

一种新的蛋白酪氨酸磷酸酶基因的克隆、定位和组织表达谱研究

瞿祥虎, 翟云, 魏汉东, 鱼咏涛, 贺福初

(军事医学科学院放射医学研究所基因组学与蛋白质组学研究室, 北京 100850)

摘要: 从人胎肝 cDNA 文库分离出一长度为 5248bp 的 cDNA 克隆, 该基因包含 26 个外显子和 25 个内含子, 染色体定位于在某些肿瘤细胞中易缺失的 3p21.1-21.33。其可读框编码 1636 个氨基酸, 该蛋白属于蛋白酪氨酸磷酸酶(PTP)家族, 其 C 端有一个典型的 PTP 结构域, N 端含有约 800 氨基酸残基的 BRO1 样结构域及随后 2 个可能的 SH3 结构域结合位点, 在这两个结构域之间及 C 末端还各有一个脯氨酸富集区。Northern 杂交和点杂交分析显示, 该基因以大约 5.4kb 的单一转录物广泛表达于人体各种组织, 而且在人部分肿瘤细胞中高表达。结果提示, 人源 PTP-TD14 是一个新的蛋白酪氨酸磷酸酶。

关键词: 蛋白酪氨酸磷酸酶; cDNA 克隆; 人染色体 3p21; PTP 结构域; BRO1 样结构域; 组织表达谱

中图分类号: Q343.22

文献标识码: A

文章编号: 0253-9772(2001)06-0503-08

Cloning, Chromosomal Assignment and Tissue Expression of a Novel Human Protein Tyrosine Phosphatase (PTP-TD14) Gene

QU Xing-hu, ZHAI Yun, WEI Han-dong, YU Yong-tao, HE Fu-chu

(Department of Genomics and Proteomics, Beijing Institute of Radiation Medicine,

Academy of Military Medicine Science, Beijing, 100850 China)

Abstract: A human cDNA of 5248bp encoding a novel protein tyrosine phosphatase PTP-TD14(1636aa) has been isolated from fetal liver. The gene is located at chromosome 3p21.3, an area frequently deleted in many types of cancer, and composed of at least 26 exons and 25 introns. The phosphatase has unique features in its domain structure: a tyrosine phosphatase domain, a C-terminal PEST motif, two SH3-binding motifs, two proline-rich region and an N-terminal domain similar to yeast BRO1(a yeast protein that is involved in the mitogen-activated protein kinase signaling pathway). Northern blot and dot blot hybridizations indicate that it is expressed ubiquitously in human 50 tissues and 7 cancer cell lines. Thus, it is a novel protein tyrosine phosphatase gene located on 3p21.3.

Key words: protein tyrosine phosphatase; cDNA cloning; human chromosome 3p21; PTP domain; BRO1 like domain; tissue transcription profile

在胎儿发育过程中, 胎肝不仅行使肝脏的功能, 而且兼有重要的造血功能, 尤以红系造血为主, 因而

胎肝成为造血研究的一个重要模型。20 世纪 90 年代以来, 本研究室以肝脏造血为突破点, 试图对参与

收稿日期: 2001-05-08 **修回日期:** 2001-07-18

基金项目: 国家高技术发展计划项目(863-102 08-05-2)、国家自然科学基金(39730310, 30070177)和国家重点基础研究发展规划资助项目(G1999053903)。

作者简介: 瞿祥虎(1963-), 男, 湖北洪湖人, 副教授, 博士后, 专业: 分子生物学。

通讯联系人: 贺福初(1962-), 男, 湖南常德人, 研究员, 博士生导师, 专业: 分子遗传学, Tel: 010-66931216; E-mail: hfc@nic.bmm.ac.cn

造血细胞增殖分化过程的相关基因进行规范化研究^[1]。通过对 22 周孕龄人胎肝 cDNA 文库进行大规模测序,获得了 1 万 8 千余条 EST 序列,约 51% 为新基因序列;其中 1 条 EST(FLF2896)引起我们的极大兴趣,这是因为其可能编码 293 个氨基酸(aa)的部分肽段,且该肽段的氨基酸序列与蛋白酪氨酸磷酸酶(PTP)的典型结构域 PTPc 高度同源。

PTP 是一个有众多成员的大家族。估计人类基因组中有多达 500 个此类蛋白的基因^[2]。迄今在不同物种已鉴定近 80 种 PTP,它们分属于胞质 PTP 和跨膜 PTP 两类^[3]。PTP 的生化功能是催化蛋白质上酪氨酸残基的去磷酸化,因而在细胞生长、分化和转化等多种过程的信号转导通路中起极为重要的作用。研究表明,许多 PTP 在造血系统中高表达,有些甚至在造血细胞中特异高表达^[3~5]。正由于 PTP 酶在造血过程中的重要作用,本文利用网上 EST 库将克隆 FLF2896 的插入片段进行有效电子延伸及 PCR 扩增后,成功地分离到一条全长为 5248 bp 的 cDNA,已在 GenBank 登录(AF290614)。它的可读框(ORF)编码 1636 个 aa;由于其与小鼠的一种蛋白酪氨酸磷酸酶 PTP-TD14 高度同源^[6],因而被命名为人源 PTP-TD14。Northern 杂交和点杂交研究表明,该基因在多种组织中广泛表达,且在某些人肿瘤细胞中呈高表达。结果提示 PTP-TD14 蛋白在介导细胞增殖、凋亡或转化的信号转导途径中可能起重要作用。

1 材料与方 法

1.1 人 PTP-TD14 的克隆

(1)电子拼接和 PCR 扩增 对 22 周孕龄人胎肝 cDNA 文库随机测序后所获的 EST,经上网进行同源性及功能线索搜索,挑选出一个可能编码含 PTPc 肽段的克隆 FLF2896,利用本地库及 GenBank 的 EST 数据库进行重叠群分析,然后根据拼接延伸后的 cDNA 序列在两端设计并合成一对特异 PCR 引物: P1 (5' - ACGCGTCCGGCAGCTGCAG-CAGCTACGG - 3'), P2 (5' - ACGGGCCA-CAGAACAGGGTGGTTTATTA - 3')。以此从人胎肝 cDNA 扩增出预期长度的 PCR 产物。PCR 反应按以下条件在 GeneAmp PCR System 9700 (PE Applied Biosystems) 上进行: 94°C 变性 5 min 后, 94°C, 45 s, 58°C, 45 s, 72°C, 6 min, 32 个循环, 72°C

再延伸 7 min。所有片段重组至 PROMEGA 公司的 pGEMT-T 载体以备测序。(2)RT-PCR 制备人胎肝 cDNA 模板 人胎肝 mRNA (CLONTECH) 用 MMLV 反转录酶 (Superscript II, GIBCO 公司) 和 SMART™ RACE cDNA Amplification Kit (CLONTECH) 按试剂盒方法进行反转录合成第一条链。(3)DNA 测序 用 QIAGEN 公司的小量质粒提取试剂盒提取质粒 DNA,定量后用 ABI PRISM 系统 (Perkin-Elmer) 在 ABI 377-XL DNA 自动测序仪 (Perkin-Elmer) 上从双向进行测序反应。

1.2 人 PTP-TD14 的基因组结构和染色体定位分析

(1)基因组结构分析 以人 PTP-TD14 的全长 cDNA 序列于 GenBank 的 HTG 数据库搜索可能与其对应的基因组序列,将 cDNA 序列与其基因组序列进行二重比对后找出外显子和内含子,比对结果用典型的剪接模式 (GT/AG) 校正。(2)基因的染色体定位 生物信息学方法研究新基因的染色体定位的原理是以一条或多条 EST 作为靶基因与 UniGene 簇之间的桥梁。如果待定位基因所含的某 EST 属于一个已知的 UniGene 簇,而这个 UniGene 簇已通过某染色体位标 (sequence-tagged site, STS) 定位于一条染色体的某个区段,那么该基因也定位于同样的染色体区段。

1.3 人 PTP-TD14 蛋白序列分析

人 PTP-TD14 与已知蛋白的同源性在 NCBI (National Center for Biotechnology Information) 进行 BLAST 搜索,多条序列之间的比对采用 Mac Vector (OXOL MOL COM) 软件。蛋白质序列中可能的结构域、基序 (motif) 和跨膜区等的查询分别在以下网站进行^[7,8]: <http://expasy.hcugc.ch/sprot/prosite.html>, <http://coot.embl-heidelberg.de/SMART/>, http://www.ch.embnet.org/software/TMPRED_form.html。

1.4 Northern 杂交和点杂交分析

分别载有人 12 种正常组织 (脑、心、骨骼肌、结肠、胸腺、脾、肾、肝、小肠、胎盘、肺和白细胞,每个泳道含约 1 μg mRNA) 和 7 种肿瘤细胞 (早幼粒细胞 HL-60、HeLa 细胞 S3、慢性髓性白血病细胞 K-562、成淋巴细胞 MOLT-4、结肠腺癌细胞 SW180、肺癌细胞 A549 和黑素瘤细胞 G-361,每个泳道含约 2 μg mRNA) 的 Northern 杂交膜,按 CLON-

TECH 公司提供的方法进行杂交。用(α - 32 P)dCTP 和 Prime-a-Gene[®] 标记系统 (PROMEGA) 中的 DNA 聚合酶 I Klenow 片段以随机引物法标记探针。标记探针的 DNA 模板分别为 5 端的 747bp 核苷酸(第 1~747 核苷酸)和近 3 端的 881bp 核苷酸(即克隆 FLF2896 的插入片段)。载有来自人体 50 种不同组织 mRNA 的点杂交膜 (RNA Master Blot[™], CLONTECH) 也用 5 端的 747bp 探针按该公司推荐的方法进行杂交。杂交膜经杂交、洗膜、晾干后与 X 射线底片(加增感屏)于 -70℃ 压片 72h 后常规显影。将曝光底片上的阳性信号用 GS-710 Calibrated Imaging Densitometer (BIO-RAD) 扫描, 以 PD-Quest (BIO-RAD) 软件分析其 OD 值, 据此分析目的基因在各种组织中的相对表达量。

2 结 果

2.1 人 PTP-TD14 全长 cDNA 的克隆

对 22 周孕龄人胎肝 cDNA 文库随机测序后, 经 BLAST 搜索, 挑出插入片段为 881bp 的一个感兴趣克隆 FLF2896, ORF 分析显示其可能编码 293 个 aa 的肽段, 其 aa 序列与 PTP 特有的结构域 PTPc 高度同源, 但该 EST 既缺乏起始密码子又无终止密码子。因此, 将其进行网上电子拼接后, 获得一条长 5248 bp 的 cDNA 序列; 参与拼接并使其得到有效延伸的主要 EST 和 cDNA 片段分别为: AW631376、AI735183、AB040904 和 AL110210 等。随后, 据此 cDNA 序列设计并合成两端特异引物, 以人胎肝 cDNA 为模板 PCR 扩增出预期长度的 cDNA 序列, 再将其克隆至 T 载体后测序, 结果只有三个碱基与预期不同。该序列全长为 5248 bp (图 1), 含 4908 bp 长的 ORF, 编码一个具 1636 aa 的蛋白, 估计分子质量为 178.8kDa, 等电点为 6.45; 其 3' 非翻译区 (UTR) 为 279bp, 含多聚 A (20 个 A) 及典型的 AATAAA 加尾信号; 5'UTR 为 61 bp。起始密码子位于第 62~64 核苷酸, 其临近序列为典型的 Kozak 序列 (GCCGCCATGG)^[9], 因而属于强起始密码子。

2.2 人 PTP-TD14 的序列分析

对人 PTP-TD14 的一级结构分析显示, 它与近期报道的大鼠 PTP-TD14 高度同源, 全程有 83% aa 相同, 86% aa 相似^[6]; 在它们的 C 端都含有约 210 aa 组成的 PTPc 结构域, 并含有 PTP 的特征

基序: (I/V)HCXAGXXR(S/T) (图 1), 因而都属于 PTP 大家族。然而, 人 PTP-TD14 比大鼠 PTP-TD14 (1494 个 aa) 长, 当使二者所有 aa 达最佳匹配时, 前者比后者在 N 端整整多出 180 个 aa (图 2)。另外, 除 C 端含 PTPc 结构域的区段外, 人 PTP-TD14 与其他已知的 PTP 在 aa 序列上没有明显的同源性。但是其 N 端约 800 个 aa 区段内含有一个 BRO1 样结构域, 该区段与芽殖酵母中参与 MAPK 信号转导途径的一种蛋白 BRO1^[10] 有 21% aa 相同, 41% aa 相似 (图 2), 并含有保守的 BRO1 盒: KDNDFIYHEXV(S/P) (图 1)。另外, 和大鼠 PTP-TD14 一样^[6], 在人 PTP-TD14 的 BRO1 样序列和 PTPc 这两个结构域之间有 2 个可能的 SH3 结构域结合位点, 分别位于第 902~908 aa 和第 1095~1101 aa (图 1)。在 PTPc 的两侧还各有一个分别由 212 aa (第 897~1108 aa) 和 53 aa (第 1521~1573 aa) 组成的脯氨酸富集区。人 PTP-TD14 与小鼠 PTP-TD14 在 PTPc 结构域和 BRO1 样区段的同源性极高, 在 2 个脯氨酸富集区的同源性则较低 (图 2)。

2.3 基因的染色体结构及定位

用人 PTP-TD14 cDNA 序列在 GenBank 对 HTG 数据库进行 BLAST 搜索, 发现一条来自人 3 号染色体的基因组序列, 它分段与人 PTP-TD14 cDNA 高度匹配。该序列是由美国 Whitehead 研究所/MIT 基因组研究中心所测序并注册的人基因组工作草图 (AC023230, 199621bp)。经分析, 人 PTP-TD14 基因至少含有 25 个外显子和 24 个内含子, 内含子与外显子的连接处均严格遵循典型的 GT/AG 剪接模式 (表 1)。25 个外显子的大小分别为 85、128、75、50、132、81、131、47、56、59、80、115、66、146、312、155、152、180、1759、185、105、139、114 和 736 bp。内含子的大小见表 1。

随着人类基因组计划的顺利实施, 被测序的人基因组序列和在染色体上精确定位的 STS 不断增加, 从而使许多新基因的染色体定位不再局限于 FISH 或 RH 等实验方法^[11]。本文即采用 UniGene 和已定位的 STS 研究了人 PTP-TD14 的染色体定位。经网上查询 (<http://www.ncbi.nlm.nih.gov/Schuler/Unigene/>), PTP-TD14 所对应的 UniGene 簇为 Hs. 25524, 它来自 102 条人源 EST; 该 UniGene 依与其高度匹配的染色体位标 stSG46169

```

ACGCGTTCDDGACGCTGCAGCAGCTACGGAGTGCCTGGTGGCGGGGGTGCACGGCCGATGAGGCGGTGCC
76
M E A V P
5
CCCATGCCCATGATCTCTGCTGACCTGAAGAGGGCGGTGACTTTCACCTTCAGCCAGCTGTGAAGAGTTTGTG
151
R M P M I W L D L K E A G D F N F Q P A V K K F V
30
CTGAAGAATTTGGAGAGAAACCCAGAGCTACAATGAAGAAGTGAAGAAGCTGGAGTTGCTCAGACAGAAATGCT
226
L K N Y G E N P E A Y N E E L K K L E L L R Q N A
55
GTGCGTGTCCACGGAGACTTTGAGGGCTGTAGTGTCTCCGCAAGTACCTCGGCCAGCTTCATTACCTCCAGAGT
301
V R V P R D F E G C S V L R K Y L G Q L H Y L Q S
80
CGGTTDDCATGGCTTGGCCAGGAGGCGGCTGTGCTTCCACCTGGACAGAGATCTTCTCAGGCAAGTCTGTG
376
R V P M G S G Q E A A V P V T W T E I F S G K S V
105
GCCATGAGGACATCAAGTACGAGCAGCGCTGTATTCTCTACAACTTCGAGCGCTGGACTCCATGCTGGGGCC
451
A H E D I K Y E Q A C I L Y N L G A L H S M L G A
130
ATGGACAGCGCGGTGTCTGAGGAGGGCATGAGGTCTCCCTGTACCCATTTCCAGTGGCAGCGGGCGCTTCCGC
526
M D K R V S E E G M K V S C T H F Q C A A G A P A
155
TAGCTAGCGGAGCAGCTTCCCTCAAGCTACAGGCTGCAGATGAGCGCCAGATCCCTTACGCTCAAGCTCAAGCTC
601
Y L R E H F P Q A Y S V D M S R Q I L T L N V N L
180
ATGCTGGCCAGGCTCAGGAGTGCCTTCTGAGAGTGCATGTTGACACACAGAGAGCTTCTTGGTGGCGCCG
676
M L G Q A Q E C L L E K S M E D N R K S F L V A R
205
ATCAGTGCACAGGTGCTAGATTACTACAGGAGGCAATCCGGCGCTTGGAGAACCCGACACTGCCCTACTGCTG
751
I S A Q V V D Y Y K E A C R A L E N P D T A S L L
230
GGCCGATCCAGAGGACTGGAAGAAGCTTGTGCAGATGAAGATCTACTACTTGCAGCGCTGGCTCATCTGCAC
826
G R I Q K D W K K L Y Q M K I Y Y F A A V A H L N
255
ATGGGAAGCAGCGCCAGGAGCGAGAGTGTGGGAGCGGTTGCATCTCCAGCGCCCTGGACAAGCTC
901
M G K Q A E B Q G C Q G R V A Y P O S A L D K L
280
AATGAAGCCATCAGTGTGGCCAGGGCCAGCCTGACACTGTGCAAGACGGCCTTGCCTTCACTATGGATGTCATT
976
N E A I K L A K G Q P D T V Q D A L R F T M D V I
305
GGGGAAAGTACAATTTGCCAGAGGACACAGCTTCAITTTACCATGAGGCTGTCCAGCATGGACACTCTT
1051
G G K Y N S A K K D N D F I Y H E A V P A L D T L
330
CAGCGTGTAAAGGAGGCGCCCTGTGGTGAAGCCCTTCCAGTGAACCCACAGACCCAGCTGTTCACAGCCCTGAC
1126
Q P V K G A P L Y K P L P V N P T D P A V Y T G P D
355
ATCTTTGCAAGTGTACCCAGCTGCCAGAGGCTGCTGACTGTACAGTGAAGGAGAGGCGCAAGCTGCTC
1201
I F A K L V P M A A H E A S S L Y S E E K A K L L
380
CGGAGATGATGGCCAGATTGAGGACAAGAATGAGTCTCGACAGTTCATGATTCATGACTGGATGCC
1276
R E M M A K I E D K N E V L D Q P M D S M Q L D P
405
GACAGCGTGGACAGCTTGCATGCTACAGCCAGTCAACAGCCAGCTCATGGAGAGTGGGGCGCTCTCAGCGTC
1351
E T V D N L D A Y S H I P P Q L M E K C A A L S V
430
CGGCGGAGACTGTGAGCAAGCTGTACAGTCCATGCAAGTGTGTCAGGTGTGTTCAAGGATGTGAGGCTCC
1426
R P D T V R N L V Q S M Q V L S G V F T D V E A S
455
CTGAAGGACATCAGAGATCTGTTGGAGGAGGATGAGCTGCTAGAGCCAGAGTTTCAAGAGCGCTGGGGCAGGCA
1501
L K D I R D L L E E D E L L E Q K F Q E A V G Q A
480
GGGCCATCTCCATCACTCCAGGCTGAGCTGGCAGAGTCAAGCGGAGAAATGGCCAAATACATGGAAGTCCAT
1576
G A I S I T S K A E L A E V R R E W A K Y M E V H
505
GAGAAGCCCTCTTCAACAGAGTGAAGTGCAGCTGACCTGTCATGAGCTCCAGCTGGCAAGCTGGCGCTGCTCAGC
1551
E K A S F T N S E L H R A M N L H V G N L R L L S
530
GGCGCGTGTGACAGTTCORRGTGCGCTGGCCAGACAGGCGCTCTGCGCAGAGCAAGCGCTGCTGCAAAAC
1726
G P L D Q V R A A L P T P A L S P E D K A V L Q N
555
CTAAAGGCACTCTGCTGAGTGAAGTGCAGAGATGCGGAGCAGCGCGTCTGCTGAGCAGCACTGCTGAGCTT
1801
L K R I L A K Y Q E M R D Q R V S L E Q Q L R E L
580
ATCCAGAAAGATGACATCACTGCTGCTGCTGCCACAGACCACTGAGAGATGAAGAGTTGTTGGAGGAGCAG
1876
I Q K D D I T A S L Y T T D H S E M K K L F E E Q
605
CTGAAGAAATGATGACAGCTGAGGTTGACTGCGAGCAGAACTGGCGGCCAGGAGCGTGTCTCTGTGTCAGCTG
1951
L K K Y D Q L K Y Y L E Q N L A A Q D R V L C A L
630
ACAGAGGCCAAGCTGCAGTACCGAGCGCTGGCGGGTACTCAGCGACTTGGACCAAAAGTGAAGTCCAGCTG
2026
T E A N V Q Y A A Y R R V L S D L D Q K W N S T L
655

```

定位于 3p21.1-21.33 (<http://www.gdb.org/gdb/>); stSG46169 也与 *PTP-TD14* cDNA 的第 3740-3893 核苷酸高度匹配(图 1)。因此,人 *PTP-TD14* 亦定位于此。

2.4 *PTP-TD14* 在人正常组织及肿瘤细胞中的表达

为了确定人 *PTP-TD14* 转录物的实际大小及组织表达谱,进行了多组织 Northern 杂交和点杂交分析。分别用 5'端 747bp(第 1~747 bp)和近 3'端的 881bp(结果未给出)为探针对成人 12 种正常组织和 7 种肿瘤细胞的 Northern 膜杂交,结果均显示,该基因呈单一转录物,大小为 5.4kb 左右(图 3),与本文获得的 *PTP-TD14* 全长 cDNA 相当。该基因在 12 种正常组织中广泛表达,相对而言,其在脑中表达最高。另外,该基因在 7 种肿瘤细胞的表达均不低于正常组织,其中在 K-562、SW480 和 G-361 的表达明显高于正常组织。来自人 50 种组织的点杂交结果亦支持该基因在各种成人和胎儿组织中广泛表达,其中以胎盘、胰腺、垂体、脑、前列腺和胎肝等组织中的表达较高(图 4)。

3 讨 论

许多 PTP 酶在造血系统中高表达,并在信号转导途径中起重要调节作用。例如,在造血细胞中特异高表达的 CD45,在 TCR 和 BCR 介导的信号转导途径中起正调节作用^[4];主要在造血细胞表达的 SHP-1(又名 PTP1C),在 BCR、TCR、EpoR、CSF-1 受体和 C-kit 受体等多种造血受体介导的信号转导途径中起负调节作用^[5],但促进由 Fas 介导的淋巴细胞凋亡及 CSF-1 诱导的巨噬细胞 MAPK 的活化^[11];在造血组织和其他组织广泛表达的 SHP-2(又名 PTP1D),则促进由蛋白酪氨酸激酶

CAGACDCTGGTGGCTGATGAGGCTATGAGGACCTGATGAAGAAGTGCAGGAGGGCAGGACTTCTAGCCA 2101
 Q T L V A S Y E A Y E D L M K K S Q E G R D F Y A 680
 GATCTGGAGACCAAGTGGCTGCTCTGCTGGAGGCCACCCAGTGCACCTGGCCAGGCCCGGAGGCTGCCGCCAG 2176
 D L E S K V A A L L E R T Q S T C Q A R E A A R Q 705
 CAGCTCTGACAGGGAGCTGAGAGAGAGCCGCGCCACGGCCACAGGCCCAGAGGCGCTGCTCCGCGCAGCG 2251
 Q L L D R E I K K K P P P R P T A P K P L L P R R 730
 GAGGAGAGTACGACGTCGAGAGCAGGAGACCCDCTGAGGAGCTGCGCAGCCTCCGCGCTGACATGCTGGCTGOC 2326
 E E S E A V E A G D P P E E L R S L P P D M V Y A G 756
 CCAGCAGTGGCTGACAGCTTCTGGGAAATGCGACCTTCAGCTTCTGCTGCGAGCCCTTCCGCGCTGCACA 2401
 P R L P D T F L G S A T P L H F P P S P F P S S T 780
 GGCCGAGGCCACTATCTCTGAGCCCTTGGCCCTGCTACTGAGGCGCCACCCAGCTGATACAGCC 2476
 G P G P H Y L S G P L P P G T Y S G P T Q L I Q P 805
 AGGCGCCAGGGCCCAATGCAATGCGCTGACGCTGGCCCTGCGCTTACCGAGCCCTGCTACACAGCCGAG 2551
 R A P G P H A M P V A P G P A L Y P A P A Y T P E 830
 CTGGGCTTGTGCGCCATCTCCGCGCAGCAGTGGGCTGCTGAGCAGTCCCTATGCTGGGGTACGGCCGGCCCA 2626
 I G L V P R S S P Q H G V Y S S P Y V G V G P A P 855
 CCAGTTCACAGTCTGDCCTGGCCGACCTCTCTAATCTCAGGCGCCAGTTCGCGCAGCGCTTGGCGCCACC 2701
 P V Y A G L S A P P P Q F S G P E L A M A V R P A 880
 AOCACCAAGTATAGATAGCATGAGGCGCCGATTCGCCAGCCACAGCCGCAAGCCGCAAGCCGCTGCTGCT 2776
 T T T V D S I Q A P I P S H T A P R P N P [] 905
 CCCCCTGCTTCCCTGCTGGCCGACCGAGCCACTCCGCGCCCTTACAGCTACCCGCGCGGCTTAGCAA 2851
 [] C F P V P P P Q P L P T P Y T Y P A G A K Q 930
 CCGATCCGCGCACAGCAGCACTCTCTCTGGGATCCCGCCAGGTTTCCAGCCCAAGGATTCGGCCCGCAGCC 2926
 P I P A Q H H F S S G I P A G F P A P R I G P Q P 955
 CAGGCGCATCTCAGCCCACTCTCAGAGCGTTTGGGCTCAGCCGCGCCAGCAGCCGCTTCCACTCAGCAT 3001
 Q P H P Q P H P S Q A F G P Q P P Q Q P L P L Q H 980
 CCAGCATCTTGGCAGCCGAGGAGCTCTACCCGCAATGCGCTAGCCGCTATGCGCTCAGCCTGGG 3076
 P H L F P P Q A P G L L P P Q S P Y P Y A P Q P C 1005
 GTCTGGGAGCCGCGCCGCGCCCTTACAGCAGCAGCTTACCCAGTCCCGCTCAGAGCCCTCTGCGAGCCAC 3151
 V L G Q P P P P L H T Q L Y P G P A Q D P L P A H 1030
 TCGGGGCTTGGCTTCCGCGCCCTGCGCCCTTACCCCTCCGCTCCGCGCTGCGCATACTGCTGCGCCCT 3226
 S G A L P P P S P G P P Q P P H P P L A Y G P A P 1055
 TCTCAGAGCCATGGGCGCCGAGCGCCCTCTTACCAATTCAGGCGCCCTGCTGCTGCGAGCTCCAGCCCT 3301
 S T R P M G P Q A A P L T I R G P S S A G Q S T P 1080
 AGTCCDCACTGGTGTCTACTGCGCCATCTCCAGGCGCTGCTGGTACCCCTGCGCCCGCAGCAGAGAA 3376
 S P H L V P S P A P S G P P [] P A A E 1105
 CCACCCCTTGGCTGGCGGAGCCGCGCTGCGAGCCTGCTCTCTCCGCGCGAGAGCCGAGCAGTGGCGCG 3451
 P P P C L R R G A A A A D L L S S S P E S Q H G G 1130
 ACTCAGTCTCTGGGGTGGCGCCGCTGCTGCGCCACCAAGCTGGATGAGCTGAGGGTGGTGGCGGCGAG 3526
 T Q S P G G G Q P L L Q P T K V D A A E G R R P Q 1155
 GGCCCTGGCTGATTGAGCGGAGCCCTATGAGCATCTGAGAGGCTGGCGCAGTTGCAAGCAGGAGCTGGAGGCC 3601
 A L R L I E R D P Y E N P E R L R Q L Q Q E L E A 1180
 TTTCGGGTCAGCTGGGGAATGGGAGCTCTGCGCAGCTTCTGGCGAGAGCTGCAAGATGCCAGGAGCATGAT 3676
 F R G Q L G D V G A L D T Y V W R E L Q D A Q E H D 1205
 GCGCGAGGCGCTTCCATGGCCATTCGCGCGCTGCTACTGCAAGAGCGCGAGCAGGATGCAAGCCCATGAGC 3751
 A A R G R S I A I A R C Y S L K N R H Q D V M P Y D 1230
 AGTAAACCTGTGCTGCTGGCTGAGCCAGGATGACTACATCAATGCGAGCTGGCTGGAGCGGCTCTCCGCAATC 3826
 S N R V V L R S G K D D Y I N A S C V E G L S P Y 1255
 TGCCCGCCCTAGTGGCAGCCAGCCGCTGCTGGCAGCCTGCTGAGCTTCTGGCTCATGCTGCATGAGCAG 3901
 C P P L V A T Q A P L P G T A A D F W L M V H F Q 1280
 AAAGTGTCACTGATCTTCACTGCTGTTTCTGAGGCTGAGATGGAGAGCAAAAAGTGGCAGCCTACTTCCGACC 3976
 K V S V I Y M L V S E A E M E K Q K V A R Y F P T 1305
 GAGAGGCGCCAGCCATGCTGCAGCCGCTGCGCTGAGCCCTGGCATTGAGCAGGCTGGCGAGCAGCCAAACCAATG 4051
 E R G Q P M Y H G A I S L A L S S V R S T E T H Y 1330
 GAGGCGCTGCTGAGCTGAGTTCGAGCAGCAGCCGCTCAAGCGCTCTCTTTCGCACTGCACTTCCGCACTTGG 4126
 E R V L S L Q P R D Q S L K R S L V H L H F P T W 1355

(PTK) 启动的信号转导途径^[12]。CD45 属于膜受体型 PTP，而 SHP-1 和 SHP-2 则属于胞质型 PTP。人源 PTP-TD14 显然属于后者，其原因主要有三：①其 C 末端有 PTP 特有的 PTPc 结构域和特征基序 (VHCSS-GVGRG)；②它与大鼠 PTP-TD14 高度同源，可能属直系同源体，而后者已由免疫组化实验定位于胞浆细胞器中；③在多个网站采用不同计算机软件分析，未在其序列中发现明显提示膜定位的跨膜区、定位胞外的信号肽或定位于核内的核定位信号。

PTP-TD14 与其他已知的 PTP 在序列上的一个显著区别是，其 N 端有一个 BRO1 样结构域，该结构可能对该酶发挥正常功能具有举足轻重的作用。有研究表明，与该区段具有同源性的几个已知蛋白均在信号转导途径中起重要作用。例如，97kDa 的酵母蛋白 BRO1 参与 MAPK 信号转导；小鼠 AIP1 蛋白(869aa)通过与能诱导细胞凋亡的 22kDa 钙结合蛋白 ALG-2 的相互作用，而与之共同调节钙依赖的细胞凋亡^[13]；AIP1 在非洲爪蟾中的同源体 Xp95(867aa)近期也被证实参与爪蟾卵母细胞成熟过程中的 PTK 信号转导途径^[14]。PTP-TD14 N 端的 BRO1 样序列、酵母 BRO1 蛋白和 Ca²⁺ 依赖的细胞凋亡相关蛋白 (AIP1)^[11]三者比较，每两两之间的同源性均极为相近(21%~22% 相同，41%~43% 相似)。

本研究所克隆的人 PTP-TD14 cDNA 全长为 5248 bp，与 Northern 杂交分析的转录本大小相当。大鼠各组织中的 PTP-TD14 也以大小相近(5.3bp)的单一转录本存在^[6]。Northern 杂交和点杂交结果均显示，人 PTP-TD14 广泛表达于造血系统和其它各种组织，提示其作为一个重要的持家基因的可能。然而，North-

```

CTGTGAGTTAGCCCTCCCGACAGCCCACTGCTGCCTTCATCCAGGAGTGCACCCACATTACTGCAT 4201
P E L G L P D S P S N L L R F I Q E Y H A H Y L H 1390
CAGCGCCGCTGCACACGCCCATCATTTGTCACTGCACCTCTGGTGTGGGGCCACGGGAGCCTTTCACACTGCTC 4276
Q R P L H T P I T I V H C S S G Y G R T G A F A L L 1405
TATGCAGCTGTGCAGAGGTGCAGGCTGGGAACGAAATCCCTGAGCTGCAGCTGGTCCGGCCATGGCCAG 4351
Y A A Y Q E Y E A G N G I P E L P Q L Y R R M R Q 1430
CAGAGAAAGCAGATGCTGCAGGGAAGCTGCACCTCAGCTTCTGCTATGAGGCACTGCTGAGACACCTGGAGCAG 4426
Q R K H M L Q E K L H L R P C Y E A Y V R H V E Q 1455
GTCCTGCAGGCCATGCTGTGCTCTCTCATATGCAAAACCTTGGCCAGTGCAMGATCAGCCAGGAAGAACCCATT 4501
Y L Q R N G V P P P C K P L A S A S I S Q K N H L 1480
CCTCAGACTCCAGGACTGCTGCTGCTGGTGGGATGTCGCCATCAGCTCCATCCAGGCCACCATTTGCCAAGCTC 4576
P Q D S Q D L V L G G D V P I S S I Q A T I A K L 1505
AGCATTCCGCTCTGGGGGTGGAGTCCCCGTTGCCAGCTTCCAGGCCCTGCAGAGCCCCAGGCCCTCCCG 4651
S T R P P G G L E S P V A S L P G P A E P P G L P 1530
CCAGCCAGCTCCAGAGTCTAACCAGATCCCATCTTCTCTCCACCCCTCTTCTCTCCCACTACTGAGGCT 4726
P A S L P E S T P I P S S S P P P L S S P L P E A 1555
CCCAGCCTANGCAGGACCCAGCTGCTGAGGCCACCTCCGGCCGCCCTCTCTCTCTCTGGAATTCCTG 4801
P Q P K E E P P V P E A P S S G P P S S L E L L 1580
GCCTCTTGACCCAGAGCCCTTCTCTCTGACAGCTCCCTCCGGGCCAAMCAGGATGAGGACAGCATAAATTT 4876
A S L T P E A F S L D S S L R G K Q R M S K H N F 1605
CTGAGGCCCATAAAGCCAGAGCTGCGGCCACCCGCCCTCTGACGACCCCTCAGCCCTTCTGATCCAGTC 4961
L Q A H N G Q G L R A T R P S D D P L S L L D P L 1630
TGGACACTCAACAGACCTTGCACAGGTTTGGCTACTGGCTTACACTACATCATCATCTCATGCCACAC 5026
W T L N K T * 1636
TGCCACACCCACAGCagcttctcagtgggccagctctctactccattctgctgctcttggccctgcttggc 5101
ccagcctgcaccctgtggggtaaatgtactgcaggctctgggtcaggtctgctctcttlatggaccggaca 5176
tttttcagctctttgtATTGAAATKATAAACCCCTGTCTGTGGCCGTAAMAAAAAAAAAAAAAAAAAAAA 5248

```

图1 人 PTP-TDI4 的 cDNA 序列及其推导的氨基酸序列

框内分别显示起始密码子、终止密码子和加尾信号，与克隆 FLF2896 的插入片段匹配的核苷酸序列用单下划线表示，3' UTR 中与染色体位标 stSG46169 匹配的核苷酸序列用小写字母表示；双下划线分别标示 BR01 盒和 PTP 特征基序，阴影部分分别显示 2 个可能的 SH3 结构域结合位点

Fig. 1 Nucleotide sequence of human PTP-TDI4 cDNA and its predicted amino acid sequence

The initiating ATG codon, a TGA stop codon and a classical polyadenylation signal are boxed. The nucleotide sequence present in the insert of FLF2896 is underlined. The nucleotide sequence in 3' UTR matched with the sequence - tagged site, stSG1698, is written in lowercase. BR01 and PTP motifs are double underlined. The two putative SH3 binding sites are shaded.

表 1 人 PTP-TDI4 基因外显子的大小及内含子与外显子的剪接点

Table 1 Intron sizes and 5' and 3' intron/exon boundaries of human PTP-TDI4 gene

Intron	Size(bp)	5' junction sequence	3' junction sequence
1	?	CTGTGAAGAAggtgagcttgccttc	cttetctctctagTTTGTCTGAA
2	?	TGCTCAGACAGgttaggagatgat	catctgcccacagAATGCTGTCCG
3	227	CCTGTCACTGgtgagagccgagc	gtgccatgtctgagGACAGAGATCT
4	669	CTACAACCTTgtgagctgctgctc	ccccctcccagGAGCGCTGCAC
5	117	TGCTGTAGGAGgtgaggagagggga	cccactcccacagGGCATGAAGGT
6	279	ACCTCATGCTGgttaagaggccctc	cagtgctgcccagGGCCAGGCTCA
7	97	TCAGTGCACAGgttagggacgggctg	ccaccctgtrctcagGTGGTAGATTA
8	92	CCGTGGCTCATgtgaggcctggggc	cttctcttcccagCTGCACATGGG
9	381	TCGGGGAGCGgtgagctacagcgag	tgccacctcccagGTTGCACTACTT
10	124	AGTGGCCAAgtaagctgaggag	tccctaaccccacagGGCCAGCCTGA
11	139	ATTGGGGGAAgtagtctgtggggg	gctcctgaccccagGTACAATCTGT

ern 杂交分析还显示，该基因在部分肿瘤细胞中表达上升，提示其与细胞增殖、分化和转化的可能联系。有报道，转染大鼠 PTP-TDI4 基因的 NIH-3T3 细胞，其由 Ha-ras 介导的转化作用被抑制 3 倍以上^[6]。无独有偶，在胰岛素介导的细胞应答过程中起负调节作用的 PTP1B，近期也被证实通过与一种锚定蛋白 p130^{CAS} 的 SH3 结构域的结合而抑制 v-crk、v-src 或 v-ras 介导的 3Y1 细胞的转化。如果在该 PTP 的 SH3 结构域结合位点发生突变，则其蛋白结合活性和转化抑制活性丧失^[15]。这也提示了人 PTP-TDI4 的 BR01 样结构域中的 SH3 结构域结合位点的重要功能。另外，由于 PTP 的功能之一是反转 PTK 的作用，许多 PTK 都是原癌基因，提示某些 PTP 作为抑癌基因的可能^[16]。人 PTP-TDI4 定位于 3 号染色体 3p21.3，而该区带已被发现容易在肾和肺等组织的肿瘤细胞中缺失^[16]，这些结果进一步支持 PTP-TDI4 与肿瘤发生的可能联系。至于在肿瘤细胞中表达的 PTP-TDI4 是否可能在几个重要功能位点发生突变，则有待进一步研究。另外，PTP-TDI4 的天然底物尚未找到，它的表征将有助于进一步认识 PTP-TDI4 的确切生物功能和作用机制。

Intron	Size(bp)	5' junction sequence	3' junction sequence
12	113	GCCTGTAAAAGgtcggggagctgaga	ttcccaccccacagGAGCCCUCTTG
13	80	TCACTGTACAGgtgggtggagggtgg	ctgtcctctctcagTGAGGAGAAGG
14	391	GAGGTCTGGAgtgagtgaggactt	tgttgcctcccacagCCAGTTCATGG
15	287	GTCCATGCAAGgtgagtaaggggcag	aatccacaaccccagTGCTGTCAGGT
16	76	CCTCTCCCAGgtgagcccaccaga	cactgtgcccacagAGGACAAGGCC
17	101	CAGAGATGAAGgtgggctgggtgagc	ccacccttctcagAAGTTGTTCGA
18	88	TTGACCAAAAgtcagtgcccagtc	ccttctgcccacagGTGGAAC TUCA
19	95	CTCTGGACAGgttctgtggccctg	ttcctgtgcacagGGAGCTGAAGA
20	1856	AGATGGAGAAGgtgagaaggggggt	ttatctgtcctcagCAAAAAGTGGC
21	128	TGGCCTGAGTTgtgagtcactgctc	catcctgtcccacagAGGCCTGCCCG
22	86	gtgcactgcagtagaggtgggct	cccttctggcacagCTCTGGTGTGG
23	94	TGCAGGAGAAGgtgaggtatgggca	aatgcacctgtcagCTGCACCTCAG
24	80	TCAGCCAGAAGgtgaggaaagtctcg	actctgttctcagAACACCTTCC

大写字母示外显子,小写示内含子;第 1 和第 2 个内含子的大小尚不能确定

Exon sequences are written in capital and intron sequences in lowercase;the sizes of intron 1 and intron 2 are uncertain.

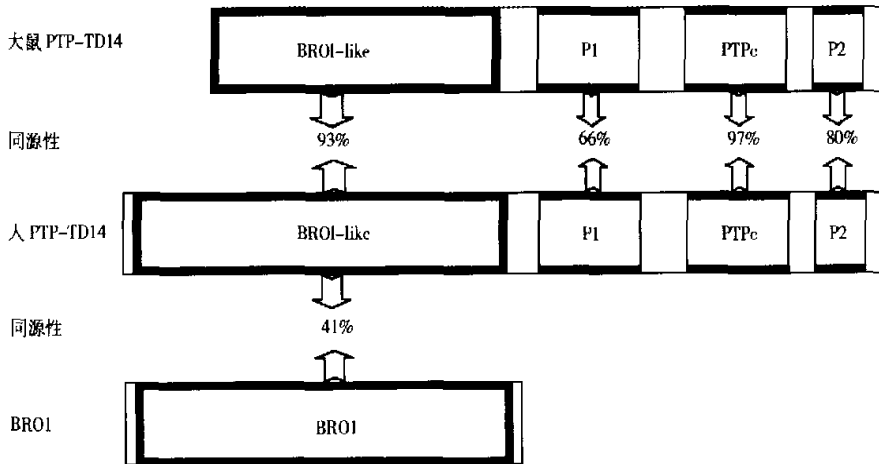


图 2 人 PTP-TD14 的结构域及其与大鼠 PTP-TD14 和酵母蛋白 BRO1 的同源性比较

P1 和 P2 分别表示 2 个脯氨酸富集区

Fig. 2 The domains in human PTP-TD14 and amino acid sequence homology among human PTP-TD14, rat PTP-TD14 and yeast BRO1

P1 and P2 indicate proline-rich regions

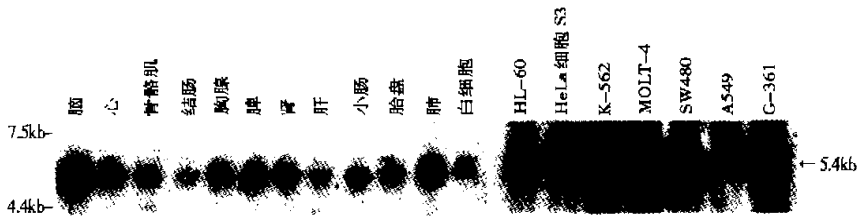


图 3 Northern 分析 PTP-TD14 在人正常组织及肿瘤细胞中的表达

Fig. 3 Northern blot analysis of PTP-TD14 transcription in normal human tissues and human cancer cell lines

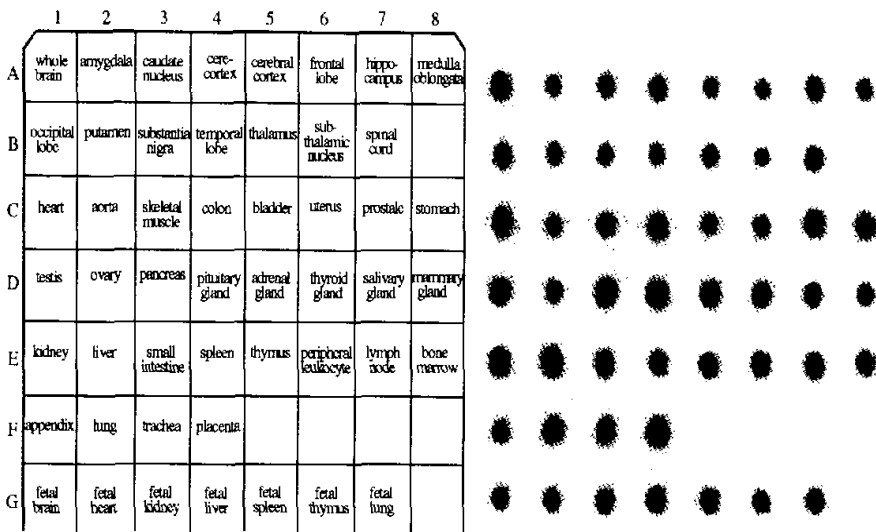


图4 点杂交分析在人50种组织中的表达

Fig. 4 Dot blot analysis of PTP-TD14 expression in human 50 tissues

参考文献(References):

- [1] 翟芸,刘孟琨,张成岗,景祥虎,周钢桥,贺福初.应用RH技术定位人胎肝的7个新基因[J].遗传,2001,23(2):93~96.
- [2] Tamir I, Dal Porto J M, Cambier J C. Cytoplasmic protein tyrosine phosphatases SHP-1 and SHP-2, regulators of B cell signal transduction[J]. Curr Opin Immunol, 2000, 12(3): 307~315.
- [3] Frearson J A, Alexander D R. The role of phosphotyrosine phosphatases in haematopoietic cell signal transduction[J]. Bioessays, 1997, 19(5): 417~427.
- [4] Cyster J G, Healy J I, Kishihara K, Mak T W, Thomas M L, Goodnow C C. Regulation of B-lymphocyte negative and positive selection by tyrosine phosphatase CD45[J]. Nature, 1996, 381(6580): 325~328.
- [5] Klingmuller U, Lorenz U, Cantley L C, Neel B G, Lodish H F. Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. Cell[J]. 1995, 80(5): 729~738.
- [6] Cao L, Zhang L, Ruiz-Lozano P, Yang Q, Chien K R, Graham R M, Zhou M. A novel putative protein-tyrosine phosphatase contains a BRO1-like domain and suppresses Ha-ras-mediated transformation[J]. J Biol Chem, 1998, 273(33): 21077~21083.
- [7] Schultz J, Copley R R, Doercks T, Ponting C P, Bork P. SMART: A Web-based tool for the study of genetically mobile domains[J]. Nucleic Acids Res, 2000, 28: 231~234.
- [8] Hofmann K, Bucher P, Falquet L, Bairoch A. The PROSITE database, its status in 1999[J]. Nucleic Acids Res, 1999, 27: 215~219.
- [9] Kozak M. Initiation of translation in prokaryotes and eukaryotes[J]. Gene, 1999, 234: 187~208.
- [10] Nickas M E, Yaffe M P. BRO1, a novel gene that interacts with components of the Pkc1p-mitogen-activated protein kinase pathway in *Saccharomyces cerevisiae*[J]. Mol Cell Biol, 1996, 16(6): 2585~2593.
- [11] Krautwald S, Buscher D, Kummer V, Buder S, Baccarini M. Involvement of the protein tyrosine phosphatase SHP-1 in Ras-mediated activation of the mitogen-activated protein kinase pathway[J]. Mol Cell Biol, 1996, 16(11): 5955~5963.
- [12] Tang T L, Freeman R M Jr, O'Reilly A M, Neel B G, Sokol S Y. The SH2-containing protein tyrosine phosphatase SH-PTP2 is required upstream of MAP kinase for early *Xenopus* development[J]. Cell, 1995, 80(3): 473~483.
- [13] Vito P, Pellegrini L, Guet C, D'Adamo L. Cloning of AIP1, a novel protein that associates with the apoptosis-linked gene ALG-2 in a Ca²⁺-dependent reaction[J]. J Biol Chem, 1999, 274(3): 1533~1540.
- [14] Che S, El-Hodiri H M, Wu C F, Nelman-Gonzalez M, Weil M M, Etkin I D, Clark R B, Kuang J. Identification and cloning of Xp95, a putative signal transduction protein in *Xenopus* oocytes[J]. J Biol Chem, 1999, 274(9): 5522~5531.
- [15] Liu F, Sells M A, Chernoff J. Transformation suppression by protein tyrosine phosphatase 1B requires a functional SH3 ligand[J]. Mol Cell Biol, 1998, 18(1): 250~259.
- [16] LaForgia S, Morse B, Levy J, Barnea G, Cannizzaro L A, Li F, Nowell PC, Boghosian-Sell L, Glick J, Weston A. Receptor protein-tyrosine phosphatase gamma is a candidate tumor suppressor gene at human chromosome region 3p21[J]. Proc Natl Acad Sci USA, 1991, 88(11): 5036~5040.