^{(2)}$. In this paper, we run 2000 permutations to obtain the P-values. The P-values of our method, the $L_{2,1}$-iPaD and iPaD methods on the CCLE data set are listed in Table 1, in which the superior results are shown in bold type. In this paper, the known drug-pathway associations are served as the validated information. In Table 1, the drug named Nutlin-3 is related with the Chronic myeloid leukemia pathway. A published paper suggests that Nutlin-3 can up-regulate the expression of Notch1 in both lymphoid and myeloid leukemic cells for a part of the negative feedback antiapoptotic mechanism [14]. Besides, the authors in [15] confirm that IAP inhibition using a small synthetic inhibitor (LBW-242) increases the sensitivity of CML cells to TKI. This drug-pathway association pair is not validated in the CancerResource, but our method can find their associations. And in Table 1, only 5 drug-pathway association pairs are validated in the CancerResource. For the rest 10 drug-pathway associations which are not validated in the CancerResource, we retrieve published papers to prove their associations. Only one association pair is not found from published papers. Besides, similarly,

the authors in [16] suggest that cotreatment with LBH589 and 17-AAGcan induce more apoptosis of IM-resistant primary CML-BC and acute myeloidleukemia cells than treatments with either agent alone. And in this paper, our new method also can find that 17-AAG is associated with the Chronic myeloid leukemia pathway. Moreover, the drug-pathway pairs corresponding to nonzero elements in the matrix $\mathbf{B}^{(2)}$ are selected as the identified drug-pathway association pairs. In this experiment, our method identifies 413 drug-pathway pairs that have p-value no more than 0.05, and 70 drug-pathway pairs are validated in the CancerResource database. The $L_{2,1}$-iPaD method identifies 368 drug-pathway pairs that have p-value no more than 0.05, and 66 drug-pathway pairs are validated in the CancerResource database. However, iPaD identifies 88 drug-pathway pairs that have p-value no more than 0.05, and only 25 drug-pathway pairs are validated in the CancerResource database. When we set the P-value cutoff as 0.005, 51 drug-pathway association pairs are identified by the iPaD method, with 16 association pairs validated in the CancerResource database. But for our method and the $L_{2,1}$-iPaD method, 53 drug-pathway association pairs are identified, with 16 association pairs validated in the CancerResource database. In addition, in [8], we can easily find that the results of iFad and GSEA methods are poorer than the iPaD method, so in this paper, we does not compare our method with the iFad and GSEA methods.

Then we compute the identification and verification rates of drug-pathway association pairs on the CCLE data set. Specifically, we make the ratio of number of identification (or verification) and total number of drug-pathway pairs as the identification (or verification) rate. The identification and verification rates of drug-pathway association pairs on the CCLE data set for the $L_1L_{2,1}$-iPaD, $L_{2,1}$-iPaD and iPaD methods can be found in Table 2 and Table 3. It is obvious that our method can identify more drug-pathway association pairs than other existing methods.

The results on the NCI-60 data set

Similar to the $L_{2,1}$-iPaD and iPaD methods, we also apply our method to the NCI-60 data set. For the NCI-60 data set, we also run 2000 permutation to evaluate the significance of the coefficients in the drug-pathway association matrix $\mathbf{B}^{(2)}$. The P-values of our method, the $L_{2,1}$-iPaD and iPaD methods on the NCI-60 data set are listed in Table 4. The authors in [17] prove that the mechanism of action of the EA derivatives prepared in this study is more complex than the inhibition of glutathione S-transferase p ascribed as unique effect to EA and might help to overcome tumor resistances. And in Table 4, Melphalan is associated with T cell receptor signaling pathway. A study published in 2003 suggests that Melphalan can control the expression of T cell receptor signaling pathway [18]. In Table 4, only

Table 1: The top 15 identified drug-pathway associations on CCLE data set by $L_1L_{2,1}$-iPaD, $L_{2,1}$-iPaD and iPaD methods

Drug	Pathway	$L_1L_{2,1}$-iPaD	$L_{2,1}$-iPaD	iPaD	Validated
			P-value		
Nutlin-3	Chronic myeloid leukemia	**0**	1.74E-43	1.09E-17	CR
PD-0332991	Chronic myeloid leukemia	**0**	6.93E-41	3.16E-13	CR
LBW242	Chronic myeloid leukemia	**1.51E-45**	2.80E-44	8.08E-17	[28]
17-AAG	Chronic myeloid leukemia	**5.52E-45**	9.46E-43	3.41E-16	[16]
L-685458	Chronic myeloid leukemia	**6.81E-44**	4.33E-43	3.32E-19	[29]
AZD0530	Colorectal cancer	**1.46E-43**	1.62E-41	8.81E-13	[30]
PHA-665752	Chronic myeloid leukemia	**1.03E-41**	1.09E-40	3.41E-16	[31]
Paclitaxel	Chronic myeloid leukemia	**2.11E-40**	2.14E-38	4.58E-12	[32]
AZD0530	Chronic myeloid leukemia	**5.86E-40**	7.12E-38	4.76E-18	CR
PD-0325901	Thyroid cancer	**1.25E-28**	3.59E-12	2.57E-05	[33]
ZD-6474	Chronic myeloid leukemia	**1.52E-22**	1.62E-21	2.36E-10	[34]
RAF265	ECM-receptor interaction	**2.19E-17**	1.26E-15	2.32E-04	Unfound
AZD0530	ErbB signaling pathway	**1.13E-16**	4.41E-16	5.10E-06	CR
Erlotinib	Chronic myeloid leukemia	**3.71E-15**	5.69E-15	2.34E-08	[35]
Nilotinib	ErbB signaling pathway	**4.00E-14**	1.23E-13	1.80E-05	CR

Table 2: The identification and verification rates on CCLE data set with P-value<0.005

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_1L_{2,1}$-iPaD	**53**	**16**	**0.0125**	**0.0415**
$L_{2,1}$-iPaD	**53**	**16**	**0.0125**	**0.0415**
iPaD	51	**16**	**0.0125**	0.0399

Table 3: The identification and verification rates on CCLE data set with P-value<0.05

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_1L_{2,1}$-iPaD	**413**	**70**	**0.0549**	**0.3237**
$L_{2,1}$-iPaD	368	66	0.0517	0.2884
iPaD	88	25	0.0196	0.0689

Table 4: The top 15 identified drug-pathway associations on NCI-60 data set by $L_1 L_{2,1}$-iPaD, $L_{2,1}$-iPaD and iPaD methods

Drug	Pathway	$L_1 L_{2,1}$-iPaD	$L_{2,1}$-iPaD	iPaD	Validated
		P-value			
Hydroxyurea	Neuroactive ligand-receptor interation	**0**	**0**	NAN	[36]
Rebeccamycin	T cell receptor signaling pathway	**1.78E-17**	4.12E-16	4.65E-10	Unfound
Tiazofurin	Cell cycle	**7.70E-12**	8.19E-11	7.54E-07	CR
Selenazofurin	Cell cycle	**8.27E-11**	1.75E-10	2.78E-07	CR
Mycophenolic acid	Cell cycle	**9.02E-11**	2.61E-10	2.52E-06	[19]
Lucanthone	Tight junction	2.06E-08	**9.97E-09**	4.31E-06	CR
Primaquine	Neuroactive ligand-receptor interation	**4.46E-08**	1.14E-06	2.69E-04	[37]
Ethacrynic acid	Glutathione metabolism	**1.17E-07**	2.29E-02	6.36E-03	CR
Aminoglutethimide	Primary immunodeficiency	**6.55E-07**	1.30E-06	1.16E-04	[38]
Diallyl disulfide	Acute myeloid leukemia	**8.76E-07**	8.13E-06	8.41E-05	[39]
Bleomycin	Focal adhesion	**1.46E-06**	1.17E-05	4.56E-04	[40]
Geldanamycin	Gap junction	**1.46E-06**	7.89E-06	1.87E-04	[41]
Melphalan	T cell receptor signaling pathway	**3.76E-06**	2.64E-05	6.16E-04	CR
Lomustine	Tight junction	**6.74E-06**	1.06E-05	2.64E-04	CR
Vitamin K3	Metabolism of xenobiotics by cytochrome P450	**9.98E-06**	2.22E-05	2.71E-04	[42]

6 drug-pathway association pairs are validated in the CancerResource. For the rest 9 drug-pathway associations which are not validated in the CancerResource, we retrieve published papers to prove their associations. Only one pair associations are not found from published papers. So, our method can also identify associations, which are not validated in CancerResource, in the NCI-60 data set between drugs and pathways. For example, a paper studied in 2000 writes that an *in vitro* study the effects of MPA (Mycophenolic acid) on human peripheral blood lymphocyte activation markers and on cell cycle characteristics are investigated [19]. Moreover, the drug-pathway pairs corresponding to nonzero elements in the matrix $\mathbf{B}^{(2)}$ are selected as the identified drug-pathway association pairs. In this experiment, our method identifies 593 drug-pathway pairs that have p-value no more than 0.05, and 163 drug-pathway pairs are validated in the CancerResource database. The $L_{2,1}$-iPaD method identifies 562 drug-pathway pairs that have p-value no more than 0.05, and 163 drug-pathway pairs are validated in the CancerResource database. However, iPaD identifies 247 drug-pathway pairs that have p-value no more than 0.05, and only 74 drug-pathway pairs are validated in the CancerResource database. And when we set the P-value cutoff as 0.005, the results of our method is similar to the $L_{2,1}$-iPaD method, but the number of identification and verification are higher than the iPaD method. Similar to the CCLE data set, we also compute the identification and verification rates of drug-pathway association pairs on the NCI-60 data set. The identification and verification rates of drug-pathway association pairs on the NCI-60 data set for the $L_1 L_{2,1}$-iPaD, $L_{2,1}$-iPaD and iPaD methods can be found in Table 5 and Table 6.

Table 5: The identification and verification rates on NCI-60 data set with P-value<0.05

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_1L_{2,1}$-iPaD	**593**	**163**	**0.0278**	**0.1012**
$L_{2,1}$-iPaD	562	**163**	**0.0278**	0.0959
iPaD	247	74	0.0126	0.0422

Table 6: The identification and verification rates on NCI-60 data set with P-value<0.005

Method	Number of identification	Number of verification	Verification rate	Identification rate
$L_1L_{2,1}$-iPaD	**89**	**34**	**0.0058**	**0.0152**
$L_{2,1}$-iPaD	**89**	33	0.0056	**0.0152**
iPaD	72	26	0.0044	0.0122

DISCUSSION

Identifying drug-targets is a momentous issue for the bioinformatics and a crucial step for the drug discovery. In this paper, we proposed a novel method named "$L_1L_{2,1}$-iPaD" to identify the paired drug-pathway associations, and used it to jointly analyze the gene expression and drug sensitivity data. In addition, for the $L_1L_{2,1}$-iPaD method, two parameters need to adjust. So, we use five-fold cross-validation to choose the optimal parameters. Besides, we perform permutation test to assess the significance of the identified drug-pathway association pairs. In order to evaluate the performance of our method, we apply it to two real datasets, including the CCLE and NCI-60 datasets. Moreover, we compare our method with the iPaD and $L_{2,1}$-iPaD methods. The experimental results demonstrate that our method is superior to the iPaD and $L_{2,1}$-iPaD methods. Our method can identify more drug-pathway association pairs which are validated in the CancerResource database than other methods. Besides, for the rest drug-pathway associations which are not validated in the CancerResource database, we retrieve published papers to prove their associations.

With the development of the high-throughput technology, more and more genomic data sets are generated from various fields. At present, one of our central task is to develop the feasible and efficient method to analyze them. Our method is one of the useful ways for identifying drug-pathway association pairs. In the future, we will develop more efficient and robust methods to jointly analyze high dimensional data and to solve the computational challenges.

NOTATIONS AND DEFINITIONS

In this Section, we summarize the definition of norms, which will be used in the methods. Given a matrix $\mathbf{M} = (m_{i,j})$, $\mathbf{m}^i$ denotes the i-th row of the matrix $\mathbf{M}$.

For the matrix $\mathbf{M} \in R^{d \times n}$, the L_1-norm of a matrix was first presented in [20], whose definition can be written as

$$\|\mathbf{M}\|_1 = \sum_{i=1}^d \sum_{j=1}^n |m_{i,j}|. \tag{1}$$

The Frobenius norm of the matrix $\mathbf{M}$ can be defined as

$$\|\mathbf{M}\|_F = \sqrt{\sum_{i=1}^d \sum_{j=1}^n m_{i,j}^2} = \sqrt{\sum_{i=1}^d \|\mathbf{m}^i\|_2^2}. \tag{2}$$

And the $L_{2,1}$-norm of a matrix was first introduced in [21]. Until now, the $L_{2,1}$-norm has been used in many fields, such as the feature extraction [22] and the image processing [23, 24] etc. The $L_{2,1}$-norm of the matrix $\mathbf{M}$ is defined as

$$\|\mathbf{M}\|_{2,1} = \sum_{i=1}^d \sqrt{\sum_{j=1}^n m_{i,j}^2} = \sum_{i=1}^d \|\mathbf{m}^i\|_2. \tag{3}$$

The $L_{r,p}$-norm of the matrix $\mathbf{M}$ can be written as follows [16]:

$$\|\mathbf{M}\|_{r,p} = \left(\sum_{i=1}^d \left(\sum_{j=1}^n |m_{i,j}| \right)^{p/r} \right)^{1/p} = \left(\sum_{i=1}^d \|\mathbf{m}^i\|_r^p \right)^{1/p}. \tag{4}$$

RELATED WORKS

iPaD method

Let $\mathbf{Y}^{(1)} \in R^{N \times G^{(1)}}$ and $\mathbf{Y}^{(2)} \in R^{N \times G^{(2)}}$ denote drug sensitivity data and transcription data matrices, respectively. N is the number of the samples (usually cell lines). $G^{(1)}$ and $G^{(2)}$ are the number of genes and number of drugs, respectively. Besides, $\mathbf{X} \in R^{N \times K}$ denotes a pathway activity matrix, that is, it indicates the activity level of K pathways in the N samples. For the traditional iPaD method [8], the authors decompose the matrix $\mathbf{Y}^{(1)}$ into the matrices $\mathbf{X}$ and $\mathbf{B}^{(1)}$, and the matrix $\mathbf{Y}^{(2)}$ into the matrices $\mathbf{X}$ and $\mathbf{B}^{(2)}$. Therefore, the model can be introduced as follows:

$$\begin{aligned} \mathbf{Y}^{(1)} &= \mathbf{X}\mathbf{B}^{(1)} + \mathbf{E}^{(1)} \\ \mathbf{Y}^{(2)} &= \mathbf{X}\mathbf{B}^{(2)} + \mathbf{E}^{(2)}, \end{aligned} \tag{5}$$

where $\mathbf{E}^{(1)}$ and $\mathbf{E}^{(2)}$ are the error matrices. $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ denote the gene-pathway association and drug-pathway association matrices, respectively. In general, the model (5) can be formulated as follows:

$$\min_{\mathbf{X},\mathbf{B}^{(1)},\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(1)} - \mathbf{X}\mathbf{B}^{(1)} \right\|_F^2 + \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2, \qquad (6)$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

For the Eq.(6), the optimization model of iPaD method [8] is defined as follows:

$$\min_{\mathbf{X},\mathbf{B}^{(1)},\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(1)} - \mathbf{X}\mathbf{B}^{(1)} \right\|_F^2 + \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2 + \lambda \left\| \mathbf{B}^{(2)} \right\|_1$$
$$\text{subject to} \quad \sum_i \mathbf{X}_{i,j}^2 \leq 1, \forall j = 1,\ldots,K, \qquad (7)$$
$$\mathbf{B}_{i,j}^{(1)} = 0, \forall (i,j): \mathbf{H}_{i,j}^{(1)} = 0,$$

where $\left\| \mathbf{B}^{(2)} \right\|_1$ is the L_1-norm of the matrix $\mathbf{B}^{(2)}$. λ is a crucial parameter and used to adjust the sparsity of the matrix $\mathbf{B}^{(2)}$. In general, the bigger the value of λ is, the more sparse the matrix $\mathbf{B}^{(2)}$ is. Since a drug is usually related to a few pathways, the matrix $\mathbf{B}^{(2)}$ may be sparse. Because the L_1-norm can produce sparsity, the authors in [8] add the L_1-norm constraint on the matrix $\mathbf{B}^{(2)}$. And, the prior knowledge matrix $\mathbf{H}^{(1)} \in \{0,1\}^{K \times G^{(1)}}$ is an indicating matrix, that is, the matrix $\mathbf{H}^{(1)}$ is used to indicate elements in the matrix $\mathbf{B}^{(1)}$. If $H_{i,j}^{(1)} = 1$, it indicates that the j-th gene is associated with i-th pathway. However, if $H_{i,j}^{(1)} = 0$, it indicates that the j-th gene is not associated with the i-th pathway. The authors in [8] assume that the known gene-pathway associations are complete, so, they pay attention to discover the novel drug-pathway associations. Therefore, the second constraint condition is used to incorporate known gene-pathway associations. Besides, the first constraint condition is used to guarantee that the model is identical.

$L_{2,1}$-iPaD method

Since the $L_{2,1}$-norm penalty can produce rows sparsity [25], another effective method named $L_{2,1}$-iPaD is proposed to identify the novel drug-pathway association pairs [12]. In this paper, the $L_{2,1}$-norm regularization is used to replace the L_1-norm regularization to enforce sparse constraint on rows. And it modifies the optimization problem (7) as follows:

$$\min_{\mathbf{X},\mathbf{B}^{(1)},\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(1)} - \mathbf{X}\mathbf{B}^{(1)} \right\|_F^2 + \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2 + \lambda \left\| \mathbf{B}^{(2)} \right\|_{2,1}$$
$$\text{subject to} \quad \sum_i \mathbf{X}_{i,j}^2 \leq 1, \forall j = 1,\ldots,K; \qquad (8)$$
$$\mathbf{B}_{i,j}^{(1)} = 0, \forall (i,j): \mathbf{H}_{i,j}^{(1)} = 0,$$

where $\left\| \mathbf{B}^{(2)} \right\|_{2,1}$ denotes the $L_{2,1}$-norm of the matrix $\mathbf{B}^{(2)}$.

METHODOLOGY

Since the gene-pathway association information is available and complete, the central interest lies in the inference of the matrix $\mathbf{B}^{(2)}$, that is, the paired drug-pathway associations. We further consider enhancing the sparsity of the matrix $\mathbf{B}^{(2)}$. Thus, we add the L_1-norm regularization to impose the penalty among all elements in the matrix $\mathbf{B}^{(2)}$ and propose our new $L_1L_{2,1}$-iPaD method. The objective function of the $L_1L_{2,1}$-iPaD method is defined as follows:

$$\min_{\mathbf{X},\mathbf{B}^{(1)},\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(1)} - \mathbf{X}\mathbf{B}^{(1)} \right\|_F^2 + \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2 + \lambda_1 \left\| \mathbf{B}^{(2)} \right\|_1 + \lambda_2 \left\| \mathbf{B}^{(2)} \right\|_{2,1}$$
$$\text{subject to} \quad \sum_i \mathbf{X}_{i,j}^2 \leq 1, \forall j = 1,\ldots,K, \qquad (9)$$
$$\mathbf{B}_{i,j}^{(1)} = 0, \forall (i,j): \mathbf{H}_{i,j}^{(1)} = 0,$$

where λ_1 and λ_2 are two adjustable parameters, which can increase or decrease the sparsity of the matrix $\mathbf{B}^{(2)}$. In general, the bigger the λ_1 and λ_2 are, the more sparse the matrix $\mathbf{B}^{(2)}$ is. Optimization problem (9) is convex. That is, when we fix $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$, optimizing $\mathbf{X}$ is a convex problem, and when we fix $\mathbf{X}$, optimizing $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ are both convex optimization subproblems. Since it is difficult to directly obtain the solution, an effective method is proposed to solve the optimization problem in Eq.(9).

Optimizing X

$$\min_{\mathbf{X}} \left\| \mathbf{Y} - \mathbf{X}\mathbf{B} \right\|_F^2$$
$$\text{subject to} \quad \sum_i \mathbf{X}_{i,j}^2 \leq 1, \forall j = 1,\ldots,K, \qquad (10)$$

where $\mathbf{Y} = \left[\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)} \right]$ and $\mathbf{B} = \left[\mathbf{B}^{(1)}, \mathbf{B}^{(2)} \right]$. The iPaD method [8] used the traditional gradient descent method to optimize $\mathbf{X}$. Here, we also apply the gradient descent method to solve this problem. The objective function of Eq.(10) can be written as follows:

$$\left\| \mathbf{Y} - \mathbf{X}\mathbf{B} \right\|_F^2$$
$$= \text{Tr}\left[\left(\mathbf{Y} - \mathbf{X}\mathbf{B} \right)^T \left(\mathbf{Y} - \mathbf{X}\mathbf{B} \right) \right]. \qquad (11)$$

By computing the derivative of Eq.(11), we can obtain

$$\frac{\partial \left\| \mathbf{Y} - \mathbf{X}\mathbf{B} \right\|_F^2}{\partial \mathbf{X}} = -2 \left(\mathbf{Y} - \mathbf{X}\mathbf{B} \right) \mathbf{B}^T$$
$$= 2 \left(\mathbf{X}\mathbf{B}\mathbf{B}^T - \mathbf{Y}\mathbf{B}^T \right). \qquad (12)$$

So that $\mathbf{X}$ can be updated by

$$\mathbf{X}_{k+1} = \mathbf{X}_k - 2\mu \left(\mathbf{X}\mathbf{B}\mathbf{B}^T - \mathbf{Y}\mathbf{B}^T \right), k = 0,1,2,\cdots. \qquad (13)$$

Here, μ is the iteration step size. At every iteration, we check whether $\mathbf{X}_{k+1}$ is located in the feasible fields $\left\{ \mathbf{X} : \sum_i \mathbf{X}_{i,j}^2 \leq 1, \forall j = 1,\ldots,K \right\}$. If $\mathbf{X}_{k+1}$ satisfies this condition, we perform next step, otherwise, we make

$X_{k+1} = X_{k+1} / \|X_{k+1}\|$. The Nesterov's method [26] can achieve a convergence rate becoming $1/t^2$ (The original convergence rate is $1/t$). So, when we update $\mathbf{X}$, we also apply this method to quicken the convergence speed.

Optimizing $\mathbf{B}^{(1)}$

$$\min_{\mathbf{B}^{(1)}} \left\| \mathbf{Y}^{(1)} - \mathbf{X}\mathbf{B}^{(1)} \right\|_F^2$$
$$\text{subject to} \quad \mathbf{B}_{i,j}^{(1)} = 0, \forall (i,j): \mathbf{H}_{i,j}^{(1)} = 0. \tag{14}$$

Least squares method is a common optimization algorithm to solve the unconstrained optimization problems. In order to use the prior knowledge matrix $\mathbf{H}^{(1)}$ to optimize $\mathbf{B}^{(1)}$, similar to [8], we also decompose the original problem into $G^{(1)}$ OLS (ordinary least squares) problems. That is, we optimize each column of $\mathbf{B}^{(1)}$, separately. The problem (14) can be introduced as follows:

$$\text{For } q \in \left\{1, 2, \cdots, G^{(1)}\right\},$$
$$\min_{\mathbf{B}_{:,q}^{(1)}} \left\| \mathbf{Y}_{:,q}^{(1)} - \mathbf{X}_{:,H_{:,q}^{(1)}} \mathbf{B}_{H_{:,q}^{(1)},q}^{(1)} \right\|_F^2, \tag{15}$$

where $\mathbf{Y}_{:,q}^{(1)}$ is a vector with the elements corresponding to the q-th column of the matrix $\mathbf{Y}^{(1)}$, and $\mathbf{X}_{:,H_{:,q}^{(1)}}$ denotes the sub-matrix of the matrix $\mathbf{X}$, which is composed of the columns corresponding to the non-zero parts of the prior knowledge matrix $\mathbf{H}_{:,q}^{(1)}$. And $\mathbf{B}_{H_{:,q}^{(1)},q}^{(1)}$ is a vector with the q-th column vector of $\mathbf{B}^{(1)}$ corresponding to the non-zero parts of the prior knowledge matrix $\mathbf{H}_{:,q}^{(1)}$.

Optimizing $\mathbf{B}^{(2)}$

For optimizing $\mathbf{B}^{(2)}$, we observe each column of the matrix $\mathbf{B}^{(2)}$ and decompose the original problem into $G^{(2)}$ sum of L_1-norm and $L_{2,1}$-norm minimization problems:

$$\text{For } q \in \left\{1, 2, \cdots, G^{(2)}\right\},$$
$$\min_{\mathbf{B}_{:,q}^{(2)}} \left\| \mathbf{Y}_{:,q}^{(2)} - \mathbf{X}\mathbf{B}_{:,q}^{(2)} \right\|_2^2 + \lambda_1 \left\| \mathbf{B}_{:,q}^{(2)} \right\|_1 + \lambda_2 \left\| \mathbf{B}_{:,q}^{(2)} \right\|_{2,1}. \tag{16}$$

In order to use the prior knowledge on the drug-pathway associations (matrix $\mathbf{H}^{(2)}$), we add the L_2-norm on the matrix $\mathbf{B}^{(2)}$. Thus, the problem (16) can be rewritten as follows:

$$\text{For } q \in \left\{1, 2, \cdots, G^{(2)}\right\},$$
$$\min_{\mathbf{B}_{:,q}^{(2)}} \left\| \mathbf{Y}_{:,q}^{(2)} - \mathbf{X}\mathbf{B}_{:,q}^{(2)} \right\|_2^2 + \lambda_1 \left\| \mathbf{B}_{(1-H_{:,q}^{(2)}),q}^{(2)} \right\|_1 + \lambda_2 \left(\left\| \mathbf{B}_{(1-H_{:,q}^{(2)}),q}^{(2)} \right\|_{2,1} + \left\| \mathbf{B}_{H_{:,q}^{(2)},q}^{(2)} \right\|_2 \right). \tag{17}$$

Similar to the prior information matrix $\mathbf{H}^{(1)}$, $\mathbf{H}_{i,j}^{(2)} \in \{0,1\}^{K \times G^{(2)}}$ is also a prior knowledge matrix, which can indicate drug-pathway association matrix $\mathbf{B}^{(2)}$. λ_1 and λ_2 are two regulable parameters which are used to control the sparsity of the paired drug-pathway association matrix $\mathbf{B}^{(2)}$.

For the part of $\mathbf{B}^{(2)}$ being indicated by $\mathbf{H}^{(2)}$, the problem is written as follows:

$$\min_{\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2 + \lambda_2 \left\| \mathbf{B}^{(2)} \right\|_2^2, \tag{18}$$

where we omit the symbol of the objective function. The objective function can be converted into the following equation:

$$\begin{aligned} \mathrm{J}_1\left(\mathbf{B}^{(2)}\right) &= \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2 + \lambda_2 \left\| \mathbf{B}^{(2)} \right\|_2^2 \\ &= \mathrm{Tr}\left[\left(\mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)}\right)\left(\mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)}\right)^T \right] + \lambda_2 \mathrm{Tr}\left(\left(\mathbf{B}^{(2)}\right)^T \mathbf{I} \mathbf{B}^{(2)} \right), \end{aligned} \tag{19}$$

where $\mathbf{I}$ is a unit matrix with the size of $K \times K$, and $\mathrm{J}_1(\cdot)$ $\mathrm{J}_1\left(\mathbf{B}^{(2)}\right)$ is an auxiliary function. Then we compute the derivative of $\mathrm{J}_1\left(\mathbf{B}^{(2)}\right)$ and set its result to zero,

$$\frac{\partial \mathrm{J}_1\left(\mathbf{B}^{(2)}\right)}{\partial \mathbf{B}^{(2)}} = -2\mathbf{X}^T\mathbf{Y}^{(2)} + 2\left(\mathbf{X}^T\mathbf{X} + \lambda_2\mathbf{I}\right)\mathbf{B}^{(2)} \tag{20}$$
$$= \mathbf{0}.$$

Hence, we can obtain:

$$\mathbf{B}^{(2)} = \left(\mathbf{X}^T\mathbf{X} + \lambda_2\mathbf{I}\right)^{-1} \mathbf{X}^T\mathbf{Y}^{(2)}. \tag{21}$$

For the part of $\mathbf{B}^{(2)}$ being indicated by $1 - \mathbf{H}^{(2)}$, we will provide a novel and available method to obtain the interesting drug-pathway association matrix $\mathbf{B}^{(2)}$. The optimization problem can be described as follows:

$$\begin{aligned} &\min_{\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{B}^{(2)} \right\|_F^2 + \lambda_1 \left\| \mathbf{B}^{(2)} \right\|_1 + \lambda_2 \left\| \mathbf{B}^{(2)} \right\|_{2,1} \\ &= \min_{\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(2)} - \mathbf{X}\mathbf{D}^{-1/2}\mathbf{D}^{1/2}\mathbf{B}^{(2)} \right\|_F^2 + \lambda_1 \mathrm{Tr}\left[\left(\mathbf{B}^{(2)}\right)^T \tilde{\mathbf{D}} \mathbf{B}^{(2)} \right] + \lambda_2 \mathrm{Tr}\left[\left(\mathbf{B}^{(2)}\right)^T \mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{B}^{(2)} \right], \end{aligned} \tag{22}$$

where $\mathbf{D}$ is a diagonal matrix with the i-th diagonal element as $d_{ii} = 1 \big/ 2 \left\| \left(\mathbf{B}^{(2)}\right)^i \right\|_2$, and $\tilde{\mathbf{D}}$ is also a diagonal matrix with the i-th diagonal element as $\tilde{d}_{ii} = 1 \big/ 2 \left| \left(\mathbf{B}^{(2)}\right)_{i,j} \right|$.

In order to simplify the optimization problem (22), we set $\mathbf{X}_1 = \mathbf{X}\mathbf{D}^{-1/2}$ and $\mathbf{B}_1 = \mathbf{D}^{1/2}\mathbf{B}^{(2)}$. Then, the problem (22) can be equivalent to the following minimization problem:

$$\min_{\mathbf{B}_1} \left\| \mathbf{Y}^{(2)} - \mathbf{X}_1\mathbf{B}_1 \right\|_F^2 + \lambda_1 \mathrm{Tr}\left[\left(\mathbf{B}_1\right)^T \mathbf{D}^{-1/2} \tilde{\mathbf{D}} \mathbf{D}^{-1/2}\mathbf{B}_1 \right] + \lambda_2 \mathrm{Tr}\left[\left(\mathbf{B}_1\right)^T \mathbf{B}_1 \right]. \tag{23}$$

The objective function is equal to the following equation:

$$\mathrm{J}_2\left(\mathbf{B}_1\right) = \left\| \mathbf{Y}^{(2)} - \mathbf{X}_1\mathbf{B}_1 \right\|_F^2 + \lambda_1 \mathrm{Tr}\left[\left(\mathbf{B}_1\right)^T \mathbf{D}^{-1/2} \tilde{\mathbf{D}} \mathbf{D}^{-1/2}\mathbf{B}_1 \right] + \lambda_2 \mathrm{Tr}\left[\left(\mathbf{B}_1\right)^T \mathbf{B}_1 \right]. \tag{24}$$

And then we compute the derivative of $\mathrm{J}_2\left(\mathbf{B}_1\right)$, then set its result to zero, we have:

$$\frac{\partial \mathrm{J}_2\left(\mathbf{B}_1\right)}{\partial \mathbf{B}_1} = 2\left[\left(\mathbf{X}_1\right)^T \mathbf{X}_1 + \lambda_2\mathbf{I} + \lambda_1\mathbf{D}^{-1/2} \tilde{\mathbf{D}} \mathbf{D}^{-1/2} \right]\mathbf{B}_1 - 2\left(\mathbf{X}_1\right)^T \mathbf{Y}^{(2)} \tag{25}$$
$$= 0.$$

So, we can obtain:

$$\mathbf{B}_1 = \left[\left(\mathbf{X}_1^T\mathbf{X}_1 + \lambda_2\mathbf{I} + \lambda_1\mathbf{D}^{-1/2} \tilde{\mathbf{D}} \mathbf{D}^{-1/2} \right) \right]^{-1} \mathbf{X}_1^T\mathbf{Y}^{(2)}. \tag{26}$$

Therefore, we finally compute $\mathbf{B}^{(2)}$ by optimizing $\mathbf{B}_1$, that is, $\mathbf{B}^{(2)} = \mathbf{D}^{-1/2}\mathbf{B}_1$. We sum up the $L_1L_{2,1}$-iPaD method as the Algorithm 1. Besides, in this paper, we also use the soft-impute algorithm to deal with the missing values problems in solving $\mathbf{X}$. The detailed steps can be found from the iPaD [8] and $L_{2,1}$-iPaD methods [12].

Algorithm 1: The alternating updating algorithm for the $L_1L_{2,1}$-iPaD method

---

Data Input: $\mathbf{Y}^{(1)}$, $\mathbf{Y}^{(2)}$, $\mathbf{H}^{(1)}$
Parameter: λ_1, λ_2
Output: $\mathbf{B}^{(2)}$

Initialization: set $\mathbf{B}^{(1)} = \mathbf{H}^{(1)}$ and set $\mathbf{B}^{(2)} = 0$.
Optimization:
(1). Optimize $\mathbf{X}$:
$$\mathbf{X} = \arg\min_{\mathbf{X}} \left\| \mathbf{Y} - \mathbf{XB} \right\|_F^2$$
$$s.t. \sum_i \mathbf{X}_{i,j}^2 \leq 1, \forall j = 1, \ldots, K,$$
where $\mathbf{Y} = \left[\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)} \right]$ and $\mathbf{B} = \left[\mathbf{B}^{(1)}, \mathbf{B}^{(2)} \right]$.
(2). Optimize $\mathbf{B}^{(1)}$:
$$\mathbf{B}^{(1)} = \arg\min_{\mathbf{B}^{(1)}} \left\| \mathbf{Y}^{(1)} - \mathbf{XB}^{(1)} \right\|_F^2$$
$$s.t. \ \mathbf{B}_{i,j}^{(1)} = 0, \forall (i,j) : \mathbf{H}_{i,j}^{(1)} = 0.$$
(3). Optimize $\mathbf{B}^{(2)}$:
$$\mathbf{B}^{(2)} = \arg\min_{\mathbf{B}^{(2)}} \left\| \mathbf{Y}^{(2)} - \mathbf{XB}^{(2)} \right\|_F^2 + \lambda_1 \left\| \mathbf{B}^{(2)} \right\|_1 + \lambda_2 \left\| \mathbf{B}^{(2)} \right\|_{2,1}.$$
(4). Repeat step (1), (2) and (3) until convergence.

Parameter selection and significance test

Compared with the iPaD and $L_{2,1}$-iPaD methods, our new $L_1L_{2,1}$-iPaD method has two adjustable parameters (λ_1 and λ_2), which can control the sparsity of the drug-pathway association matrix $\mathbf{B}^{(2)}$. In this paper, We perform five-fold cross-validation to find the suitable parameters. In the five-fold cross-validation, we divide the matrix $\mathbf{Y}^{(2)}$ into 5 folds. At each round of the cross-validation, we make each of the 5 folds as missing values, and the rest of 4 folds as training data. It is obvious that it is difficult to find the optimal parameters λ_1 and λ_2, simultaneously. Therefore, we solve the $L_1L_{2,1}$-iPaD problem to produce a λ_1 sequence by fixing λ_2, and then we use five-fold cross-validation to find an optimal parameter λ_1. We treat the minimum residual sum of squares (RSS) corresponding to the λ_1 value as the optimal value. Similarly, we find an optimal parameter λ_2 by fixing λ_1. Since our method is a sparse optimization algorithm, there are many zero coefficients in the matrix $\mathbf{B}^{(2)}$. We usually treat those nonzero coefficients as the identified core drug-pathway association pairs. After searching for the optimal parameters, we perform permutation test [27] to assess the significance of the coefficients. We first obtain the estimated value of the matrix $\mathbf{B}^{(2)}$, and then compute the P-value of the coefficient for the matrix $\mathbf{B}^{(2)}$, the computational formula is as follows:

$$P_{i,j} = \frac{1}{T} \sum_{t=1}^{T} \left(\left| \tilde{\mathbf{B}}_{i,j}^{(2)(t)} \right| \geq \left| \hat{\mathbf{B}}_{i,j}^{(2)} \right| \right), \qquad (27)$$

where $\tilde{\mathbf{B}}_{i,j}^{(2)(t)}$ denotes the t-th permutation estimated values of the matrix $\mathbf{B}^{(2)}$, and T is the total number of permutations, $\hat{\mathbf{B}}_{i,j}^{(2)}$ is the estimated values of the matrix $\mathbf{B}^{(2)}$ in the original data.

CONFLICTS OF INTEREST

There is no conflicts of interest.

REFERENCES

1. Yang Y, Dai C, Cai Z, Hou A, Cheng D, Wu G, Li J, Cui J, Xu D. The pathway analysis of micrornas regulated drug-resistant responses in HeLa cells. IEEE Trans Nanobioscience. 2016; 15:113-118.

2. Ma H, Zhao H. FacPad: Bayesian sparse factor modeling for the inference of pathways responsive to drug treatment. Bioinformatics. 2012; 28:2662-2670.

3. Hong Y, Chu Q, Ying HL, Lin T, Jin Z, Yu CY, Feng X, Zhe C, Feng Z, Yu ZC. Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. Nucleic Acids Res. 2015; 44:D1069-D1074.

4. Chen X, Yan CC, Zhang X, Zhang X, Dai F, Yin J, Zhang Y. Drug-target interaction prediction: databases, web servers and computational models. Brief Bioinform. 2016; 17:696-712.

5. Ma H, Zhao H. iFad: an integrative factor analysis model for drug-pathway association inference. Bioinformatics. 2012; 28:1911-1918.

6. Ma H, Zhao H. Drug target inference through pathway analysis of genomics data. Adv Drug Deliv Rev. 2013; 65:966-972.

7. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545-15550.

8. Li C, Yang C, Hather G, Liu R, Zhao H. Efficient drug-pathway association analysis via integrative penalized matrix decomposition. IEEE/ACM Trans Comput Biol Bioinform. 2016; 13:531-540.

9. Shankavaram UT, Varma S, Kane D, Sunshine M, Chary KK, Reinhold WC, Pommier Y, Weinstein JN. CellMiner:

a relational database and query tool for the NCI-60 cancer cell lines. BMC Genomics 2009; 10:324-331.

10. Gasparini M. Markov chain monte carlo in practice. International Encyclopedia of the Social & Behavioral Sciences. 1997; 2:9236-9240.

11. Liu JX, Wang D, Gao YL, Zheng CH, Shang JL, Liu F, Xu Y. A joint-L2,1 -norm-constraint-based semi-supervised feature extraction for RNA-seq data analysis. Neurocomputing. 2016; 228:263-269.

12. Wang D, Zheng C, Gao Y, Liu J, Wu S, Shang J. L21-iPaD: an efficient method for drug-pathway association pairs inference. IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 15-18 December 2016, Shenzhen, China, pp. 664-669.

13. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 2012; 40:D109-D114.

14. Secchiero P, Melloni E, Di IM, Tiribelli M, Rimondi E, Corallini F, Gattei V, Zauli G. Nutlin-3 up-regulates the expression of Notch1 in both myeloid and lymphoid leukemic cells, as part of a negative feedback antiapoptotic mechanism. Blood. 2009; 113:4300-4308.

15. Airiau K, Mahon FX, Josselin M, Jeanneteau M, Turcq B, Belloc F. ABT-737 increases tyrosine kinase inhibitor–induced apoptosis in chronic myeloid leukemia cells through XIAP downregulation and sensitizes CD34+ CD38− population to imatinib. Exp Hematol. 2012; 40:367-378.e2.

16. George P, Bali P, Annavarapu S, Scuto A, Fiskus W, Guo F, Sigua C, Sondarva G, Moscinski L, Atadja P, Bhalla K. Combination of the histone deacetylase inhibitor LBH589 and the hsp90 inhibitor 17-AAG is highly active against human CML-BC cells and AML cells with activating mutation of FLT-3. Blood. 2005; 105:1768-1776.

17. Mignani S, El Brahmi N, El Kazzouli S, Eloy L, Courilleau D, Caron J, Bousmina MM, Caminade AM, Cresteil T, Majoral JP. A novel class of ethacrynic acid derivatives as promising drug-like potent generation of anticancer agents with established mechanism of action. Eur J Med Chem. 2016; 122:656-673.

18. Aggarwal BB. Signalling pathways of the TNF superfamily: a double-edged sword. Nat Rev Immunol. 2003; 3:745-756.

19. Heinschink A, Raab M, Daxecker H, Griesmacher A, Müller MM. *In vitro* effects of mycophenolic acid on cell cycle and activation of human lymphocytes. Clin Chim Acta. 2000; 300:23-28.

20. Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc. 1996; 58:267-288.

21. Ding C, Zhou D, He X, Zha H. R1-PCA: rotational invariant L1-norm principal component analysis for robust subspace factorization. Proceedings of the 23rd international conference on Machine learning: ACM, 25-29 June 2006, New York, USA. pp. 281-288

22. Huang J, Nie F, Huang H, Ding C. Robust manifold nonnegative matrix factorization. ACM TKDD. 2014; 8:1-21.

23. Wang H, Nie F, Huang H. Multi-view clustering and feature learning via structured sparsity. International Conference on Machine Learning, pp. 352-360.

24. Zhou Z, Wang Y, Wu QM, Yang CN, Sun X. Effective and Efficient Global Context Verification for image copy detection. IEEE Trans Inf Forensic Secur. 2016; 12:48-63.

25. Nie F, Huang H, Cai X, Ding CH. Efficient and robust feature selection via joint ℓ2,1-norms minimization. Advances in neural information processing systems 23: Conference on Neural Information Processing Systems 2010 Proceedings of A Meeting Held 6-9 December 2010, Vancouver, British Columbia, Canada, pp. 1813-1821.

26. Nesterov Y. A method of solving a convex programming problem with convergence rate O (1/k2). Soviet Mathematics Doklady; 1983. pp. 372-376.

27. Oden A, Wedel H. Arguments for Fisher's permutation test. Ann Stat. 1975; 3:518-520.

28. Weisberg E, Kung AL, Wright RD, Moreno D, Catley L, Ray A, Zawel L, Tran M, Cools J, Gilliland G. Potentiation of antileukemic therapies by Smac mimetic, LBW242: effects on mutant FLT3-expressing cells. Mol Cancer Ther. 2007; 6:1951-1961.

29. Netzer WJ, Dou F, Cai D, Veach D, Jean S, Li Y, Bornmann WG, Clarkson B, Xu H, Greengard P. Gleevec inhibits β-amyloid production but not Notch cleavage. Proc Natl Acad Sci U S A. 2003; 100:12444-12449.

30. Eng C, Kopetz S, Morris J, Malik Z, Stewart D, Chang H, Ohinata A, Abbruzzese J, Gallick G. Phase II study of the novel oral Src-kinase inhibitor, AZD0530, in previously treated advanced colorectal cancer patients. Cancer Res. 2008; 68.

31. Smolen GA, Sordella R, Muir B, Mohapatra G, Barmettler A, Archibald H, Kim WJ, Okimoto RA, Bell DW, Sgroi DC. Amplification of MET may identify a subset of cancers with extreme sensitivity to the selective tyrosine kinase inhibitor PHA-665752. Proc Natl Acad Sci U S A. 2006; 103:2316-2321.

32. Kang CD, Yoo SD, Hwang BW, Kim KW, Kim DW, Kim CM, Kim SH, Chung BS. The inhibition of ERK/MAPK not the activation of JNK/SAPK is primarily required to induce apoptosis in chronic myelogenous leukemic K562 cells. Leuk Res. 2000; 24:527-534.

33. Ball DW, Jin N, Rosen DM, Dackiw A, Sidransky D, Xing M, Nelkin BD. Selective growth inhibition in BRAF mutant thyroid cancer by the mitogen-activated protein kinase kinase 1/2 inhibitor AZD6244. J Clin Endocrinol Metab. 2007; 92:4712-4718.

34. Jia H. (2009). ZD6474 inhibits Src kinase Leading to Anti-tumoral Effects In imatini-resistant cells. Zhongshan University (in Chinese).

35. Li Z, Xu M, Xing S, Ho WT, Ishii T, Li Q, Fu X, Zhao ZJ. Erlotinib effectively inhibits JAK2V617F activity and polycythemia vera cell growth. J Biol Chem. 2007; 282:3428-3432.

36. Markasz L, Stuber G, Vanherberghen B, Flaberg E, Olah E, Carbone E, Eksborg S, Klein E, Skribek H, Szekely L. Effect of frequently used chemotherapeutic drugs on the cytotoxic activity of human natural killer cells. Mol Cancer Ther. 2007; 6:644-654.

37. Tripathy S, Roy S. Redox sensing and signaling by malaria parasite in vertebrate host. J Basic Microbiol. 2015; 55:1053-1063.

38. Fioredda F, Calvillo M, Bonanomi S, Coliva T, Tucci F, Farruggia P, Pillon M, Martire B, Ghilardi R, Ramenghi U. Congenital and acquired neutropenia consensus guidelines on diagnosis from the Neutropenia Committee of the Marrow Failure Syndrome Group of the AIEOP (Associazione Italiana Emato-Oncologia Pediatrica). Am J Hematol. 2011; 57:10-17.

39. Merhi F, Auger J, Rendu F, Bauvois B. Allium compounds, dipropyl and dimethyl thiosulfinates as antiproliferative and differentiating agents of human acute myeloid leukemia cell lines. Biologics. 2008; 2:885-895.

40. Kinoshita K, Aono Y, Azuma M, Kishi J, Takezaki A, Kishi M, Makino H, Okazaki H, Uehara H, Izumi K. Antifibrotic effects of focal adhesion kinase inhibitor in bleomycin-induced pulmonary fibrosis in mice. Am J Respir Cell Mol Biol. 2013; 49:536-543.

41. Boengler K, Konietzka I, Buechert A, Heinen Y, Garciadorado D, Heusch G, Schulz R. Loss of ischemic preconditioning's cardioprotection in aged mouse hearts is associated with reduced gap junctional and mitochondrial levels of connexin 43. Am J Physiol Heart Circ Physiol. 2007; 292:H1764-H1769.

42. Edson KZ, Prasad B, Unadkat JD, Suhara Y, Okano T, Guengerich FP, Rettie AE. Cytochrome P450-dependent catabolism of vitamin K: ω-hydroxylation catalyzed by human CYP4F2 and CYP4F11. Biochemistry. 2013; 52:8276-8285.