

文章编号:1004-7220(2016)03-0213-05

整合素黏附体信号转导通路预测研究

张晓盼, 宋 婕, 焦 雄

(太原理工大学 力学学院, 应用力学与生物医学工程研究所, 太原 030024)

摘要: **目的** 运用生物信息学方法预测整合素黏附体信号蛋白间的信号转导通路, 为以实验方法研究整合素相关信号转导通路的机制提供参考。 **方法** 把相互作用的整合素黏附体信号分子对搭建成相互作用网络; 利用相互作用的置信概率构建边权; 然后, 采用动态规划算法计算出最小权重的线性通路, 再把这些线性通路组装成通路网络。 **结果** 从 147 个整合素黏附体蛋白所组成的 736 对相互作用中预测出 7 个信号通路网络, 并计算出各个通路网络中基因本体论注释蛋白的占有率。 **结论** 对信号转导通路的研究能够在分子水平上探索疾病的发病机制。预测出可能的信号通路网络, 不仅为基础医学研究疾病机制提供有用信息, 也为探索包括力学、化学等外界信号刺激下的信号转导通路提供有益参考信息。

关键词: 整合素黏附体; 信号转导通路; 动态规划算法

中图分类号: R318.01 文献标志码: A

DOI: 10.3871/j.1004-7220.2016.03.213

Prediction study on signal transduction pathways of integrin adhesome

ZHANG Xiao-pan, SONG Jie, JIAO Xiong (*Institute of Applied Mechanics and Biomedical Engineering, College of Mechanics, Taiyuan University of Technology, Taiyuan 030024, China*)

Abstract: **Objective** To predict signal transduction pathways of the integrin adhesome via bioinformatics method, so as to provide references for experimental study on the mechanism of integrin-related signal transduction pathways. **Methods** First, the interaction network between the interactive integrin adhesome was constructed, and the confidence probability of each interaction was used as its link weight, respectively. Secondly, the pathways of the minimum weight were identified via a standard dynamic programming algorithm. Finally, all pathways calculated by the algorithm were aggregated into some probable networks. **Results** Seven signal transduction pathways were obtained from the integrin adhesome interaction network that contained 147 components with 736 interactions. In every predicted signal transduction pathway, the coverage rate of the proteins with Gene Ontology annotation was calculated. **Conclusions** By research on signal transduction pathways the pathogenesis of some diseases can be explored at the molecular level. Several possible signal transduction pathways obtained in this study have some reference value for exploring disease mechanism in basic medical sciences, and also provide some useful information for such exploration under the stimulation of external signals including mechanical or chemical signals.

Key words: Integrin adhesome; Signal transduction pathway; Dynamic programming algorithm

收稿日期:2016-01-19; 修回日期:2016-02-18

基金项目:国家自然科学基金项目(31070828, 31300770), 博士后科学基金项目(2012T50247, 20100471587), 山西省自然科学基金项目(2009021018-2, 2013021003-2)。

通信作者:焦雄, 副教授, E-mail: jiaoxiong@tyut.edu.cn。

整合素作为一种极其重要的细胞黏附分子,属于跨膜蛋白,对信号转导有着至关重要的作用^[1]。它通过各种复杂的途径来转导外界的力化学信号,包括细胞迁移、生长、分化、凋亡和侵入等在内的大多数细胞表型都由整合素介导调节。因此,预测整合素黏附信号分子之间的信号转导通路对探索生命活动有着举足轻重的作用。然而,整合素黏附体信号分子之间的信号通路是一个极其复杂的调控过程,对于不同的物理、化学或者生物因素的刺激,将引起不同的细胞生理反应或诱导不一样的基因表达,形成不同的信号通路^[2]。近年来,许多学者着手研究在力学信号刺激下整合素信号转导事件对机体的影响,如对血管细胞增殖的影响,对骨组织中成骨分化的调节以及在力学信号刺激下循环血流中整合素 LFA-1 的调控机制^[3]等。其中,整合素介导细胞力化学信号转导主要可激活 FAK 和 She 这两种酪氨酸依赖性通道^[4]。

由于通过实验方法寻找信号转导通路是一个极其繁琐的过程,需要很大的工作量,故借助计算方法预测信号转导通路来为实验研究提供理论导向,可以极大地降低研究成本,缩短研究周期。例如,Steffen等^[5]和 Scott等^[6]采用动态规划的方法从蛋白质相互作用网络中预测信号转导通路,用穷举搜索的策略对加权的蛋白质网络进行通路识别,均取得不错的预测效果;近年来,在从蛋白质网络出发预测信号通路的研究中,这一方法已被广泛应用^[7-8]。本文采用动态规划算法对人类整合素黏附体信号分子间相互作用数据集进行信号转导通路的预测,探索其中信号转导的机制。

1 材料与方法

1.1 构建整合素黏附体相互作用网络

本课题组前期利用 147 个整合素黏附体计算出了 736 对相互作用的整合素黏附体信号分子相互作用对,以及表示每一个蛋白相互作用 (protein-protein interaction, PPI) 对的相互关系可靠度的度量——置信概率 (confidence probability)^[9]。基于每一个整合素黏附体只能作为一个节点在网络中出现和网络中不能包含重复的相互作用以及自相互作用的原则,利用 Cytoscape 软件^[10]把这些 PPI 搭建成整合素黏附体相互作用网络 (见图 1)。

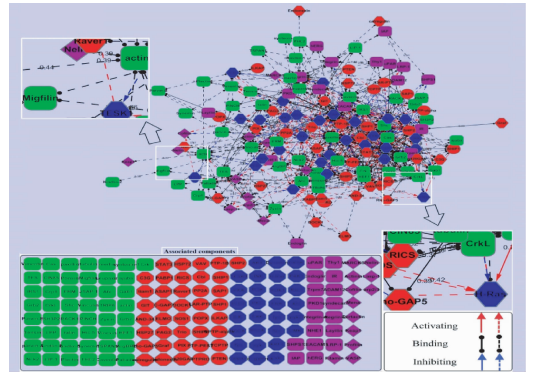


图1 整合素黏附体信号蛋白相互作用网络

Fig. 1 Interaction network of the integrin adhesome

如图 1 所示,把此相互作用网络记为:

$$G = (V, E) \quad (1)$$

式中: V 为黏附体信号分子个数, E 为边数。各边上所加的权值为对应 PPI 的置信概率。红色箭头表示此 PPI 对应的生化相互作用关系为激活 (activation) 作用,蓝色箭头表示抑制 (inhibition) 作用,黑色线条两端各有一个圆点表示 binding 作用;虚线条表示该 PPI 已在 Zaidel-Bar 等^[11]的研究中报道过,实线条表示没有被其报道。另外,节点的颜色与形状是根据整合蛋白的功能而设定的;具体如下:紫色菱形表示肌动蛋白调控 (actin regulation),绿色矩形表示接头蛋白 (adaptor),紫色矩形代表黏附受体 (adhesion receptor) 或通道蛋白 (channel protein),红色六边形代表 RNA/DNA 代谢 (RNA/DNA metabolism) 或 GTP 酶活化蛋白 (GTPase-activating protein, GAP),蓝色六边形表示激酶 (kinase),红色椭圆形表示磷酸酶 (phosphatase)。

1.2 用动态规划算法计算最小权重通路

用 (u, v) 表示整合素黏附体中信号分子 u 和 v 之间的相互作用, $P(u, v)$ 记为对应于边 (u, v) 的置信概率,则用 $w(u, v) = -\log P(u, v)$ 来表示边 (u, v) 的权重。设定集合 I 表示可能的起始整合素黏附体信号分子。对于一个非空集合 $S \subseteq V$ 和每一个 $v \in S$ 的点来说,用 $W(v, S)$ 表示始于 I 中的点、包含 S 中的所有点、终止于 v 的简单通路的最小权重。 $W(v, S)$ 可用递归公式(2)^[6]来计算:

$$W(v, S) = \min_{u \in S - \{v\}} \{W(u, S - \{v\}) + w(u, v), |S| > 1\} \quad (2)$$

式中: $\{v\}$ 表示包含点 v 的集合; $|S|$ 表示非空集合 S 中元素的个数。

在多数的信号通路中,蛋白质按照一定的本质顺序出现,例如:从膜蛋白到核蛋白到转录因子;按照这种顺序依次设定多组连贯的起点和终点,例如:起点 A , 终点 B ; 起点 B , 终点 C ; 起点 C , 终点 D 即为一组连贯的起点和终点。对于每一个 $v \in V$ 的整合素黏附体信号分子,按照式(2)用标准的动态规划算法找到起始于设定的起点、终止于设定的终点的有最小权重的线性通路。然后,按顺序把这些最小权重通路连接成完整的线性通路。

1.3 把线性通路网络化

信号转导通路很少以单一的线性通路存在,通常是由多个线性通路相互连接合并组成复杂的信号通路网络^[12]。一条线性通路上的信号分子能够共价修饰并改变其他线性通路中组分的活性;另外,一条线性通路中的信号传递也可以影响、调节和控制其他线性通路中的信号传递^[13]。本文通过以下规则把计算出的线性通路组装成信号转导通路网络:
① 起点相同且终点也相同的整合素黏附体信号线性通路,才能进行合并;
② 如果各个线性通路之间的蛋白节点存在交互作用,应当予以保留。

2 结果

2.1 通路网络的预测结果

用动态规划的算法从含有 736 对 PPI 的整合素黏附体相互作用网络中计算出 22 条有最小权重的线性通路;然后,把所有起点相同且终点也相同的线性通路合并成信号通路网络。例如:起点为信号分子 PKA (Swiss-Prot ID: P17612)、终点为信号分子 tubulin (Swiss-Prot ID: P68363) 的线性通路有 3 条,可合并成如图 2 所示的通路网络。把剩余的 19 条线性通路按照此规则进行合并后,共得到 6 个通路网络(见图 3)。

2.2 通路网络的评价

在现有的研究中,往往使用基因本体论 (gene ontology, GO)^[14] 中关于生物通路的注释数据来对预测的信号转导通路网络的合理性进行评价。如果一个信号转导通路网络中,由 GO 注释的蛋白质占有率越高,则这个通路网络将有越高的正确性和合理性。利用

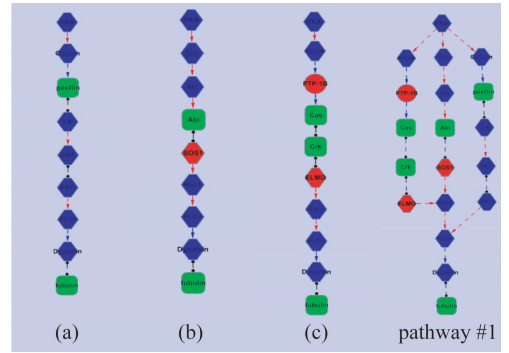


图2 起点为 PKA 且终点为 tubulin 的所有线性通路的合并
Fig.2 Aggregation of all the paths between PKA as starting point and tubulin as ending point

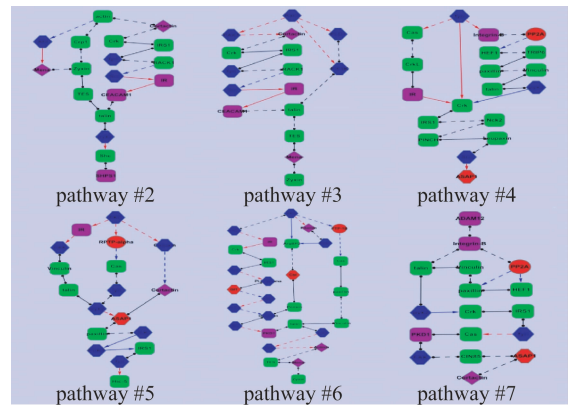


图3 预测出的整合素黏附体信号转导通路网络
Fig.3 Predicted signal transduction pathways of the integrin adhesome

$$P = \frac{N}{M} \times 100\% \quad (3)$$

计算出各个通路网络中获得 GO 注释的蛋白质所占该网络中所出现的全部蛋白质的比例(见图 4)。其中, N 为通路网络中具有 GO 注释的蛋白质数量; M 为同一通路网络中所含的全部蛋白质个数。

3 讨论

研究信号转导通路是了解生物体生命过程的重要途径。近年来,随着研究方法的改进和实验结果的积累,已经发现了部分信号转导通路,其相关调控作用机制也被探索。但是,随着研究的不断深入,仅单纯依靠实验方法来研究信号转导通路已经不能满足后基因组时代各种组学数据大规模爆发的趋势。本文通过用生物信息学的方法预测出可能的整合素

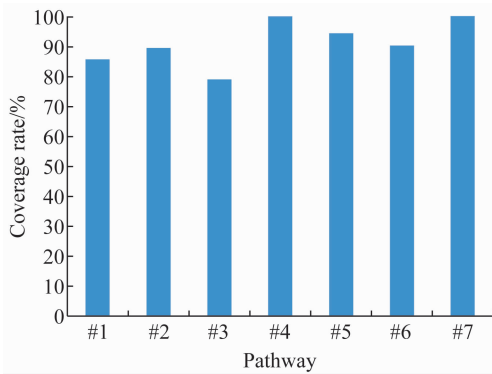


图4 各个通路网络中GO注释蛋白质的占有率

Fig. 4 Coverage rate of proteins which are GO annotation

黏附体相关信号转导通路,为进一步利用实验开展相关研究提供参考。

随着越来越多PPI数据的出现,应用生物信息学的方法预测信号转导通路将会更加普遍。另外,预测方法在不断地改进,数据信息的可靠度也在不断提高,则预测结果的准确率也将会进一步提高。本研究采用目前得到认可的从蛋白质相互作用网络预测信号转导通路的方法,并且使用的整合素信号分子相互作用数据也具有较高的可靠度,故预测结果具有较高的准确度。此外,在现有对蛋白质相互作用网络的研究中,作用网络多以整体网络或功能模块的形式呈现;本文基于蛋白质功能采用数学方法计算出的信号转导通路网络不仅具有功能模块的特点,还能有效指导信号通路的研究。在后续的研究中,将进一步结合实验探索相关预测结果涉及的通路以及其与其他通路之间的关联关系。

目前,通过研究整合素黏附体信号分子介导的信号转导通路探索相关疾病的机制已成为相关研究的热门领域^[15-16]。例如:通过研究整合素激活FAK介导的信号转导通路探索肿瘤的转移^[17];通过研究整合素-肌动蛋白信号转导通路探索骨关节炎的发病机制^[18-19]等。本文预测了整合素黏附体信号分子介导的信号转导通路,对探索相关疾病机制的研究提供帮助。

4 结语

本研究通过动态规划的算法,在以前关于整合素黏附体信号分子间PPI预测结果的基础上,预测了新的整合素黏附体信号分子间可能的信号转导通

路。这将为用实验方法探索整合素黏附体相关信号转导通路提供参考信息,也有助于认识细胞在生命过程中生长、代谢、分化、癌变以及凋亡等方面的表现,进而理解生命的本质,为探索包括力学、化学等在内的外界信号刺激下整合素的信号转导通路机制提供参考。另外,利用生物信息学的方法研究信号转导通路能够在分子的水平上探索疾病的发病机制,为发展新的诊疗手段提供理论指导。

参考文献:

- [1] Luo BH, Carman CV, Springer TA. Structural basis of integrin regulation and signaling [J]. *Annu Rev Immunol*, 2007, 25(25): 619-647.
- [2] Li HP, Huang HY, Lai YR, *et al.* Silencing of miRNA-148a by hypermethylation activates the integrin-mediated signaling pathway in nasopharyngeal carcinoma [J]. *Oncotarget*, 2014, 5(17): 7610-7624.
- [3] 童洁, 刘黎, 吴建华, 等. 力诱导循环血流中白细胞整合素LFA-1的激活机制[J]. *医用生物力学*, 2015, 30(1): 14-20.
- [4] Tong J, Liu L, Wu JH, *et al.* Mechanism of force-induced activation of integrin LFA-1 under circulating flows [J]. *J Med Biomech*, 2015, 30(1): 14-20.
- [5] Seon GJ, Tajik A, Sun J, *et al.* Distinct mechanisms of FAK mechanoactivation by different extracellular matrix proteins [J]. *J Med Biomech*, 2013, 28(S1): 69-71.
- [6] Steffen M, Petti A, Aach J, *et al.* Automated modelling of signal transduction networks [J]. *BMC Bioinform*, 2002, 3(1): 1-11.
- [7] Scott J, Ideker T, Karp R M, *et al.* Efficient algorithms for detecting signaling pathways in protein interaction networks [J]. *J Comput Biol*, 2006, 13(2): 133-144.
- [8] Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis [J]. *BMC Bioinform*, 2009, 9(559): 1-13.
- [9] Ideker T, Krogan NJ. Differential network biology [J]. *Mol Syst Biol*, 2012, 8(1): 565-565.
- [10] Zhang X, Jiao X, Song J, *et al.* Prediction of human protein-protein interaction by a domain-based approach [J]. *J Theor Biol*, 2016, 396: 144-153.
- [11] Shannon P, Markiel A, Ozier O, *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks [J]. *Genome Res*, 2003, 13(11): 2498-2504.
- [12] Zaidel-Bar R, Itzkovitz S, Ma'ayan A, *et al.* Functional atlas of the integrin adhesome [J]. *Nat Cell Biol*, 2007, 9

- (8): 858-868.
- [12] Mustelin T. Biochemistry of signal transduction and regulation [J]. *BBA-Gen Subjects*, 2002, 1571(3): 249-250.
- [13] Mcevoy A, Greenfield D, Shroff H, *et al.* Self-organization of the Escherichia coli chemotaxis network imaged with super-resolution light microscopy[J]. *PLoS Biol*, 2010, 9(3): e1000137.
- [14] The Gene Ontology Consortium. Gene ontology consortium: Going forward [J]. *Nucleic Acids Res*, 2015, 43(D1): D1049-D1056.
- [15] Poonam BP, Sumona S. The role of integrin antagonists in the treatment of inflammatory bowel disease [J]. *Expert Opin Biol Ther*, 2014, 14(12): 1815-1823.
- [16] Winograd-Katz SE, Ssler FR, Geiger B, *et al.* The integrin adhesome: From genes and proteins to human disease [J]. *Nat Rev Mol Cell Biol*, 2014, 15(4): 273-288.
- [17] 周军. 整合素激活 FAK 介导的信号转导通路与肿瘤的侵袭转移[J]. *医学综述*, 2003, 9(6): 321-323.
- [18] 李锦, 罗明秀, 王加丽. 整合素-肌动蛋白信号转导通路与骨关节炎[J]. *国外医学: 医学地理分册*, 2011, 32(3): 215-219.
- [19] 张鹏, 房兵, 江凌勇. 机械刺激对成骨细胞骨架的影响[J]. *医用生物力学*, 2011, 26(1): 87-91.
- Zhang P, Fang B, Jiang LY. Effect of mechanical stimulation on osteoblast cytoskeleton [J]. *J Med Biomech*, 2011, 26(1): 87-91.

· 致读者 ·

关于图表的要求

表和图的设计应科学、简洁、合理,有自照性,均分别按其正文出现先后次序连续编号,并冠以图(表)序号和题目。说明性的资料应置于图(表)下方注释中,并在注释中标明图标中使用的全部非公知公用的缩写及表中的统计学处理。均采用三线表,表内数据同一指标有效位数一致,均数及标准差小数点后保留位数一致。图题、表题及图注、表注均应中、英文对照书写。

黑白图片必须清晰度及对比度良好,层次分明,彩色照片要求色彩鲜明,图像清晰。图片或照片大小要基本一致。图不宜过大,最大宽度半栏图不超过 7.5 cm,通栏图不超过 16.5 cm,高与宽比例以 5 : 7 为宜。图注应放在图题之上及固定脚注之前,照片中需说明的部位请以箭头或字母标注,在图注中说明。图片及照片不得折损。若刊用人像,应征得本人书面同意,或遮盖其能辨认出系何人部分(眼睛)。大体标本照片在图内最好有尺度标记。病理照片要求注明染色方法和放大倍数。

试验数据的曲线图必须根据测试数据绘制,线条必须光滑清晰,有横纵坐标的统计图必须有确切的标目名称和规范的计量单位;实验装置等示意图一定要精心绘制,对图内的结构要有必要的注释,使读者一目了然;引用他人的图(表)一定要用文献角码表明出处。

本刊编辑部