

# 基于高通量数据的人体外源性 植物 miRNA 跨界调控建模

张 浩<sup>1,2</sup>, 占萌苹<sup>2,3</sup>, 郭刘香<sup>2,3</sup>, 李 誌<sup>4</sup>, 刘元宁<sup>1,2</sup>,  
张春鹤<sup>1,2</sup>, 常浩武<sup>1,2</sup>, 王志强<sup>1,2</sup>

(1. 吉林大学 计算机科学与技术学院, 长春 130012; 2. 吉林大学 符号计算与知识工程教育部重点实验室, 长春 130012; 3. 吉林大学 软件学院, 长春 130012; 4. 长春理工大学 应用技术学院, 长春 130022)

**摘 要:**以基于人体血清的高通量测序数据为研究对象,采用生物计算方法对数据进行分析,以目前已知的人类和植物的 RNA 数据作为参照集,从中筛选出可能的人体外源植物 miRNA;基于 miRNA 的调控机制原理,获得可能的植物 miRNA 干扰人类 mRNA 的靶基因;构建植物 miRNA-人类 mRNA 调控网络模型,分析可能的跨界调控作用。本文通过构建植物-人体跨界调控模型,共挖掘出 9 个核心模块。研究结果表明,外源性植物 miRNA 参与调控人体胺代谢、酒精代谢以及脂肪酸代谢等生物过程。

**关键词:**人工智能;高通量数据;外源性植物 miRNA;跨界调控

**中图分类号:**TP399 **文献标志码:**A **文章编号:**1671-5497(2018)04-1206-08

**DOI:**10.13229/j.cnki.jdxbgxb20170521

## Human exogenous plant miRNA cross-kingdom regulatory modeling based on high-throughput data

ZHANG Hao<sup>1,2</sup>, ZHAN Meng-ping<sup>2,3</sup>, GUO Liu-xiang<sup>2,3</sup>, LI Zhi<sup>4</sup>, LIU Yuan-ning<sup>1,2</sup>,  
ZHANG Chun-he<sup>1,2</sup>, CHANG Hao-wu<sup>1,2</sup>, WANG Zhi-qiang<sup>1,2</sup>

(1. College of Computer Science and Technology, Jilin University, Changchun 130012, China; 2. Symbol Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China; 3. College of Software, Jilin University, Changchun 130012, China; 4. College of Applied Technique, Changchun University of Science and Technology, Changchun 130022, China)

**Abstract:** Exogenous miRNAs are cross-kingdom regulatory in bacteria and viruses, but whether exogenous plant miRNAs are stable in the human body and are involved in cross-kingdom regulation is still controversial. Only by cross-border modeling on the basis of a comprehensive analysis of the human high-throughput miRNA data, and design of effective biological experiments, the human exogenous plant miRNA presence, access pathways and regulatory role can be revealed. In this paper, high-throughput sequencing data based on human serum were used as the study objective and were analyzed by bio-calculation method. The possible human exogenous plant miRNAs were screened by

收稿日期:2017-04-27.

基金项目:吉林省科技发展计划项目(20150101056JC);中国博士后科学基金项目(2015M570273).

作者简介:张浩(1971-),男,教授,博士.研究方向:生物信息学. E-mail:zhangh@jlu.edu.cn

通信作者:刘元宁(1962-),男,教授,博士生导师.研究方向:生物信息学. E-mail:liuyn@jlu.edu.cn

using the RNA data of human and plant as the reference group. Based on the principle of miRNA regulatory mechanism, the possible target genes of plant miRNAs interfering with human mRNA were obtained. A network model of plant miRNA-human mRNA was constructed to analyze possible cross-kingdom regulation. Nine core modules were excavated by constructing a plant-human cross-kingdom regulation model. The results show that exogenous plant miRNAs are involved in the regulation of biological process such as human amine metabolism, alcohol metabolism and fatty acid metabolism.

**Key words:** artificial intelligence; high throughput data; exogenous plant miRNA; cross-species regulation

## 0 引言

miRNA 是由 21~23 个核苷酸组成的非编码小分子 RNA<sup>[1]</sup>。它通过与 mRNA 碱基互补配对结合的方式,降解靶基因 mRNA 或者抑制靶基因 mRNA 的翻译,对蛋白水平的表达产生影响,在生物体细胞的生长、分化、凋亡过程中起到重要的调控作用<sup>[2]</sup>。自 miRNA 发现以来,其相关研究主要集中于生物体内源性 miRNA 的表达和调控机制,包括 miRNA 与肿瘤、复杂疾病、流感病毒等的发生和发展关系上,以及在其他免疫系统之间调控关系的研究<sup>[3]</sup>。随着研究的深入,研究人员发现 miRNA 的调控作用不仅局限于生物体内源,还可以作为一种信号分子在不同的生命体间进行传导,对异源细胞的 mRNA 表达进行调控。现已证实细菌、病毒的 miRNA 能够在不同宿主间进行传导,并对宿主 mRNA 的调控产生影响<sup>[2]</sup>。真菌及病毒跨界调控的发现为 RNA 干扰研究领域提出了新的科学问题,如 miRNA 的跨界调控是否普遍存在;如果存在,外源性 miRNA 是如何进入宿主体内并参与宿主体内调控;内源性 miRNA 与外源性 miRNA 调控机制有何相关性等。

近期研究还发现,外源性植物 miRNA 很可能通过某种途径进入到人体中,在人体中稳定存在,并且参与其相关生理调控<sup>[4,5]</sup>。Pastrello 等<sup>[5]</sup>实验发现芸苔属蔬菜中含有的 miR160 致使 MTHFD2 作为转录物的含量减少;Zhou 等<sup>[6]</sup>发现金银花中的 miR2911 可以抑制流感病毒的 PB1 和 NS 蛋白的合成,有效地减轻 H1N1、H5N1 和 H7N9 亚型流感病毒感染后的症状。Zhang 等<sup>[7]</sup>发现植物 miR168a 可以与人体的相关靶基因结合,阻碍 LDLRAP1(低密度脂蛋白受体适配蛋白 1)的转录,从而阻碍血浆中低密度脂

蛋白 LDL 的清除。这些研究让科学家们意识到,外源性植物 miRNA 对人体也可能存在类似的跨界调控。

由此可见,外源性植物 miRNA 可能会对动物产生跨物种调控作用,但外源性 miRNA 与内源性 miRNA 在调控机制上有何关联尚需要进一步研究。目前,对于在人体检测到的植物 miRNA,是通过饮食等方式进入人体还是由于实验污染所致还未有定论。外源性植物 miRNA 是否能像内源 miRNA 一样具有调控作用也需进一步研究。然而,由于现有可能的外源性植物 miRNA 均是通过生物实验得到的,数量少不具有普遍性、且实验结果各异,严重制约了植物 miRNA 跨界调控方面的研究。

近年来,高通量测序技术的迅速发展为解决上述问题提供了数据基础和更好的解决方案。人和动物高通量测序数据的产生、生物计算方法的进步使得全面分析植物 miRNA 的跨界调控成为可能。

本文通过对收集的 miRNA 高通量测序数据进行筛选分析,找出可能存在的外源性植物 miRNA,对其进行靶基因预测,挖掘出核心模块,构建植物 miRNA-人体 miRNA 跨界调控模型,进而研究外源性植物 miRNA 在人体中可能存在的调控作用。

## 1 本文方法模型

首先,以 NCBI GEO 中基于人体血清的高通量测序数据作为研究对象,采用生物计算方法对数据进行分析,将目前已知的人类和植物的 miRNA 数据作为参照集,从中筛选出可能的人体外源植物 miRNA。其次,基于 miRNA 的调控机制,采用分级打分方法预测可能的植物 miRNA 干扰人类 mRNA 的靶基因。最后,构建植物

miRNA-人类 mRNA 调控网络,分析可能的调控

作用。具体流程如图 1 所示。

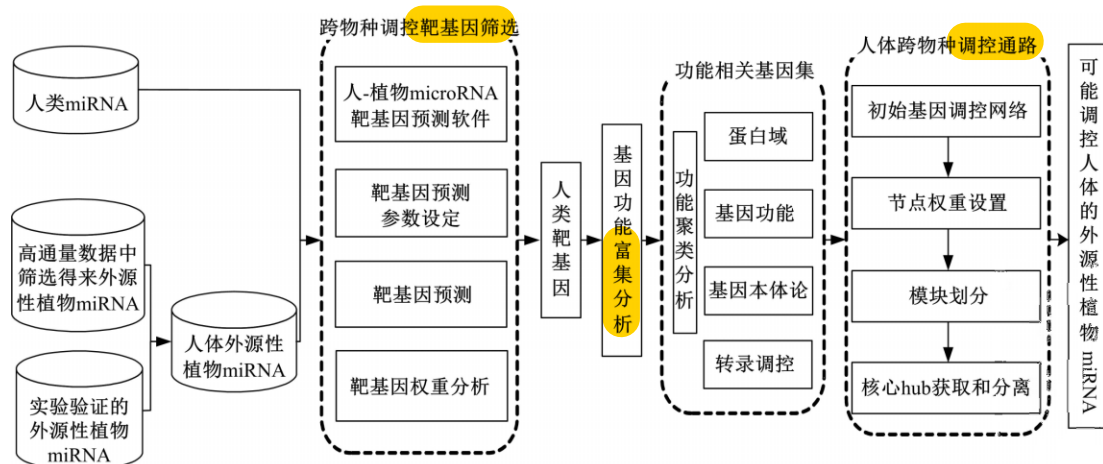


图 1 外源性植物 miRNA 功能预测模型

Fig. 1 Prediction model of exogenous plant miRNA function

### 1.1 基于高通量数据筛选外源性植物 miRNA

在很多疾病的发生过程中,miRNA 在血清、血浆等体液中都异常表达,参与生物体的调控,并成为评估身体健康状态的生物标志物。外源性 miRNA 在人体血清、血浆和其他无细胞样品中被发现,有可能在参与人体生命活动时发挥重要的调控作用<sup>[8]</sup>。由于实验过程中检测到的外源性植物 miRNA 都是在血清组织中,因此本文采用人体血清组织的高通量测序数据作为研究对象。

本文从 NCBI GEO 下载人体血浆的高通量小 RNA 测序数据,在 miRBase 数据库下载植物和人的 miRNA 数据,在 Silva rRNA 数据库下载人类的 rRNA 数据,在 Ensembl 数据库下载 ncRNA 数据,在 Gencode 数据库下载 mRNA 数据。本文在 miRBase 下载包括 ath(拟南芥),osa(水稻),tae(小麦)和 sly(番茄)的 71 种植物的所有 miRNA 数据作为标准比对数据集。

#### (1) 高通量数据预处理

由于实验原因,高通量测序数据中会包含有许多不需要的数据,如低复杂性的 reads 和接头会导致错误的组装、PCR 扩增中产生的 reads 重复等。这些片段常出现在长片段插入的文库中,影响高通量测序数据中成对 read (mate-pair) 的统计量。因此,本文需要对数据进行预处理来获

取合适的数据库。

从测序数据中去除衔接子、测序接头、来自 PCR 扩增的几乎大量相同 reads、制备文库中插入的 DNA 片段以及 poly A 的人工污染物。

#### (2) 筛选有效数据

对高通量测序数据而言,低质量的数据可能存在测序上的错误,因此需要通过质量分数和  $k$ -mer 频率修正初步筛选的数据。首先通过脚本将测序数据中低质量分数过滤;然后过滤出长度小于 17 nt 和大于 26 nt 的读数(17 和 26 是 miRBase 中最短和最长的植物 miRNA 的长度);最后得到有效的数据。

#### (3) 挖掘人体外源性植物 miRNA

首先,将获得的合格数据与人体 miRNA、rRNA、tRNA 进行 blast 比对,在比对过程中允许两个错配,比对之后将人体的这些小 RNA 删除;其次,将未确认的数据与人类其他 RNA(包括 ncRNA 和 mRNA)进行比对,这次比对中允许一个错配,并将匹配的数据删除;最后,将剩下的数据与植物 miRNA 比对,匹配过程中等于或小于一个不匹配,获得人体所有可能的植物外源 miRNA 数据。在这个过程中,使用 bowtie 软件进行序列比对。植物 miRNA 筛选流程如图 2 所示。

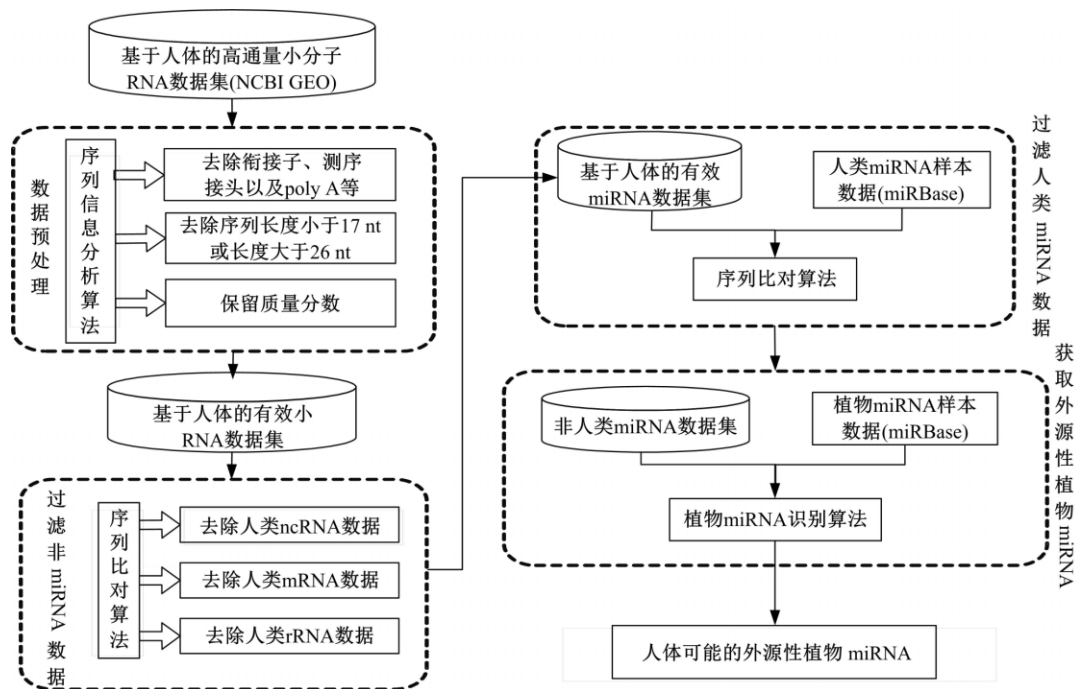


图2 基于高通量数据挖掘外源性植物 miRNA

Fig. 2 Based on high-throughput data mining exogenous plant miRNA

### 1.2 外源性植物 miRNA 靶基因预测

本文基于动物体内源性 miRNA 调控原理,考虑多靶点,同时引入结构信息,确定可能的 miRNA-mRNA 关联关系。基本规则如下:

(1) 将 miRNA 5' 端的第 1~8 个核苷酸作为种子区域。

(2) 若 miRNA 与靶标结合的位点位于 mRNA 的茎区上,则认为该位点成为 miRNA 的靶标作用点的可能性大。

(3) 基于多步打分的算法 miRScore 来寻找与植物 miRNA 对应的人类 mRNA 靶基因。

miRScore 具体算法如下:

(1) 针对人体 mRNA 序列,基于人体外源植物 miRNA 共同的起始结合位点打分。

本文采用滑动窗口方法完成此步。滑动窗口大小为 8,步长为 1。当 miRNA 的种子区域与序列完全匹配时,记录下该起始结合位点。打分规则如下:首先对于同一条序列的同一个位点,为种子区域能够与从该位点起的 8 个核苷酸完全匹配的 mRNA 加 5 分;其次考虑 miRNA 除种子区域外的其他核苷酸,当 miRNA 的核苷酸与 mRNA 的核苷酸匹配时,给该组 miRNA-mRNA 加 5 分。

(2) 基于全部序列的共同二级结构打分;

根据在第(1)步中已经得到的每个 miRNA

在序列上的起始结合位点,以 miRNA 的长度为基准打分,打分规则如下:如果 miRNA 与靶 mRNA 相结合的核苷酸在共同二级结构的茎上,那么给该组 miRNA-mRNA 加 10 分。

(3) 计算 miRNA-mRNA 组分数的平均值,分数大于平均值的为结果,其中分数最高的即为最有可能存在调控关系。如果具有相同分数的 miRNA 多于一个,那么使用 Vienna RNA 软件包中的 RNAfold,结合能最低的 miRNA-mRNA 组为最终结果。

### 1.3 LeaderRank 算法选取调控网络核心节点

由于靶基因数量较多,因此核心节点选择的算法尤为重要。在调控网络节点计算权值时常用 PageRank 算法<sup>[9]</sup>。其主要思想是在整个网络中,初始时赋值给各节点相同的值,采用二维矩阵迭代相乘,直到网络中全部的节点分值达到稳定。在该算法中,从一个节点出发,访问其他节点的概率相等,即每个节点的随机跳转概率是相同的。然而在调控网络中,信息少的节点访问信息多的节点的概率小得多。PageRank 算法具体公式如下:

$$PR_i(t) = (1 - c) \sum_{j=1}^n a_{ji} \frac{PR_j(t-1)}{k_j^{\text{out}}} + \frac{c}{n} \quad (1)$$

由式(1)可知,PageRank 算法中涉及到随机跳转概率  $c$  的选取,在不一样的条件下参数也不



具有通用性。通过比较分析,本文采用 LeaderRank 算法计算调控网络节点权值。它能够很好地解决上述两个问题。本文通过在网中加了一个背景节点还有该节点与其他全部节点的边来取代 PageRank 算法中的跳转概率,得出没有参数并且更简单的一个算法。在 LeaderRank 算法中,通过一个节点的概率与该节点浏览背景节点的概率是一样的,当一个节点信息量大时,本算法使其访问背景节点的概率与节点信息量成反比。在开始阶段除了背景节点  $v_g$  之外赋值给其他节点一个数值,即  $LR_i(0) = 1, \forall i \neq g$ ;  $LR_g(0) = 0$ 。

直到使用下述公式迭代达到稳定状态:

$$LR_i(t) = \sum_{j=1}^{n+1} \frac{a_{ji}}{k_j^{\text{out}}} LR_j(t-1) \quad (2)$$

特别地,式(2)中邻接矩阵是  $n+1$  阶(包含背景节点)。达到稳定状态时把背景节点的分值  $LR_g(t_c)$  平均分配给其他  $n$  个节点,因此也就得到了节点  $v_i$  用 LeaderRank 算法打分的值:

$$LR_i = LR_i(t_c) + \frac{LR_g(t_c)}{n} \quad (3)$$

LeaderRank 算法在用来表示网络中节点重要性等方面有不错的结果<sup>[10]</sup>,研究发现,LeaderRank 相比于 PageRank 收敛更快;也可以从网络中更好地发现重要性大的节点,挖掘出这些有影响力的节点,同时还有更强的鲁棒性。这些都使 LeaderRank 算法应用很广。

#### 1.4 挖掘核心节点为中心的模块

现有研究显示,在基因网络中一个基因通常能够与其周围互相作用的基因形成基因模块,从而发挥其特定的生物学功能<sup>[11]</sup>。上述过程中计算了基因调控网络中的节点权值,并根据 LR 权值排序各个节点,将权值较高的选择出来作为核心节点研究,并挖掘核心网络模块。接下来本文要设计以核心节点为中心的核心调控网络模块挖掘方法,找到网络中的核心调控模块,并对挖掘出的模块进行相关分析,进而探究其参与的调控通路和产生的主要功能。

调控网络是由基因及其之间的相互作用组成。在网络中节点可以是 DNA、mRNA、miRNA 以及蛋白等活性分子,边则是这些活性分子间存在的相互作用,通常具有特定的生物学含义。此外因基因间的作用是相互的,所以选择无向图。建立网络的关键是找到基因之间的相互关系,从

而组织成基因调控网络。本文以  $V$  为节点(基因), $E$  为边(基因间的相互作用),用  $G = (V, E)$  表示二者之间的关系。核心模块的挖掘就是在图  $G$  中寻找包含核心基因节点  $V_k$  的子图  $G_k = (V_k, E_k)$ ,子图  $G_k$  符合的条件如下:

(1)与核心节点  $V_k$  相邻的节点数必须大于 1,也就是若存在节点与核心节点连接,那此类节点就不止一个,即核心节点的度必须比 1 大。

(2)核心节点  $V_k$  与节点  $V_i$  之间存在边  $E_{ki}$ ,即  $G_k$  中除核心节点外的节点  $V_i$  和核心节点  $V_k$  一定直接相邻。

(3)核心节点相邻关系的传递性:如果  $G_k$  中核心节点  $V_k$  与节点  $V_j$  相邻,那么  $V_j$  与其他和  $V_k$  相邻的节点  $V_i$  一定相邻,即若  $V_k$  与  $V_j$  存在边  $E_{kj} = (V_k, V_j)$ , $V_k$  与  $V_i$  存在边  $E_{ki} = (V_k, V_i)$ ,那么也就一定存在边  $E_{ji} = (V_j, V_i)$ 。

在调控网络中,挖掘以核心节点为中心的核心网络模块,具体步骤如下:

(1)导入数据,包含选取的核心节点、调控网络节点、节点的边对应关系的第 1 列以及第 2 列。

(2)创建模型 CoreModule 类,在类中以一维数组 Array 代表节点,以二维数组 Matrix 来表示节点以及节点存在联系,同时定义了初始值为 0 的一维判定数组 Judge,用来判定核心节点是否在其他核心模块中出现过。

(3)二维矩阵 Matrix 中根据两节点间是否存在关系进行初始化,1 表示两个节点之间存在关系;0 表示二者不存在关系。

(4)载入核心节点 Vcore,首先判断该核心节点与  $\text{Array}[i] (i=1, 2, \dots, n)$  是否相等,同样的  $\text{Judge}[i] (i=1, 2, \dots, n)$  是否等于 0,来确认该核心节点是否已经在其他核心模块中出现过。若核心节点与  $\text{Array}[i]$  相等且  $\text{Judge}[i]$  等于零,则计算核心节点的度。若度大于 1,向量 Vector1 中添加从数组 Node 中的位置开始的所有与核心节点直接相邻的节点。为了保证  $A > B$  和  $B > A$  的输出只有一个,为向量 Vector1 的值设置了标记,即定义了一维数组 Mark,其初始值设为 0。

(5)设置标记位 tag\_1,初始值为 0,对 Vcore 的每个邻接点  $\text{Array}[x]$  的处理如下:

(a)计算邻接点  $\text{Array}[x]$  的度,如果其度大于 1,那么在向量 Vector2 中添加  $\text{Array}[x]$  的直接邻接点。

(b)判断向量 Vector1 和 Vector2 中存放的

值是否相等,如果相等,获取节点在数组 Array 中的位置  $y$ ,把 tag\_1 值设为 2(输出标识);同时 Vector1 对应的位置为  $d$ ,若 tag[d] 不等于 1,就输出  $\text{Array}[x] > \text{Array}[y]$  节点关系,并把 Judge[x] 设为 1(表示节点  $\text{Array}[x]$  已经存在某核心模块中),并且把 tag[k] 赋值等于 1,以此来保证  $A > B$  和  $B > A$  只输出一个。

(c) 判断 tag\_1 的值是否等于 2,如果 tag\_1 的值等于 2,则代表邻接点模块已经全部输出,则需输出核心节点满足  $V_{\text{core}} > \text{Array}[x]$ 。

(6) 重复执行步骤(4)处理完所有核心节点。

## 2 研究结果

本文提取了实验验证的 531 个过滤靶点的相互作用和信号网络,然后创建了包含 782 个基因和 2444 个相互作用的主要调节网络。采用 LeaderRank 算法计算调控网络中 531 个节点的权值,并且根据权值大小选取出调控网络中的核心节点,从而挖掘核心模块。对网络中的节点进行计算后,按照权值大小排序,部分节点的 LR 值如下表 1 所示:

表 1 节点分值表  
Table 1 Node Score Table

节点名	LR 值	节点名	LR 值
AXIN1	3.5600379	C17orf62	3.0125687
HOXB1	3.4388115	DNAJB2	2.9904926
ATP5D	3.42415	C7orf43	2.9609702
MAPK8IP2	3.3535602	CLPTM1	2.94651
ZAP70	3.2993844	NFKBIL1	2.93037
MTNR1B	3.210638	EXOC3L2	2.9290342
GPR35	3.1470814	CLSTN3	2.9201922
GYS1	3.1165192	CABIN1	2.8639722

本文选取其中分值最高的前 9 位 (AXIN1, HOXB1, ATP5D, MAPK8IP2, ZAP70, MTNR1B, GPR35, GYS1, C17orf62) 作为核心节点进行研究。以下将挖掘以这 9 个核心节点为中心的核心基因调控网络模块。

本文在选出核心节点后,挖掘以核心节点为中心的周围与其相互作用的基因节点形成的核心模块,然后针对核心基因网络模块进行 GO 功能富集分析,探究其在人体中存在的生物学意义。部分核心网络模块如图 4 和图 5 所示。

根据功能富集分析的结果,对各模块进行分析,部分结果如下:AXIN1 涉及核苷三磷酸酶调节

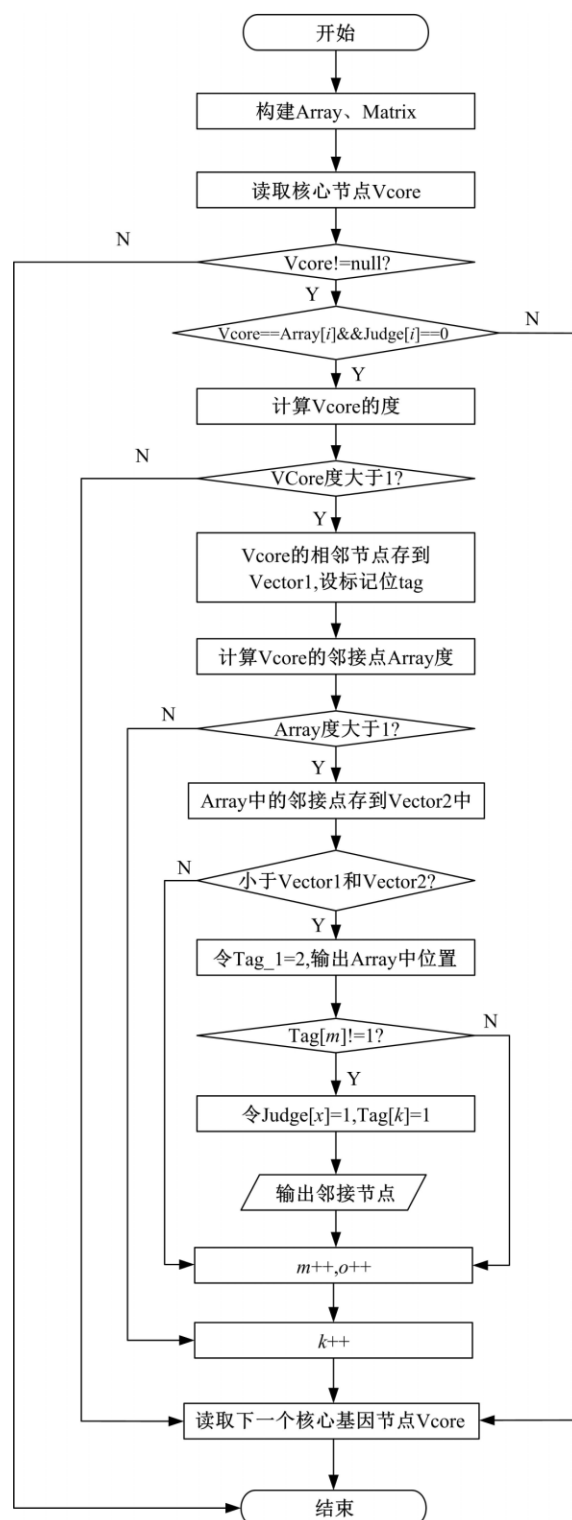
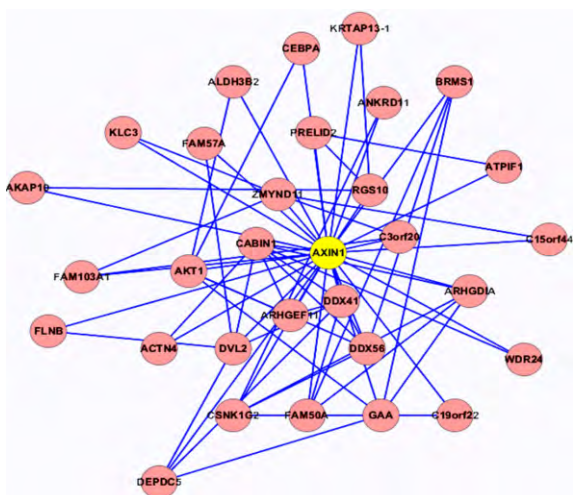


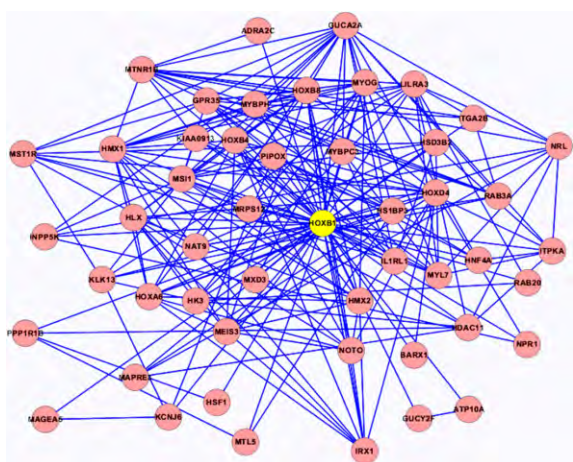
图 3 核心网络模块挖掘方法流程图

Fig. 3 Core network module mining method flow chart

活性 (GO: 0060589)、GTPase 激活活性 (GO: 0005096)、酒精的代谢过程 (GO: 0006066)、脂肪酸代谢过程 (GO: 0006631) 等;HOXB1 涉及胚胎器官发育、铁离子体内平衡、核苷酸的化合物合成过程、



**Fig. 4** AXIN1 core network module



**Fig. 5** HOXB1 core network module

由以上的分析结果可以看出,外源性植物 miRNA 跨物种调节人类靶基因时,其靶点多参与人体的日常代谢过程,这表明饮食习惯的差异,可能会由于其含有的 miRNA 不同,导致摄入人的生命过程受到影响,人体的代谢过程会发生变化,导致生理机能的变化。

Module	Primary Function	Secondary Function	Configuration	Access Point
Core Router	Inter-network routing	Load balancing	Static routing	WAN
Edge Router	Network edge management	Access control	Dynamic routing	LAN
Switch	Local network switching	VLAN management	Static configuration	LAN
Firewall	Network security	Access control	Dynamic configuration	WAN
Load Balancer	Load distribution	Session persistence	Static configuration	LAN
Proxy Server	Client request forwarding	Cache management	Dynamic configuration	LAN
Web Server	Content delivery	Static content hosting	Static configuration	LAN
Database Server	Data storage and retrieval	Query processing	Dynamic configuration	LAN
Application Server	Application logic execution	Session management	Dynamic configuration	LAN
Client	User interaction	Request generation	Static configuration	LAN

功能	具体过程	ID
离子运输	钙离子传输	GO:0006816
	脂多糖运输	GO:0015920
	铁离子跨膜转	GO:0046873
	蛋白质运输	GO:0015031
	囊泡运输	GO:0016192
应激反应和免疫	炎症反应	GO:0006954
	乙醇反应	GO:0045471
	免疫应答调节	GO:0050776
	催化活性调节	GO:0050790
	细胞刺激反应	GO:0051716
	细胞应激反应	GO:0033554
生物代谢	生物胺代谢过程	GO:0006576
	酒精代谢过程	GO:0006066
	脂肪酸代谢过程	GO:0006631
	氮代谢过程	GO:0006807
	葡萄糖代谢过程	GO:0006006
	核苷酸的代谢过程	GO:0009117
	ATP 代谢过程	GO:0046034
	嘌呤核苷酸代谢过程	GO:0006163
	氮化合物的分解代谢过程	GO:0044270
	三磷酸核苷代谢过程	GO:0009141
	大分子代谢过程的正调节	GO:0010604
	磷酸盐代谢过程的调控	GO:0019220
	磷代谢过程的调控	GO:0051174
	正调节胞蛋白质代谢过程	GO:0032270
	正调节蛋白质代谢过程	GO:0051247
	磷代谢过程	GO:0006793
	RNA 代谢过程的调控	GO:0051252
	初级代谢过程	GO:0044238
	细胞代谢过程	GO:0044237
	大分子代谢过程	GO:0043170
	蛋白质代谢过程	GO:0019538
	蛋白质分解代谢过程	GO:0030163
	核苷磷酸代谢	GO:0006753
	DNA 代谢	GO:0006259
	杂环化合物的代谢过程	GO:0046700
	细胞内脂质代谢过程	GO:0044255
核苷酸分解代谢过程	GO:0009166	
嘌呤核苷酸分解代谢过程	GO:0006195	
细胞氮代谢过程	GO:0034641	

### 3 结束语

本文从生物信息学角度,以从高通量数据中挖掘出的外源性植物 miRNA 作为出发点,基于动物 miRNA 的调控机制,构建了植物 miRNA-人体 mRNA 跨界调控模型,研究了外源性植物 miRNA 在人体中可能存在的调控作用。发现筛选出的植物 miRNA 与人类的代谢过程关系密切,并从中挖掘出 AXIN1、HOXB1、ATP5D、

MAPK8IP2 等 9 个核心节点。由此可见,饮食习惯可能会由于人体摄入的 miRNA 的不同导致人体代谢发生变化,从而导致人体生理机能的变化,该结果与欧亚人种的饮食和身体机能的关系是相吻合的,目前还需相关生物实验进行验证。本文提出的跨界调控模型为预测植物 miRNA 与人体生理及病理相关性提供了可能,并可以为动物与植物间共进化提供理论依据,进而从营养学角度揭示日常的饮食习惯对人体生理机能产生的重大影响。同时,外源性植物 miRNA 跨界调控研究将为中草药小分子 RNA 研究找到新的途径,对我国传统中草药制药具有指导意义,将有助于探索治疗多种疾病的新思路和新方法。

#### 参考文献:

- [1] Nelson P, Kiriakidou M, Sharma A. The miRNA world; small is mighty[J]. Trends in Biochemical Sciences, 2003, 28(10): 534-540.
- [2] Bushati N, Cohen S M. miRNA functions[J]. Annual Review of Cell and Development Biology, 2007, 23: 175-205.
- [3] Xiao Chang-chun, Klaus Rajewsky. miRNA control in the immune system; basic inciples[J]. Cell, 2009, 1(9): 26-36.
- [4] Liang G F, Zhu Y L, Sun B, et al. Assessing the survival of exogenous plant miRNA in mice[J]. Food Science & Nutrition, 2014, 2(4): 380-388.
- [5] Pastrello C, Tsay M, McQuaid R, et al. Circulating plant miRNAs can regulate human gene expression in vitro[J]. Scientific Reports, 2016, 6: 32773.
- [6] Zhou Z, Li X, Liu J, et al. Honeysuckle-encoded atypical miRNA2911 directly targets influenza A viruses[J]. Cell Research, 2015, 25(1): 39-49.
- [7] Zhang L, Hou D. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by miRNA[J]. Cell Research, 2012, 22(1): 107-126.
- [8] Di Leva G, Garofalo M, Croce C M. Micro RNAs in Cancer[J]. Annu Rev Pathol, 2014, 9: 287-314.
- [9] Grolmusz V. A note on the pagerank of undirected graphs[J]. Information Processing Letters, 2015, 115(6): 633-634.
- [10] Lü L, Zhang Y C, Yeung C H, et al. Leaders in social networks, the delicious case [J]. PloS One, 2011, 6(6): 1-9.
- [11] Licatalosi D D, Darnell R B. RNA processing and its regulation: global insights into biological networks[J]. Nature Reviews Genetics, 2010, 11(1): 75.