

癌症相关促凋亡蛋白 Par-4 的信号 转导途径预测研究*

赵艳超¹, 张淑波¹, 陈维田², 何 森¹
(1. 中山大学生命科学学院, 广东 广州 510275;
2. 香港理工大学屋宇设备工程系, 香港 九龙)

摘 要: 利用支持向量机(SVM)技术构建 Par-4 关联的蛋白质相互作用网络, 预测出与 Par-4 有相互作用的蛋白质 82 个; 这些蛋白质按照功能划分为 8 大类, 主要包括: 蛋白激酶、泛素化蛋白酶、死亡受体相关因子、与细胞周期或 DNA 复制相关蛋白质、调节蛋白质、与疾病相关蛋白质、具有特定结构域结合蛋白质和其他蛋白质等。结合文献挖掘和数据库检索信息, 推断出了 Par-4 的 2 条可能新的信号转导途径。首次预测到 Par-4 与一大类泛素化蛋白有密切的关系。研究发现, Par-4 与多种蛋白质具有复杂的相互作用, 并且, 在多个细胞凋亡途径中扮演了重要角色。

关键词: Par-4; SVM; 蛋白质相互作用; 细胞凋亡; 信号转导

中图分类号: Q354 **文献标志码:** A **文章编号:** 0529-6579(2010)06-0083-06

Prediction Signaling Transduction Pathways of Cancer-related Apoptosis Protein Par-4

ZHAO Yanchao¹, ZHANG Shubo¹, CHAN Danile Waitin², HE Miao¹

(1. School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China;

2. Department of Building Services Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong)

Abstract: Based on the Support Vector Machine algorithm, the protein-protein interactions (PPI) networks between Par-4, the cancer-related apoptosis protein, and related proteins have been built. There are 82 proteins have been predicted interacting with Par-4. The proteins are classified into eight groups by biological functions, which include protein kinase, death receptor associate factor, ubiquitin-protein ligase, special disease associate proteins, DNA replication related proteins, special structure associate protein, regulatory protein and others. Combined with literatures mining and molecular chaperones of Par-4 searching from BOND database, two new signal transduction pathways of Par-4 have been deduced. In this paper, we first discover that Par-4 interacts with a group of ubiquitin-proteins. We find that Par-4 has complicated interactions with many proteins. Par-4 plays a major role in some apoptosis pathways.

Key words: Par-4; SVM; protein-protein interactions; apoptosis; signal transduction

1994 年 Sells 等^[1]在对前列腺癌研究的过程中, 首先发现了 Par-4 (Pro-apoptosis response-4), 它是一种重要的促凋亡蛋白。Par-4 蛋白质广泛存在于脊椎动物细胞, 尤其是哺乳类动物细胞中,

与鼠的 NLS1, NLS2 区域 100% 的同源。近期研究发现, Par-4 与前列腺癌、肝癌、子宫癌, 以及神经变性疾病等有密切关系^[2]; 相关研究包括 Par-4 基因的缺失, 在细胞中的定位差异, 以及与其

* 收稿日期: 2009-10-09

基金项目: 国家自然科学基金重点资助项目(60736028); 广东省科技计划项目重大专项资助项目(2003A3080503)

作者简介: 赵艳超(1986年生), 男, 硕士生; 通讯作者: 何森; E-mail: lsshem@mail.sysu.edu.cn

它凋亡蛋白的相互作用等。研究还发现 Par-4 有多个伴侣分子,在细胞凋亡过程中共同起作用^[2]。最近文献报道,该蛋白质与异构拓扑酶 Top1 具有相互作用,通过影响 DNA 的复制,来阻止细胞的正常功能^[3]。美国肯塔基大学研究小组 2007 年利用转基因的方法将 Par-4 基因转入老鼠体内,培育出了对多种癌症免疫的老鼠^[4]。由于 Par-4 可以选择性地引发癌细胞的凋亡,对于正常或永久性细胞却没有影响,因此,引起了研究人员的极大兴趣。

Par-4 在细胞中的功能十分复杂。常规实验条件下,对 Par-4 与其它蛋白质间相互作用的研究可以借助于酵母双杂交、免疫共沉淀等方法。但是,实验方法只能涉及成对蛋白质间的相互作用、或单一功能和途径的研究。目前,很多实验性工作主要针对单一类型癌细胞的研究以探索 Par-4 的促凋亡机理,往往会得出相互矛盾的结果;主要原因是不同细胞类型可能具有不同的凋亡途径,这也说明了 Par-4 促凋亡途径的复杂性和多样性。

本文基于支持向量机(SVM)技术和相关数据库数据,系统性预测 Par-4 蛋白质相互作用(Protein-protein Interaction, PPI)网络;同时,借助于文献信息挖掘技术和数据库实验信息,预测 Par-4 功能及其可能的凋亡信号转导途径。

1 数据与计算方法

1.1 数据来源

Par-4 和相关蛋白质序列获取自 NCBI 的相关数据库。从 HPRD (Human Protein Reference Database) 中下载了人类的全部蛋白质参考数据。蛋白质相互作用数据参照 Guo 等人的方法搜集和整理^[5],构建本地数据集。

使用 Perl 语言编程,从本地数据集筛选目标蛋白质。选取的目标蛋白质要求具有以下特征:①该蛋白质在细胞中定位于细胞质或细胞核;②该蛋白质的细胞功能为细胞凋亡。

SVM 算法所用的正集来自 DIP (Database of Interaction Protein) 的酿酒酵母子数据库,数据库版本是 20070219,总共获得 5 944 对蛋白质相互作用数据;剔除其中同源性 > 40% 的数据对,剩余 5 594 对蛋白质相互作用数据。负集是按照细胞内定位的方法进行收集;在不同细胞内位置,依据蛋白质在生物学意义上没有相互作用关系,选取了 5 594 对非蛋白质相互作用数据。

从 DIP 的人类子数据库中下载所有人类的蛋白

质相互作用对,作为 SVM 测试数据集,版本是 20090319。

1.2 SVM 算法与检验

参照 Shen 等的建模方法^[6],将蛋白质序列的氨基酸按照不同的极性、电偶性、疏水性等特征分成 7 类,计算氨基酸三联子(即连续的 3 个氨基酸)在蛋白质序列中的频率,可以得到 343 (7 × 7 × 7) 维特征向量。在此基础上,计算蛋白质对的 686 (即 343 + 343) 维特征向量;进一步,处理为 SVM 算法的输入格式。我们选用的 SVM 类型为 C-SVC,核函数类型为 RBF 函数 ($\exp(-r|u-v|^2)$)。利用 Lib-SVM-2.69 软件中的 Grid. by 程序寻找最优参数。具体做法参照 <http://www.csie.ntu.edu.tw/~cjlin>。

本文采用 3-Fold 检验方法,用于确定模型的准确率和稳定性。检验计算涉及 4 个参数和 3 个评判指标。相关参数包括:TP (真阳性,即相互作用的蛋白质对预测为相互作用),FP (假阳性,即非相互作用的蛋白质对预测为相互作用),TN (真阴性,即非相互作用的蛋白质对预测为非相互作用),FN (假阴性,即相互作用的蛋白质对预测为非相互作用);3 个评判指标为:Sn (敏感性)、Sp (特异性)和 Acc (准确率)。参数计算公式如下:

$$Sn = \frac{TP}{TP + FN}$$

$$Sp = \frac{TN}{TN + FP}$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \times 100\%$$

1.3 PPI 预测结果的验证和信号转导途径推测

依据 SVM 计算结果,构建以 Par-4 为核心的蛋白相互作用网络。利用 Cytoscape 软件进行可视化处理和 PPI 网络输出。

从 BOND (Biomolecular Object Network Database) 检索与 Par-4 有相互作用关系的蛋白质,该数据库中记录的相互作用蛋白质对是通过酵母双杂交技术或免疫共沉淀等实验方法得到,具有很高的可信度;另一方面,结合文献挖掘技术,检索和阅读了数百篇相关文献或摘要;交叉验证 PPI 预测结果。进一步,综合利用以上信息,推测凋亡蛋白 Par-4 的信号转导途径。

2 结果与分析

2.1 SVM 模型及预测结果

通过参数优化,得到 C-SVC 的主要参数 $C =$

32, $g = 0.03125$, 校正准确率为 92.11%。进一步, 利用 3-fold 计算模型的敏感性、特异性和准确率等指标, 对测试数据集的预测结果见表 1。检验结果表明该模型的可信度高。

表 1 对测试数据集的预测结果

Table 1 The prediction results of the test sets

测试集 Test sets	TP	FN	TN	FP	Sn/%	Sp/%	Acc/%
1	1 703	162	1 759	106	91.31	94.14	92.82
2	1 717	148	1 759	106	92.06	94.19	93.19
3	1 689	175	1 745	119	90.61	93.42	92.11
平均值	1 703	162	1 754	110	91.31	93.93	92.71

文中还利用该模型测试从 DIP 人类子数据库下载的 2 171 对 PPI 数据, 得到的准确率为 80.7%, 并且, 对独立的非相互作用蛋白质数据预测结果也达到 80% 的准确率。

依据目标蛋白质特征要求, 从 HPRD 中筛选相关蛋白质, 得到 152 个符合条件的蛋白质。基于 SVM 模型, 预测出其中有 82 个蛋白质与 Par-4 有相互作用。依据生物学功能注释信息, 对这些蛋白质进行了有效分类, 分类结果见表 2。

根据生物学功能, 可以将这些蛋白质划分为 8 类, 主要类别包括: 蛋白激酶、泛素化蛋白酶、死亡受体相关因子、与细胞周期或 DNA 复制相关蛋白质、调节蛋白质、与疾病相关蛋白质、具有特定结构域结合蛋白质和其他蛋白质。

表 2 与 Par-4 可能有相互作用的蛋白质功能分类

Table 2 Function classification of proteins which interact to Par-4

类别 Classification	蛋白质 Proteins	功能 Function
蛋白激酶类 (15)	S/TPK17A, S/TPK17B, CSRN3, RIPK1, RIPK3, PTPRH, DAPK2, S/TPK3, S/TPK2, DAPK3, HIPK3, Caspase-1, Caspase-3, Caspase-8, CASP8AP2	激活蛋白质, 促使蛋白质磷酸化, 自身磷酸化引起细胞凋亡。
泛素化蛋白酶 (11)	RNF34, PACAP, RNF216, TRAF7, RIFIYLI, RYBP, XAF1, IAPAF1, BIRC2, BIRC8, PEA-15	促使蛋白质的降解。
死亡受体相关蛋白质 (9)	TAADP, TRAF2, TRAF3, TRAF4, RABP1, RASSF5, RBP, FAF1, TP53INP1	接受死亡信号, 传递死亡信号的相关蛋白质。
与细胞周期或 DNA 复制相关蛋白质 (5)	NBPC1D, DFF α , DFF β , CdcARP1, AENase	影响 DNA 的复制, 从而影响细胞周期。
调节蛋白质 (11)	NAIF1, PDCD4, PDCD7, Siva, Apaf1, ASPP2, TP53INP3, ASC, Bel2LP, PIAP1, NP3	在细胞中起调节作用。
与疾病相关蛋白质 (4)	HIP1, A1P1, β -Clp1, BNIP2	引起特定疾病的相关蛋白质。
具有特定结构域结合蛋白质 (12)	ZFP4, SAP30BP, DDAP6, BID, BIF1, DEDD2, PHLPP, NALP3, TRIM35, TRIM39, NOBDP1, BMCC1	与特定序列结构作用, 这些结构与凋亡有关。
其他蛋白质 (14)	Cal-12, Gal-7, Keratin, BC13, NOIG, RTDPEG10, Peg3P, PHLDA1, PHLDA3, AKTIP, FEM1B, Killin, EAF2, Jade-1	具有多种功能的蛋白质。

基于 SVM 模型预测结果, 成功构建了以 Par-4 为核心的 PPI 网络 (见图 1: A), 并按照其不同的功能把相关 PPI 蛋白质有效进行了聚类。参照以上几类蛋白质在细胞中的功能, 依托相关文献信息挖掘结果, 初步推测出凋亡蛋白 Par-4 在细胞中可能的信号转导通路 (见图 1-B)。我们推测细胞首先通过细胞膜的受体接受胞外信号, 信号激活与受体相联系的受体相关蛋白, 受体相关蛋白与 Par-4 相互作用, 把胞外信号传递给 Par-4; 并且, 同时蛋白激酶激活 Par-4, 使其磷酸化, 进入细

胞核内; 在细胞核内其可以与相关的转录因子结合, 从而影响基因的转录和翻译, 达到调节基因活性的作用; 当细胞没有接到胞外信号时, Par-4 被泛素化蛋白降解, 或者和调节蛋白相结合而失活。根据计算模型, 发现 Par-4 可能和某些癌症标记物有相互作用, 可能还与具有特定结构域的蛋白相互作用, 引起细胞的凋亡。本文首次预测到 Par-4 与一大类的泛素化蛋白有密切的关系。泛素化蛋白在细胞中的作用主要是促进蛋白的降解。

间的相互关系和信号转导途径。

3 讨论

Par-4 作为一种促凋亡蛋白, 多年研究发现其与癌症和神经退化性疾病具有密切关联; 并且, 在不同的细胞中发现其存在相互矛盾的作用; 由此判断, Par-4 在细胞凋亡过程中可能具备多种功能并参与了多种凋亡通路^[9]。

近期研究表明, Par-4 在行使其功能的过程中, 与其在细胞中的定位, 以及是否和其它蛋白形成复合物具有密切的关系。在细胞质中, Par-4 被 PKA 磷酸化时, 可以促进 Par-4 进入细胞核; 而 Par-4 一旦被 Akt 磷酸化, 则被滞留在细胞质内^[7]。在细胞核中, Par-4 可以和具有磷酸化的蛋白质如 Dlk、Amdia 形成复合物而出核^[10], 并可以和细胞骨架结合促进肌球蛋白的磷酸化, 从而导致细胞骨架裂解。Par-4 可以与 E2F1 形成的复合物促进 smac 基因的转录^[11], 还可以与 WTI 形成复合物抑制 Bcl-2 的转录^[2]。综上所述, 正是由于 Par-4 在疾病发生过程中, 与多种蛋白质相互作用, 可能参与了多条信号通路。因此, 只有从系统的角度开展研究, 才能够对全面地整合 Par-4 的作用机制, 准确理解、掌握其功能和在疾病产生过程所扮演的角色。

本文依据 SVM 计算结果、结合文献挖掘和数据库检索信息, 发现了 Par-4 的 2 条可能的新的信号转导途径 (图 3)。信号转导途径中相关蛋白质之间的相互作用可能导致细胞凋亡, 或维持细胞生存的某种平衡状态, 一旦这种平衡状态被破坏, 就会引发相应的疾病。例如, Par-4 过量表达会抑制 ζ PKC, 但 ζ PKC 可以促进 Akt 的活性, Akt 使 Par-4 与 14-3-3 结合滞留在细胞质, 使其无法进入细胞核起作用。Par-4 的缺失可能导致 ζ PKC 和 PKA 的过量表达, 其中, ζ PKC 可以激活 NF-kB, 引起细胞的过量增殖或癌症。本文关键的发现是 Par-4 可能通过 ζ PKC 和 NF-kB 两种途径来阻止 NF-kB 起作用。

此外, 我们还发现 Par-4 与一些疾病相关蛋白质具有重要相互作用, 例如, 与亨廷顿蛋白结合蛋白 1 (HIP1)、淀粉样蛋白 (AIP1)、 β -连接样蛋白 (β -Clp1) 和 BCL2/腺病毒 E1B 19kDa 结合蛋白 (BNIP2) 等具有相互作用。这几个蛋白质分别与亨廷顿病、老年痴呆症、结肠癌、心肌炎等相关疾病具有密切的关系。实际上, 现有研究文献已经证明这些疾病与 Par-4 具有密切的关系^[2]。

蛋白质间如果发生相互作用, 可能会处于相同的调控途径, 但不一定有物理上的接触。具有相同功能的蛋白质可能有效类聚在一起。为减少预测的假阳性, 将预测有相互作用关系的蛋白质按照功能分成 8 大类。我们构建了 Par-4 为核心的蛋白质相互作用网络。例如, 丝氨酸/苏氨酸蛋白激酶 (S/TPK17A, S/TPK17B), 在 Par-4 的结构中拥有丝氨酸和苏氨酸位点, 文献证明 Par-4 可以被 PKA、Akt 等蛋白磷酸化^[8-9], 与预测结果吻合。预测 Caspases 凋亡途径与 Par-4 有关, 在文献中得到了印证^[9]。并且, 预测到 Par-4 与 DNA 复制相关蛋白质具有相互作用, 最新的研究表明, Par-4 可以与 DNA 复制有关的 Top1 形成复合物, 抑制 DNA 的复制, 从而使细胞周期停留在 DNA 复制阶段^[3]。我们还预测到一组与细胞周期和 DNA 复制有关的蛋白质, 包括 DNA 裂解因子 (DFF α) 等; 初步推测 Par-4 与 DNA 复制过程的相互作用, 可能也是一种促细胞凋亡的途径。同时, 预测发现 Par-4 与一些死亡受体蛋白可以相互结合, 与相关的多种死亡途径建立了联系, 例如, Rab、Ras、Fas、TNF 等; 该结果同样获得了实验的支持^[2, 9]。预测指出 Par-4 与一些特定的结构域有关, 例如, 锌指、亮氨酸富集区、IAP、BH3 等结构。在 Par-4 结构分析时, 发现 Par-4 具有 C 端亮氨酸拉链结构, 其特殊的结构可以和许多其他特殊结构相结合, 伴侣分子 WT1 就具有锌指结构; 而且, Par-4 还可以通过亮氨酸链与自身发生相互作用, 这可能也是一种平衡 Par-4 在细胞中起作用的主要机制。此外, 本文首次预测到 Par-4 与一大类的泛素化蛋白有密切的关系, 泛素化蛋白在细胞中的作用主要促进蛋白的降解, 这对于稳定 Par-4 在细胞中的功能可能具有重要意义。

本研究对于完整地理解 Par-4 在细胞中的功能具有重要意义和价值; 对于开展进一步的实验研究具有指导意义。由于 Par-4 可以选择性诱导细胞凋亡, 在这一点上, 其作为药物靶点具有很大的吸引力。

参考文献:

- [1] SELLS S F, WOOD D P, RANGNEKAR V M, et al. Commonality of the gene programs induced by effectors of apoptosis in androgen-dependent and-independent prostate cells [J]. Cell Growth & Differentiation, 1994, 5:457-466.
- [2] NADIA E G, VIVERK M R. Apoptosis by Par-4 in

- cancer and neurodegenerative diseases [J]. *Experimental Cell Research*, 2003, 283:51-66.
- [3] GOSWAMI A, QIU S, THOMAS S D, et al. Par-4 binds to topoisomerase I and attenuates its DNA relaxation activity [J]. *Cancer Research*, 2008, 68(15):6190-6198.
- [4] ZHAO Yanming, BURIKHANOV R, RANGNEKAR V M, et al. Cancer resistance in transgenic mice expressing the SAC module of Par-4 [J]. *Cancer Research*, 2007, 67(19):9276-9285.
- [5] GUO Yanzhi, YU Lezheng, WEN Zhining, et al. Using support vector machine combined with auto covariance to predict protein-protein interactions from protein sequences [J]. *Nucleic Acids Research*, 2008, 36(9):3025-3030.
- [6] SHEN Juwen, ZHANG Jian, LUO Xiaomin, et al. Prediction protein-protein interactions based only on sequences information [J]. *PNAS*, 2007, 104(11):4337-4341.
- [7] GURUMURTHY S, GOSWAMI A, VASUDEVAN K M, et al. Phosphorylation of Par-4 by protein kinase A is critical for apoptosis [J]. *Molecular and Cellular Biology*, 2005, 25(3):1146-1161.
- [8] GOSWAMI A, RANGANATHAN P, RANGNEKAR V M. The Phosphoinositide 3-Kinase/Akt1/Par-4 axis: A cancer-selective therapeutic target [J]. *Cancer Research*, 2006, 66:2889-2892.
- [9] 秦天洁, 马巍, 刘陕西, 等. Par-4 诱导细胞凋亡发生机制的进展 [J]. *现代肿瘤医学*, 2008, 16(12):2216-2219.
- [10] PARK S K, NGUYEN M D, FISCHER A, et al. Par-4 links dopamine signaling and depression [J]. *Cell*, 2005, 122:275-287.
- [11] LU Chao, CHEN Jieqing, ZHOU Guoping, et al. Multimolecular complex of Par-4 and E2F1 binding to Smac promoter contributes to glutamate-induced apoptosis in human-bone mesenchymal stem cells [J]. *Nucleic Acids Research*, 2008, 36(15):5021-5032.

(上接第82页)

- [18] SMART G R. Investigations of the toxic mechanisms of ammonia to fish - gas exchange in rainbow trout (*Salmo gairdneri*) exposed to acutely lethal concentrations [J]. *J Fish Biol*, 1978, 12(1):93-104.
- [19] WHITFIELD M. The hydrolysis of ammonium ions in sea water—a theoretical study [J]. *J Mar Biol Assoc U K*, 1974, 54(3):565-580.
- [20] ARMSTRONG D A, CHIPPENDALE D, KNIGHT A W, et al. Interaction of ionized and un-ionized ammonia of short-term survival and growth of prawn larvae *Macrobrachium rosenbergii* [J]. *Biol Bull*, 1978, 154:15-31.
- [21] SHARMA B, AHLERT R C. Nitrification and nitrogen removal [J]. *Water Res*, 1977, 11:879-925.
- [22] CHEN J C, CHIN T S. Acute toxicity of nitrite to tiger prawn *Penaeus monodon* larvae [J]. *Aquaculture*, 1988, 69(3/4):253-262.
- [23] 丁美丽, 林林, 李光友, 等. 有机污染对中国对虾体内外环境影响的研究 [J]. *海洋与湖沼*, 1997, 28(1):7-12.
- [24] 吴中华, 刘昌彬, 刘存仁, 等. 中国对虾慢性亚硝酸盐和氨中毒的组织病理学研究 [J]. *华中师范大学学报: 自然科学版*, 1999, 33(1):119-122.
- [25] BRAY W A, LAWRENCE A L, LEUNG D K, et al. The effect of salinity on growth and survival of with observations on the interaction of IHNV virus and salinity [J]. *Aquaculture*, 1994, 122:133-146.