

[文章编号] 1000-4718(2008)09-1734-04

人 CD44 新剪接变异体克隆及分析*

向秋¹, 何晓松^{2△}, 朱华¹, 雷迅², 凌月福², 肖胜军³, 范才文¹, 黄岚珍¹, 王建红¹
(桂林医学院¹生物技术重点实验室,²耳鼻喉科,³病理学科, 广西桂林 541004)

[摘要] 目的:应用逆转录多聚酶链式反应(RT-PCR)分析 CD44 在鼻咽癌组织和鼻咽癌细胞株中表达,来寻找新的人 CD44 剪接变异体。方法:在 CD44 基因起始和终止密码子两端及变异剪接外显子 v10 中和剪接位点处设计特异性引物,以鼻咽癌组织、细胞株 5-8F 和 HNE1 cDNA 为模板进行 RT-PCR 扩增,产物测序。然后,应用生物信息学对克隆序列进行分析。结果:新克隆的 CD44 剪接变异体有 1634bp,包含一个完整的阅读框,起始密码子在序列的 12 位,终止密码子在序列的 1301 位,可变剪接区只有变异型剪接外显子 10,预测编码 429 氨基酸。GenBank 登陆号:EF581837。结论:一个新的、预测编码 429 个氨基酸的 CD44 剪接变异体存在于研究的人鼻咽癌组织和细胞株中,它的功能有待于进一步研究。

[关键词] CD44; 剪接变异体; 鼻咽肿瘤

[中图分类号] R363 **[文献标识码]** A

Cloning and analysis of a novel human CD44 variant

XIANG Qiu¹, HE Xiao-song², ZHU Hua¹, LEI Xun², LING Yue-fu², XIAO Sheng-jun³, FAN Cai-wen¹, HUANG Lan-zhen¹, WANG Jian-hong¹

(¹Key Laboratory of Biotechnology, ²Department of Otolaryngology, ³Department of Pathology, Guilin Medical College, Guilin 541004, China. E-mail: kexue@glmc.edu.cn)

[ABSTRACT] **AIM:** To detect unknown CD44 variants (CD44v) in nasopharyngeal cancer by using reverse transcription-polymerase chain reaction (RT-PCR) to analyze the expression of cell adhesion protein CD44 gene in nasopharyngeal cancer tissue and cell lines. **METHODS:** Specific primers at up start code, down terminal code of CD44 and primers at the middle, splicing points of variable splicing exon v10 of CD44 were designed. cDNA of nasopharyngeal cancer tissues, 5-8F and HNE1 cell lines were analyzed by RT-PCR. Products of RT-PCR were sequenced and further analyzed by bioinformatics. **RESULTS:** The new CD44v sequence possessed 1634 bp with a completed open reading frame. The start code was at 12 bp site and terminal code at 1301bp site. It was predicted to code 429 amino acids, and only variable splicing exon 10 existed in the flexible region. It was given an accessible number EF581837 by GenBank. **CONCLUSION:** A new CD44 variant predicted to code 429 amino acids exists in the studied nasopharyngeal cancer tissues and cell lines.

[KEY WORDS] CD44; Spliced isoform; Nasopharyngeal neoplasms

CD44 是一个由单一基因所编码的、分布极为广泛的单链细胞表面跨膜糖蛋白家族。人类 CD44 基因定位于染色体 11p13 上,至少有 20 个外显子,长约为 50 kb^[1]。根据 CD44 外显子的表达方式不同,可分为两种类型:一种是组成型外显子,至少有 10 个,存在于所有的 CD44 转录子中;另一种是可变区(V 区)变异型剪接外显子,位于第 5 和第 6 组成型外显子之间,至少也有 10 个,不同的变异型剪接外

显子只存在于特定的 CD44 转录子中。只含组成型外显子而无变异型外显子插入的转录子称为标准型 CD44(standard form CD44, CD44s),主要在间质和造血源性细胞中表达,是透明质酸盐的主要受体,介导多种正常生理现象。如:体外淋巴细胞的黏附、B 淋巴细胞的增殖及迁移等^[1]。含有变异型外显子插入的 CD44 转录子统称为变异型 CD44 (variance form CD44, CD44v),主要在上皮源性细胞和肿瘤细胞中

[收稿日期] 2007-12-14 [修回日期] 2008-05-23

* [基金项目] 广西壮族自治区卫生厅项目(No. Z2007236);广西壮族自治区科技厅科技攻关项目(No. 0592007-1J);广西高校病理学人才小高地建设创新团队基金资助项目(No. 桂教人[2005]34)

△通讯作者 Tel:0773-5893325; E-mail:kexue@glmc.edu.cn

表达^[1,2]。研究发现,许多肿瘤细胞中有多种 CD44v 表达,CD44 尤其是 CD44v 在肿瘤侵袭、转移各环节发挥重要作用^[3]。鼻咽癌(nasopharyngeal cancer, NPC)是一种侵袭性强,区域淋巴结及远隔器官转移率高,发生侵袭、转移时间早的恶性肿瘤。迄今,NPC 侵袭、转移的机制尚不清楚^[4]。新的 CD44v 分离、克隆对鼻咽癌侵袭、转移的分子机制阐明具有重大意义。因此,本研究旨在分析克隆鼻咽癌组织和细胞中 CD44 剪切变异体。

材 料 和 方 法

1 材料

1.1 鼻咽癌组织、细胞株 鼻咽癌活检组织,来源于桂林医学院附属医院耳鼻喉科;鼻咽癌细胞株 5-8F 和 HNE1 来源于湘雅医学院肿瘤研究所。

1.2 试剂盒 Trizol RNA 提取试剂盒购自 Invitrogen;PCR 试剂盒、M-MLV 逆转录试剂盒、T-Easy 载体克隆试剂盒购自 Promega;PCR 产物纯化和质粒抽提试剂盒购自上海华舜生物工程公司。

1.3 工具酶和试剂 核酸分子质量标准 DL2000 marker、T4 DNA 连接酶、DNase I、ribonuclease inhibitor 购自大连宝生物公司(TaKaRa);琼脂糖购自 Promega。

1.4 PCR 引物和测序 引物由上海生物工程技术服务有限公司合成;测序由上海英骏生物技术有限公司完成。

2 方法

2.1 组织、细胞总 RNA 的抽提及 cDNA 合成 按照 Trizol 试剂盒操作说明书提取鼻咽癌组织、细胞株总 RNA, DNase I 消化 RNA 中痕量 DNA,紫外分光光度计测定其浓度;PCR 扩增 GAPDH 检测基因组 DNA 消化是否完全。cDNA 的逆转录(RT)合成按照试剂盒操作说明书进行。

2.2 CD44 基因 PCR 扩增 根据 GenBank 数据库中的 CD44 基因 cDNA 序列设计上游引物 P1:5'-ACC CCA TCC CAG ACG AAG ACA GTC-3',位于 CD44 组成型外显子 5 中,下游引物 P2:5'-GGG ATG AAG GTC CTG CTT TCC TTC G-3',位于 CD44 剪接变异型外显子 10(v10)中。以鼻咽癌细胞株 5-8F、HNE1 和鼻咽癌组织 cDNA 为模板进行 PCR 扩增,PCR 反应按照试剂盒操作说明书进行:50 μL PCR 的反应体系中含 KCl 50 mmol/L, MgCl₂ 1.5 mmol/L, dNTP 200 mmol/L, Tris-HCl 10 mmol/L, pH 8.0, cDNA 模板 2 μL, Taq 酶 3 U, 待检测基因引物 0.1 mmol/L。PCR 反应参数为:94 °C 4 min;94 °C

30 s, 57 °C 30 s, 72 °C 60 s, 35 个循环;72 °C 10 min。琼脂糖凝胶(10 g/L)电泳检测有 2 条扩增产物条带。

2.3 PCR 产物克隆、测序与分析 采用 PCR 产物纯化试剂盒纯化目的片段,然后将目的片段和 T-Easy 载体连接,连接产物转化 JM109 大肠杆菌,挑取阳性克隆扩大培养,按少量质粒提取试剂盒操作说明书分离、纯化质粒。克隆质粒送上海英骏生物技术有限公司进行核酸序列测定。采用 NCBI GenBank 数据库软件对所得序列进行分析,长的扩增产物有 372 bp,来源于已知的 CD44v8-10 剪接变异体的扩增产物。短的扩增产物 181 bp,为一新的 CD44 剪接变异体来源。

2.4 新 CD44 变异体编码区序列的分离、克隆 以鼻咽癌 5-8F 细胞株 cDNA 为模板,根据人 CD44 基因 cDNA 序列,在起始密码子上游设计左边引物 P3:5'-CTC CGG ACA CCA TGG ACA AGT TTT G-3',在 181 bp 片段的剪接位点处设计特殊右边引物 P4:5'-CAT CAT TCC TAT TGG TAG CAG GGA T-3',扩增起始密码子端至 v10 间的基因片段。在 181 bp 片段的剪接位点处设计特殊左边引物 P5:5'-ATC CCT GCT ACC AAT AGG AAT GAT G-3',在 CD44 终止密码子下游设计右边引物 P6:5'-GAA AGG TTG GCG ATC AGG AAT ACA-3',扩增 v10 至终止密码子端序列。PCR 扩增条件为:94 °C 4 min;94 °C 30 s, 57 °C 30 s, 72 °C 2 min, 35 个循环;72 °C 10 min。然后进行 PCR 扩增产物的分离、纯化与测序。

结 果

1 组成型外显子 5 和变异型外显子 v10 间的基因片段扩增

以鼻咽癌组织和鼻咽癌细胞株 cDNA 为模板, P1、P2 为引物进行 PCR 扩增,凝胶电泳检测有两条大小不等的扩增产物条带,如图 1 所示。长的条带测序 372 bp,为已知的人 CD44v8-10 变异体扩增产物;短的扩增产物测序 181 bp,序列为:ACCCATC-CCAGACGAAGACAGTCCCTGGATCACCGACAGCAG-AGACAGAATCCCTGCTACCAATAGGAATGATGTCAC-AGGTGGAAGAAGAGACCCAAATCATTCTGAAGGCT-CAACTACTTTACTGGAAGGTTATACCTCTCATTACTC-ACACACGAAGGAAAGCAGGACCTTCATCCC。来源于未知的 CD44 剪接变异体。

2 新 CD44 变异体完整编码区 cDNA 的分离与分析

以鼻咽癌 5-8F 细胞 cDNA 为模板, P3、P4 为引

物扩增起始密码子端序列,测序为 690 bp;以 P5、P6 为引物扩增终止密码子端序列,测序为 969 bp,如图 2 所示。2 个扩增产物序列拼接后为 1 634 bp。分析表明,新 CD44 变异体起始密码子和终止密码子分别位于组成外显子 1 和 9,组成型外显子 5 和 6 之间只有变异剪接外显子 v10,如图 3 所示。

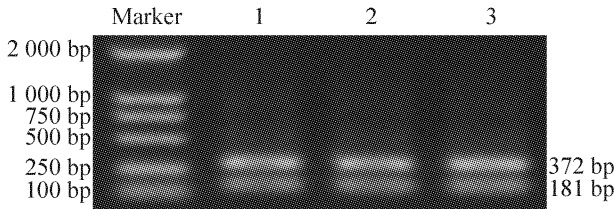


Fig 1 RT-PCR analysis of CD44 isoform expression in nasopharyngeal cancer tissue and cell lines. Marker: DNA molecular mass marker DL2000; Lane 1: Nasopharyngeal cancer cell line 5-8F; Lane 2: Nasopharyngeal cancer cell line HNE1; Lane 3: Nasopharyngeal cancer tissue.

图 1 RT-PCR 分析 CD44 变异体在鼻咽癌组织和细胞中的表达

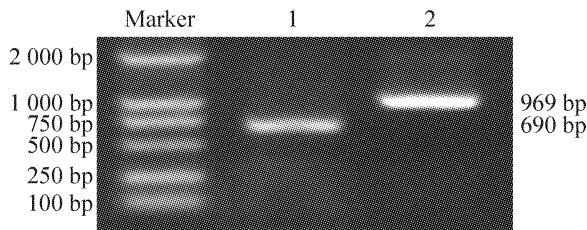


Fig 2 RT-PCR cloning of novel CD44 isoform open reading frame. Marker: DNA molecular mass marker; Lane 1: RT-PCR product with P3 and P4; Lane 2: RT-PCR product with P5 and P6.

图 2 RT-PCR 分离新 CD44 变异体完整的编码区序列

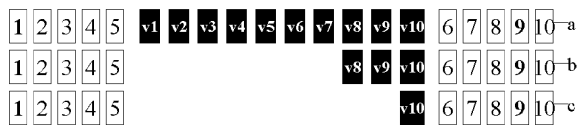


Fig 3 Analysis of CD44 gene exon constitution. 1,2,3,4,5,6,7,8,9,10: Common exons of CD44; v1, v2, v3, v4, v5, v6, v7, v8, v9, v10: Flexible splicing exons of CD44 within variable region; a: CD44 exon constitution map; b: CD44v8-10 exon constitution map; c: Novel CD44 variant exon constitution map.

图 3 CD44 基因外显子构成分析

3 新 CD44v 基因序列分析

采用 NCBI GenBank 数据库软件分析,新 CD44 变异体编码 429 氨基酸,完整的编码区序列起始密码子在 12 位,终止密码子在 1301 位,见图 4。

Length: 429 aa.

1 ctceggacacc

```

12 atggacaagtttggcgcagcagcctgggactctgcctcgtg
M D K F W W H A A W G L C L V
57 ccgctgagcctggcgcagatcgattgaatataacctgccgctt
P L S L A Q I D L N I T C R F
102 gcaggtgtattccacgtggagaaaaatggctcagacatctct
A G V F H V E K N G R Y S I S
147 cggacggaggccgctgacctctgcaaggcttcaatagcaccttg
R T E A A D L C K A F N S T L
192 cccacaatggcccagatggagaaagctotgagcatcggattgag
P T M A Q M E K A L S I G F E
237 acctgcaggtatgggttcatagaaggcgcagctgggtattcccg
T C R Y G F I E G H V V I P R
282 atccaccacaactccatctgtgcagcaaacacacaggggttac
I H P N S I C A A N N T G V Y
327 atctcacatccaacacctcccagatgacacatattgcttcaat
I L T S N T S Q Y D T Y C F N
372 gcttcagctccacctgaagaagattgtacatcagtcacagacctg
A S A P P E E D C T S V T D L
417 cccaatgcctttgatggacaattaccataactattgttaacctg
P N A F D G P I T I T I V N R
462 gatggaccgcctatgtccagaaggagaatacagaacgaatcct
D G T R Y V Q K G E Y R T N P
507 gaagacatctaccaccagcaacccctactgatgatgacgtgagcagc
E D I Y P S N P T D D D V S S
552 ggctcctccagtgaaaggagcagcacttcaggaggttacatcttt
G S S S E R S T S G G Y I F
597 tacaccttttctactgtacaccccatcccagacagacagctccc
Y T F S T V H P I P D E D S P
642 tggatcaccgacagcacagacaagaatccctgctaccaataggaat
W I T D S T D R I P A T N R N
687 gatgtcacagtggaagaagagaccacaatcattctgaaggctca
D V T G G R R D P N H S E G S
732 actactttactggaagtttactcctcattaccaccaacacagaag
T T L L E G Y T S H Y P H T K
777 gaaagcaggaccttcatcccagtgacctcagctaaactgggtcc
E S R T F I P V T S A K T G S
822 ttggagttactgcagttactgtggagattccaactcctaagtc
F G V T A V T V G D S N S N V
867 aatcgttccattcaggagaccaagacacattccacccagtgagg
N R S L S G D Q D T F H P S G
912 gggctccataccactcatgatctgaatcagatggacactcacat
G S H T T H G S E S D G H S H
957 gggagtcaagaagtgagcaaacacaacctctgtcctataagg
G S Q E G G A N T T S G P I R
1002 acaccacaattccagaatggctgatcatttgatccctctgtg
T P Q I P E W L I I L A S L L
1047 gccttgcttggatttcttgcagttgcatgagtcacacagtcga
A L A L I L A V C I A V N S R
1092 agaaggtgtggcagaagaaaaagctagatcaacagtgcaat
R R C G Q K K K L V I N S G N
1137 ggaagctgtggagacagaaagccaagtggaactcaacggagagcc
G A V E D R K P S G L N G E A
1182 agcaagctocagaaatgggtcattgttgaacaaggagctgtca
S K S Q E M V H L V N K E S S
1227 gaaactccagaccagtttatgacagctgatgagacaaggaaacct
E T P D Q F M T A D E T R N L
1272 cagaatgtggacatgaagattgggtgtaa 1301
Q N V D M K I G V *
    
```

```

1302 cacctacaccatta tcttg
1321 gaaag aaacaacctg tggaaacata accattacag ggagctggga cactaacag atgca
1381 atgtg ctactgattg ttctattgag aatctttttt agcataaaat ttctactct ttttg
1441 ttttt tgtgtttttg tctttaaagt caggtccaat ttgtaaaaac agcattgctt tctga
1501 aatta gggccaactt aataatcagc aagaatttga togttccagt toccacttgg aggcc
1561 ttcca tcctcgggtg gtgctatgga tggcttctaa caaaaactac acatattgat tcttg
1621 atgc caacctttc 1634
    
```

Fig 4 Analysis of new CD44v sequence

图 4 新 CD44v 序列分析

讨 论

真核基因在结构上的不连续性是近十年来生物学上的重大发现之一。大量研究表明,对于同一个基因,其剪接位点和拼接方式可以发生改变,从而导

致同一个基因可能有多个转录产物,表达出多个不同的相关蛋白产物,行使不同的生理功能^[5]。这种现象称为 RNA 的选择性剪接,也称可变剪接(alternative splicing)。RNA 剪接,特别是选择性 RAN 剪接是真核基因表达调控研究的重要内容之一,由于 RNA 的选择性剪接不涉及遗传信息的永久性改变,因而是一种比较灵活的真核基因表达调控方式。研究同一基因的不同剪接形式,对进一步研究基因的功能及其表达调控具有重要的意义。

人类基因 mRNA 表达量多于基因数量的事实,提示基因功能多态性与基因选择性剪接等转录后的调控有关^[6]。分析正常人子宫内膜、不典型增生子宫内膜和子宫内膜癌,发现这三种组织都有 CD44v6 mRNA 的表达,其表达率和相对含量呈显著性递增的趋势^[7]。CD44s 和 CD44v6 的表达异常在胃腺癌的发生、发展中具有重要的作用,CD44v6 表达可以用来预测胃癌的淋巴转移^[8]。Lee 等^[9]报道 CD44v5 在正常胸腺中不表达,而在胸腺瘤中的表达率高达 67%,其表达率与胸腺癌的侵袭性呈密切正相关,是一种胸腺癌侵袭的标志。结直肠癌组织中 CD44v3 和 CD44v4 的表达明显高于正常粘膜组织,CD44v3 的过度表达与癌的血管侵袭有关,CD44v4 的过度表达与癌的侵袭深度有显著相关趋势^[10]。

CD44 异常表达与多种肿瘤的发生、演变和转移密切相关。CD44s 在非小细胞肺癌(NSCLC)中表达高低与转移有关,未发生转移的肺癌病例 CD44s 阳性表达率明显高于已发生转移的病例^[11]。乳腺癌患者血清 sCD44s 和 sCD44v6 水平明显高于正常人和乳腺良性疾病患者,Ⅲ、Ⅳ期患者血清 sCD44s 和 sCD44v6 明显高于 I、II 期,检测乳腺癌患者血清中 sCD44s 和 sCD44v6 水平可作为诊断乳腺癌的辅助指标^[12]。CD44 异常表达可早于 *ras*、*p53* 等基因的异常,CD44 的异常可能与 *ras* 癌基因的激活有关,是癌形成的一个重要因素。

鼻咽癌是一种恶性的肿瘤,早期就能发生侵袭与转移。鼻咽癌组织和鼻咽癌细胞系中新发现的 CD44 剪接变异体,预测编码 429 氨基酸,在组成型外显子 5 与 6 之间,只插入 v10 一个变异型剪接外显子。这个新的 CD44 变异体与鼻咽癌的发生、发展、

侵袭和转移的关系有待于进一步研究。

[参 考 文 献]

- [1] Haynes BF, Liao HX, Patton KL, et al. The transmembrane hyaluronidase receptor (CD44): multiple functions, multiple forms[J]. *Cancer Cells*, 1991, 3(9): 347-350.
- [2] Sugahara KN, Hirata T, Hayasata H, et al. Tumor cells enhance their own CD44 cleavage and motility by generating hyaluronan fragments[J]. *J Biol Chem*, 2006, 281(9): 5861-5868.
- [3] Jothy S. CD44 and its partners in metastasis[J]. *Clin Exp Metastasis*, 2003, 20(3): 195-201.
- [4] 李桂源, 刘华英, 周 鸣, 等. 鼻咽癌癌变的分子机理[J]. *生物化学与生物物理进展*, 2006, 33(10): 922-931.
- [5] Sorek R, Shamir R, Ast G. How prevalent is functional alternative splicing in the human genome[J]. *Trends Genet*, 2004, 20(2): 68-71.
- [6] Babushok DV, Ostertaq EM, Kazazian HH. Current topics in genome evolution: molecular mechanisms of new gene formation[J]. *Cell Mol Life Sci*, 2007, 64(5): 542-554.
- [7] 王 厢, 王 敏, 张 颖, 等. 子宫内膜癌组织中 CD44 基因外显 v6 的检测及临床意义[J]. *中国实用妇科与产科杂志*, 2003, 19(3): 165-167.
- [8] Joo M, Lee HK, Kang YK. Expression of E-cadherin, beta-catenin, CD44s and CD44v6 in gastric adenocarcinoma. Relationship with lymph node metastasis[J]. *Anticancer Res*, 2003, 23(2B): 1581-1588.
- [9] Lee SC, Ham HJ, Lin TS, et al. Prognostic significance of CD44v5 expression in human thymic epithelial neoplasms[J]. *Ann Thorac Surg*, 2003, 76(1): 213-218.
- [10] Ni HM, Leong AF, Cheong D, et al. Expression of CD44 variants in colorectal carcinoma quantified by real time reverse transcriptase polymerase chain reaction[J]. *Lab Clin Med*, 2002; 139(1): 59-65.
- [11] 王晓玫, 成志强, 王 玲, 等. CD44s 在肺癌中的表达及其意义[J]. *中国病理生理杂志*, 2004, 20(8): 1489-1491.
- [12] 杨振林, 郑树森, 王荣华, 等. 血清 sCD44s 及 sCD44v6 变化与乳腺癌的初步研究[J]. *中国病理生理杂志*, 2004, 20(3): 456-458.