基于虚拟基因敲除的疾病分子网络关键节点辨识方法*

□尤章林 王 耘** 顾 浩 乔延江**

(北京中医药大学中药信息工程研究中心 北京 100102)

摘 要:目的:建立基于虚拟基因敲除的关键节点辨识方法。方法:以疾病相关分子网络为载体,针对网络中节点进行虚拟敲除,通过敲出前后其他节点对生物标记物的节点影响力值变化来评价该节点的关键性,以阿尔茨海默病相关疾病分子网络为例来验证方法的可靠性。结果:阿尔茨海默病已知靶点的虚拟敲除,使其他节点的节点影响力值均有显著下降,而随机节点的敲除无此现象。结论:本方法能够验证已知靶点的关键性并筛选出新的关键节点,具有可行性和可靠性。

关键词:关键节点 阿尔茨海默病 生物标记物 虚拟基因敲除 节点影响力

doi: 10.3969/j.issn.1674-3849.2012.04.011

生物分子网络是描述复杂系统的一种有效工具。目前复杂生物网络在网络简化[1]、网络解析[2]等方面开展了相应的研究,但由于网络规模较大、节点和边代表的分子类型和关系类型较多、动力学过程复杂以及网络动态进化等原因,目前复杂网络研究尚缺乏有效利用生物分子网络信息的分析方法,尤其在疾病层面上,疾病复杂网络中承载的生物分子间相互作用信息在疾病诊断、药物设计等方面的应用还缺乏相应的研究。针对网络中关键节点的研究可有效分析复杂网络生物分子间相互作用信息。目前关键节点研究工作主要针对网络的整体拓扑结构特征和拓扑参数特征[3,4]等进行研究,方法较为单一。

针对疾病而言,复杂疾病在发病机理上可以用复杂网络的形式表达,疾病复杂网络关键节点的针对性干预可以对复杂疾病的治疗起到作用。因此,疾病分子网络中关键节点的辨识对于发现疾病靶点和针对复杂疾病的多靶点药物设计具有重要价值。

针对上述现状,本文提出基于虚拟基因敲除的疾病分子网络关键节点辨识方法,以疾病相关分子网络为载体,通过考察网络中节点虚拟敲除前后对于其他节点的节点影响力值变化,来判定被敲除节

收稿日期:2012-07-09

修回日期:2012-08-13

* 国家自然科学基金项目(30973946):基于三阴三阳系统的药性组合模式及其作用机理研究,负责人:王耘;国家自然科学基金项目(81173568):基于三阴三阳的经方药对语法系统及其作用机制研究,负责人:王耘;科学技术部国家科技支撑计划项目(2008BAI51B01):中药有效成分群辨识技术研究,负责人:张宏桂;北京中医药大学科研创新团队支持项目(2011-CXTD-11):中药信息工程科研创新团队,负责人:乔延江。

** 通讯作者:王耘,教授,博士生导师,主要研究方向:中药方剂的系统生物学研究,Tel: 010-84738620, E-mail:wangyun@bucm.edu.cn;乔延江,本刊编委,教授,博士生导师,北京中医药大学副校长,主要研究方向:中药信息工程,中药新药研发及中药生产过程在线监测,Tel:010-84738621,E-mail: yjqiao@263.net。

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

点的关键性,从而在宏观和微观两个层面上对关键节点进行研究。其中关键节点是指在生物分子网络中存在部分节点因其在位置、连接、功能上等的重要性,在生物系统中处于主调节地位的节点,如关键的调控基因、功能强大的蛋白质、生化反应的关键酶等。基因敲除是指以胚胎干细胞技术和 DNA 同源重组技术为基础和原理,通过对某个基因的完全敲除或部分敲除、基因序列的成段敲除或单一序列的敲除,来观察生物或细胞发生的表型变化,从而达到对特定基因功能研究的目的。虚拟基因敲除是指在生物体外,利用生物分子网络模拟体内的生化反应环境及信号通路传递过程,通过对生物分子网络中特定基因节点或其他分子节点的虚拟敲除,使得该基因节点或其他分子节点在网络中丧失相应的动力学特性和功能,进而分析敲除前后对网络环境或特定蛋白节点的影响,从而模拟体内实体基因敲除前后对生物体所产生影响的方法。

一、方 法

1. 原 理

以构建的疾病分子网络为研究载体,针对网络中疾病的已知治疗靶点进行虚拟敲除,通过考察网络中靶点敲除后对其他节点的节点影响力值变化(影响的范围及影响的幅度)来验证被敲除靶点的关键性,并在后续研究中针对其他的非靶点节点进行虚拟敲除,通过与靶点节点敲除后产生影响的比较,从而发现可能的关键节点作为药物的新治疗靶点。

基于虚拟基因敲除的疾病分子网络关键节点辨识方法相关基本原理如下。

(1) 建立疾病分子网络形式化模型原理。

构建疾病分子网络,需建立网络形式化模型。本文的形式化模型中定义了一个基本生化反应单元,包含 CONTAINS、LEFT、RIGHT、CONTROLLER、ACTIVATION、INHIBITION、PARTICIPANT 7 种关系类型(图 1),一个基本生化反应单元形式化模型构建方法为:主框架为反应物 D 经反应 X 生成产物 E、F,其中反应物 D 是由上游途径中单一分子 A、B、C 经复合反应生成的复合物,此步体现出分子 A、B、C 包含于并形成分子 D,即 CONTAINS 的分子作用关系;对于反应 X,分子 D 为反应的输入分子,即 LEFT 作用关系;分子 E、F 为输出分子,即 RIGHT 作用关系;分子 a 通过控制反应 X 的激活剂 b,从而激活促进反应 X,即 CONTROLLER 和 ACTIVATION 作用关系;同理,分子 c 抑制反应 X,即 INHIBITION 作用关系;最后反应 X 参与了某一细胞信号通路 Y,记为 PARTICIPANT 作用关系。将人体内基本生化反应单元采用生物分子网络来体现,网络中节点代表基本生化反应中的所有参与分子,网络中的边代表分子间相互作用关系,并依据反应作用类型形成有向作用的边,由此形成有向网络形式化模型,如图 1 所示。

(2) 虚拟基因敲除原理。

虚拟基因敲除的基本原则是节点被敲除后,所有与该节点直接相连的边都将被删除,并同时遵循以下原则:若被敲除节点为一个生化反应的输入分子,且该生化反应的输出分子可由其他反应生成,则该生化反应将被删除,但输出分子不被删除;若被敲除节点为一个生化反应的输入分子,且该生化反应的输出分子不能由其他反应生成,则该生化反应及输出分子都将被删除;若被敲除节点为一个生化反应的激活剂或抑制剂,且该生化反应可由其他调节剂激活或抑制,则该反应调节剂的删除不影响反应继续进行,故该生化反应将不被删除;若被敲除节点为一个生化反应的激活剂或抑制剂,且该生化反应不能由其他调节剂激活或抑制,则该生化反应及输出分子都将被删除。

(3) 节点影响力值计算原理。

本文首次提出基于疾病网络生物标记物的节点影响力值的概念。生物标记物[5]是指生物体在损

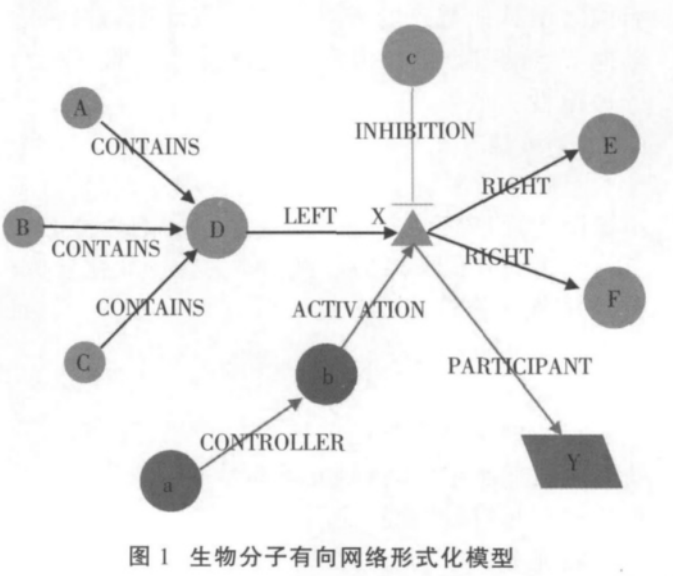

图 1 生物分子有向网络形式化模型

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

害发生的前期，在分子、细胞等水平上产生的异常变化的生物信号指标。因此各种生物分子对生物标记物的影响在一定程度上反映出了该生物分子对疾病的影响。而网络生物标记物[6,7]将疾病的临床生物标记物映射于网络中，为网络的研究提供新思路。节点影响力是指在生物分子网络中的节点通过网络动力学的传播从而对网络生物标记物节点产生影响，即节点对其具有影响力。这里定义节点影响力值(Node Impact Counter, NIC)为 x/L，其中 $x \in \{0,1\}$，即一个节点对网络生物标记物节点无影响即记为 0，有影响则记为 1；L 表示节点到网络生物标记物节点的最短路径长度，$L \in [1,+\infty)$，节点具有影响力，则 L 的最小取值为 1，若节点不具有影响力，认为该节点到网络生物标记物节点间具有无限远的路径，永远不能到达网络生物标记物节点，则 L 取值为 $+\infty$；一个节点的影响力值可根据疾病网络生物标记物的个数 n 分为 n 个部分，也就是说节点影响力值等于该节点对 n 个网络生物标记物节点各自的影响力值的总和，$NIC \in [0,n]$。

2. 步 骤

基于虚拟基因敲除的疾病分子网络关键节点辨识方法步骤如下：

(1) 选择合适的目标疾病。

(2) 构建目标疾病的相关疾病分子网络。

(3) 选择疾病对应的网络生物标记物。

(4) 验证疾病已知治疗靶点的关键性。

依据虚拟基因敲除原理敲除疾病分子网络中疾病的已知治疗靶点；已知治疗靶点虚拟敲除前后网络中其他节点的节点影响力值计算；网络中其他节点的节点影响力值的变化分析；验证靶点的关键性。

(5) 关键节点辨识。

疾病分子网络中非靶点的虚拟敲除；非靶点虚拟敲除前后网络中其他节点的节点影响力值的计算；网络中其他节点的节点影响力值的变化分析；被敲除节点若对其他节点的节点影响力值产生较大影响则可能为关键节点。

3. 适用范围

基于虚拟基因敲除的疾病分子网络关键节点辨识方法适用范围为：研究目标是疾病发生机理和作用机制相对较明确、疾病生物标记物数量较多且研究较充分的疾病；研究的网络对象为须由一步接一步的生化反应所构成的有向疾病网络。

4. 数据来源

本文以阿尔茨海默病 (Alzheimer's Disease, AD) 为例，对以上方法进行举例说明。以题名中包含“阿尔茨海默病”及“信号通路”或“阿尔茨海默病”及“发病机制”的检索方式，在中国知网 (CNKI) 中国学术文献网络出版总库中搜索文献，分别得到 13 篇和 95 篇文献，发表于 1998 年 1 月~2012 年 2 月。选择其中有文献佐证的，的确与 AD 发病机制相关的 18 条信号通路[8~16]。同理以题名中包含“阿尔茨海默病”及“生物标记物”的检索方式，在 CNKI 中搜索得到 6 篇 AD 生物标记物相关文献，发表于 2006 年 5 月~2011 年 6 月。选择其中与 AD 相关且有文献佐证的 4 种生物标记物[17]作为 AD 网络生物标记物。此外，AD 相关信号通路网络的数据来源于 Nature 所属 PID (Pathway Interaction Database) [18]数据库。

二、结果与讨论

因阿尔茨海默病(Alzheimer's Disease, AD)等神经退行性疾病是继心脑血管疾病之后的世界第二大疾病，且 AD 的发病机制较明确，便于后期结果验证，故本文以 AD 为例，对以上方法进行验证。

本文按照上述方法，首先明确并收集了与 AD 相关的 18 条细胞信号通路。其次依据疾病分子网络形式化模型原理，整合 18 条信号通路，构建阿尔茨海默病相关疾病分子网络，包含 2009 个节点，3366 条边。再次根据文献确定 AD 的 4 种网络生物标记物，如表 1 所示。

针对阿尔茨海默病分子网络，基于虚拟基因敲除原理，本文对阿尔茨海默病的主要治疗靶点如 GSK3-beta(3β-糖原合成酶)、APP(Aβ 淀粉样 A4 蛋白)、CDK5(细胞周期蛋白依赖性激酶 5)等进行了虚拟敲除，结果显示这些靶点的敲除对网络中其他节点的节点影响力值产生很大影响，其他节点的节点影响力值均有显著下降。其中 GSK3-beta 的敲除对网络中 2005 个节点中的 153 个节点的节点影响力值产生了影响，且节点影响力值平均下降 60.7%；APP 的虚拟敲除，对 109 个节点的节点影响力值产生了影响，且平均下降 43.2%；CDK5 的虚拟敲除，对 147 个节点的节点影响力值产生了影响，且平均下降 51.3%；此外，本文亦随机选择了处于网络中心的非靶点进行敲除，如 beta-catenin (β 连环蛋

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 1 阿尔茨海默病的相关网络生物标记物

网络生物标记物	与 AD 病理关联性	可以测量的体液	AD 脑脊液中发现
Aβ42	多见于老年斑,FAD 中生成增加	脑脊液,血浆,尿液	降低
t-tau	见于 AD 脑内神经纤维缠结,神经元损伤后分泌	脑脊液	增加
AGEs	促进 AD 相关病理形成,导致神经元死亡	循环组织	未知
Neuron apoptosis	AD 的发生伴随神经元的丢失和死亡	脑组织	无

白)、Ras 蛋白的敲除，分别对 24 个节点和 29 个节点的节点影响力值产生了影响，且节点影响力值下降幅度均极小。以上结果显示,靶点与非靶点的敲除对其他节点的节点影响力值是不同的，且靶点敲除后的影响面和影响程度比非靶点的敲除大得多,以上方法验证了靶点的关键性,有利于对疾病网络的解析。

三、结 论

基于虚拟基因敲除的疾病分子网络关键节点辨识方法能够验证已知治疗靶点的关键性,并具有后续研究价值。该方法的应用将进一步深入体现出生物网络中节点所处的地位和所起的作用,为复杂网络的解析提供思路与方法。但由于工作量巨大,本文尚缺乏针对非靶点的虚拟敲除,可在后续研究中进一步敲除非靶点节点,通过与靶点敲除后产生影响的比较,从而发现可能的关键节点以作为药物的新治疗靶点。只要我们充分利用系统生物学整体性、大规模的优势,充分结合生命科学内的交叉学科和其他科学,可以说网络生物学将拥有更为广泛的应用前景。

参考文献

1 Jeong-Rae Kim, Junil Kim, Yung-Keun Kwon, *et al*. Reduction of Complex Signaling Networks to a Representative Kernel. *Sci Signal*, 2011, 4(175):ra35.

2 Ortega GJ, Sola RG, Pastor J. Complex network analysis of human ECoG data. *Neurosci Lett*, 2008, 447(2-3):129~133.

3 黄海滨,杨路明,王建新.基于网络拓扑的生物网络关键节点识别研究进展.数学的实践与认识,2011,41(7):114~125.

4 黄海滨,王建新,邵平.一种新的拓扑参数及其蛋白质网络关键节点识别.计算机应用研究,2009,26(8):2842~2846.

5 周炯亮. 生物标记物在生化毒理学中的应用. 卫生毒理学杂志,1994,8(4):231~235.

6 Wang XD. Role of clinical bioinformatics in the development of network-based Biomarkers. *Journal of Clinical Bioinformatics*, 2011,1:28.

7 Barry Horwitz, James B Rowe. Functional biomarkers for neurodegenerative disorders based on the network paradigm. *Progress in Neurobiology*, 2011:505~509.

8 李晓红. 高级糖化终产物诱导阿尔茨海默病样病理改变及相关机制研究.华中科技大学博士学位论文, 2011.

9 郭正宇. 阿尔茨海默病与胰岛素信号通路. 中国动脉硬化杂志,2009,17(3):246~248.

10 汝文娟, 唐少君, 钟翎. mTOR 信号通路及其与阿尔茨海默病的关系.中国比较医学杂志,2010,20(8):70~75.

11 洪道俊,朱粹青.阿尔茨海默病相关的早老蛋白与 Notch 和 Wnt 信号通路.生命的化学,2002,22(1):47~49.

12 葛小平,潘三改,李凤,等.转化生长因子 β2 与相关性疾病的研究进展.中医药导报,2007,13(2):85~87.

13 曾凡,王叶冉,周华东,等.P75NTR 在阿尔茨海默病发病机制中的作用及防治新思路. 人民军医-认知障碍及相关疾病专刊,2011:53~56.

14 伍吉云,邓志敏.钙信号与阿尔茨海默病的关系及其作用.井冈山大学学报,2010,31(4):107~110.

15 张宁,倪普,徐凌凡,等.ApoE 基因敲除小鼠海马神经元内 β 淀粉样蛋白的表达变化.安徽医科大学学报,2011,46(6):531~534.

16 付艳玲.氧化应激与阿尔茨海默病以及抗氧化治疗研究进展.国际神经病学神经外科学杂志,2010,37(4):336~340.

17 曾凡,张吉强,王延江.体液生物标记物在阿尔茨海默病早期诊断中的作用与评价. 中国神经精神疾病杂志,2011,37(4):244~246.

18 Carl F. Schaefer, Kira Anthony, Shiva Krupa, *et al*. PID: the Pathway Interaction Database. *Nucleic Acids Research*, 2009, 37:674~679.

Essential Nodes Identification in Disease Molecular Network Based on Virtual Gene Knockout

You Zhanglin, Wang Yun, Gu Hao, Qiao Yanjiang

(Research Center of Traditional Chinese Medicine Information Engineering, Beijing University of Chinese Medicine, Beijing 100102, China)

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Abstract: This paper was aimed to establish essential nodes identification method based on the virtual gene knockout. The disease-related molecular network was used as a subject. This paper virtually knocked-out certain node of the network and thus evaluated the importance of a particular node by introducing the Node Impact Counter (NIC). At last Alzheimer's disease-related network was taken as an example to validate the method. The results showed that the NIC of other nodes decreased significantly after knocking out the known target for Alzheimer's disease, while such phenomenon is not observed when knocking out the non-target nodes. It was concluded that this method can be used to screen out new essential nodes. Examples on Alzheimer's disease-related network show that the method is feasible and reliable.

Keywords: Essential node, Alzheimer's disease, biomarker, virtual gene knockout, node impact counter

（责任编辑：申萌萌　张志华，责任译审：王　晶）

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net