研究报告

Identification and Characterization of a Nasopharyngeal Carcinoma (NPC) Related Gene Which Shows The Hallmark Characteristics of a Processed Pseudogene*

XIANG Juan-Juan, YU Ying, WANG Jie-Ru, ZHU Shi-Guo,
ZHANG Bi-Cheng, LI Zhong-Hua, LI Gui-Yuan**
(*Cancer Research Institute, Xiangya School of Medicine, Central South University, Changsha* 410078, *China*)

Abstract A nasopharyngeal carcinoma(NPC) related gene, named NAG73 which mapped to 3p25~ 26 was identified by positional candidate cloning. Structural analysis reveals that NAG73 has an intronless genomic sequence, which is homologous with the fourth exon and 3′ UTR of Homo sapiens Growth Hormone Secretagogue Precursor Gene (GHRELIN). NAG73 has no 5′ promoter sequence, and has putative polyA. The results suggest that NAG73 might be a processed pseudogene of GHRELIN. It has been showed that NAG73 is actively transcribed in some cells and tissues examined. There is expression difference between normal nasopharynx epithelia and nasopharyngeal carcinoma epithelia. And there is potential for a NAG73 encoded translation product. So, it is possible that NAG73 is translated and the product might act on the tumorigenesis of nasopharyngeal carcinoma.

Key words processed pseudogene, nasopharyngeal carcinoma, gene cloning, chromosome 3p

There are two genetic events required to mutationally inactivate tumor suppressor genes. First, that is either inherited as a germ-line lesion or somatic mutation. Second, that is complete or partial loss of chromosome which is observed as loss of heterozygosity(LOH)[1].

Genetic imbalance is considered to be an important reason for the tumorigenesis of nasopharyngeal carcinoma (NPC). It has been reported that there is high frequency LOH on short arm of chromosome 3 in genome of NPC[2,3], especially on 3p25~ 26[4], suggesting the potential tumor suppressor genes associated with NPC might be located on this region.

Here we have selected one expressed sequence tag (EST), which is located on 3p25 ~ 26 and downexpressed in nasopharyngeal carcinoma compared with normal nasopharynx tissue, as a candidate tumor suppressor gene. Using positional candidate cloning, we have identified and characterized a novel gene associated with NPC named NAG73 that map to 3p25~ 26.

Through structural analysis, NAG73 shows the hallmark characteristics of a processed gene[5], which is intronless, has no promoter, and is homologous with the fourth exon and 3′ UTR of Homo sapiens Growth Hormone Secretagogue Precursor (GHRELIN, GenBank accession No. AF296558). According to Vanin[5], pseudogene is used to describe sequences found to be both related and defective. It suggests that NAG73 might be a pseudogene of GHRELIN. And we have found the NAG73 is highly transcribed in some tissues and cells. Transcriptional analysis reveals that NAG73 might act on the tumorigenesis of nasopharyngeal carcinoma.

1 Materials and Methods

1.1 Selecting EST associated with Nasopharyngeal Carcinoma

The previous study has demonstrated that there is high frequency LOH region on short arm of chromosome 3 in the genome of nasopharyngeal carcinoma, in which 3p25~ 26 is the highest frequent LOH region. EST AA767031, which is located on 3p25~ 26, was selected from Unigene Resources (www. ncbi. nlm. nih. gov/unigene). The clone which contains EST AA767031 was obtained from Research Genetics Cooperation of USA.

1.2 Transcriptional analysis

1.2.1 Designation and synthesis of PCR primers

Primer pairs were chosen from the 3 300 bp of EST AA767031. The primer pairs of house keeping gene GAPDH were synthesized. The product of GAPDH amplification is about 600 bp.

* This work is supported by state 863 High Technology R & D Project of China (102-10-01-05) and the Special Funds for Major State Basic Research of China (G1998051008).
** Corresponding author.
Tel: 86-731-4805446, E-mail: Ligy@public.cs.hn.cn
Received: April 26, 2001 Accepted: June 28, 2001

1.2.2 RT-PCR

Total RNA was isolated from normal human embryonic nasopharynx epithelia, human nasopharyngeal carcinoma epithelia and nasopharyngeal carcinoma biopsy. The normal human embryonic nasopharynx epithelia and human nasopharyngeal carcinoma epithelia were constructed by Cancer Research Institute of Xiang-Ya School of Medicine, Central South University, China. Nasopharyngeal carcinoma biopsy was obtained from Tumor Hospital of Hunan. Possible DNA contamination was removed by Dnase treatment, and reverse transcription followed by RT-PCR was performed to amplify a 300 bp transcript. 50 μl reaction contained 50 μmol/L KCl, 10 mmol/L Tris-HCl (pH 8.3), 2.5 mmol/L $MgCl_2$, 0.2 mmol/L each dNTP, 0.2 mmol/L each primers and 3 U Taq polymerase. Reactions were incubated for 5 min at 94 ℃, and then subjected to 30 cycles of melting (1 min at 94 ℃), annealing (50 s at 57 ℃) and elongation (1 min at 72 ℃). The final extension was for 10 min at 72 ℃.

1.3 Northern-blot analysis

The plasmid was isolated according to the method of "Molecular Cloning: A Laboratory Manual" (Cold Spring Harbor Laboratory Press, 1989). The plasmid was subjected to digestion with restriction endonuclease *Not* I and *Eco*R I. The product of digestion was labeled by ^{32}P. The Multiple Tissue Northern(MTN™) Blots was obtained from Clotech Company. MTN™ Hybond N is blotted 8 kinds of tissue mRNA including heart, brain, placenta, lung, liver, kidney, SK. muscle and pancreas. The MTN™ Hybond N hybridized with ^{32}P labeled cDNA probes. Hybridization were performed at 68 ℃ in 5 × SSC, 20 mmol/L phosphate buffer, 5 × Denharts, 10% dextran sulfate, 100 mg/L denatured sheared salmon sperm DNA.

1.4 DNA sequencing and database searches

Automated DNA sequencing was performed by TaKaRa Bio Co., China. Nucleic acid and protein database searches were performed at the NCBI server (www.ncbi.nlm.nih.gov).

2 Results

2.1 Transcriptional analysis of AA767031

RT-PCR results showed that in normal nasopharynx epithelia, EST AA767031 was actively transcribed, while in 66% biopsy of nasopharyngeal carcinoma, EST AA767031 failed to express or downexpressed (Figure 1). The results suggested that EST AA767031 might affect on the tumorigenesis of nasopharyngeal carcinoma. So, EST AA767031 is thought to be a tumor suppressor gene candidate.

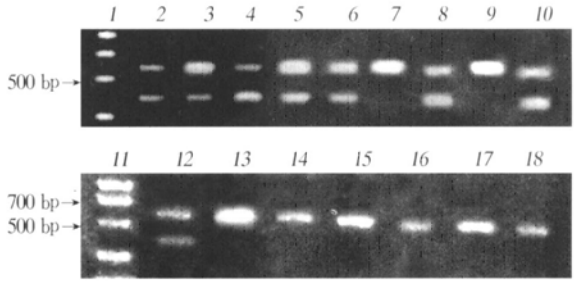

Fig. 1 RT-PCR analysis of EST AA767031

1 and *11*: markers. *2*: normal human embryonic nasopharynx epithelia. *3*: human nasopharyngeal carcinoma epithelia. *4* ~ *10*, *12* ~ *18*: nasopharyngeal carcinoma biopsy. The figure shows that EST AA767031 is down expressed or is norr-expressed in 66% nasopharyngeal carcinoma biopsy and human nasopharyngeal carcinoma epithelia, suggesting that EST AA767031 affect on the tumorigenesis of nasopharyngeal carcinoma. EST AA767031 is thought to be a tumor suppressor gene candidate.

2.2 Isolation and cloning of NAG73

Northern blot analysis of multiple tissue resulted in a transcript of about 1.3 kb, indicating that EST AA767031 which has 2.1 kb contains a full-length cDNA, named NAG73 (GenBank accession No. AF280797, Figure 2).

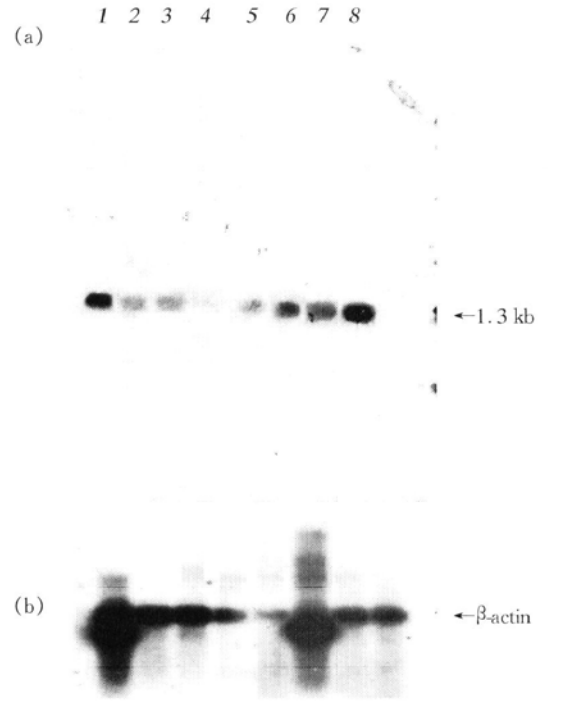

Fig. 2 Northern blot analysis of EST AA767031

(a) The MTN™ Hybond N hybridized with ^{32}P labeled AA767031 cDNA probes. Probes were obtained from digestion with restriction endonuclease *Not* I and *Eco*R I. The product of digestion was labeled by ^{32}P. The hybridization resulted in a transcript of about 1.3 kb, indicating that EST AA767031 which has 2.1 kb includes a full-length cDNA, named NAG73. NAG73 is actively transcribed in heart, brain, pancreas, placenta, liver, SK. muscle and kidney. (b) Same blot shown in upper panel probed with β-actin. The hybridization resulted in active transcript in all of the eight tissues. *1* ~ *8*: respectively heart, brain, placenta, lung, liver, SK. muscle, kidney, pancreas.

2.3 Nucleotide sequence and structural anlysis of NAG73

NAG73 was sequenced in its entirety, then the structure was analyzed. Open reading frame (ORF) analysis revealed that the open reading frame was from nt524 to nt854, the product of translation predicted to be 110 aa (Figure 3). In a search of Genbank with the sequence of NAG73, homology to the GHRELIN was identified. BLAST revealed that NAG73 was homologous with the fourth exon and 3′ UTR of GHRELIN gene (Figure 4). And NAG73 originally identified homology with Homo sapiens 3p25~26 BAC CTB-187P1 (Figure 5). Structural analysis revealed that NAG73 is an intronless sequence, which has putative polyA signals AATAAA. It was possible that NAG73 was a pseudogene of GHRELIN.

Analysis of the 5′ UTR of NAG73 revealed several potential splice sites[6] (CAG preceded by polypyrimidines) at position 62, 101, 179, 240, 326, 487, 496 (Figure 3). Potential ribosomal recognition site CTG, was shown at position 419 (Figure 3). The above structural features supported the possibility that NAG73 might be translated.

1	tgagtagggg	gcctgcagtg	tgctgtgacg	cctgggaagg	ggattctaag	ccatttccca
61	tcagccattt	actaagtagc	ctggggatct	tgttaaaatg	cagatcctga	tagattaggg
121	ctgggggtgct	gcccaagatt	ctacatttct	aagagtccct	ggatgacggt	ggcattcaca
181	gaccacagct	tctatgtgag	ggagagcagt	tgggtttctg	tctgcctgcc	gcacctgatc
241	agtagttagt	gcctgcatgt	gtggcggcca	gaattttact	tggggagacg	ctcggctcct
301	agcaccctct	gtggtagggg	ttttccagag	tgtgggcatt	accccaactg	tctctgcaga
361	cggcttcctg	catgtttccc	acaagcgctc	agatggctga	attggcaagt	ctgtggtgCT
421	Gctctttggg	gccacctcgt	tctttgcctt	tccctcccct	tagcgatgtg	tcccatccgt
481	tgcctacagt	caagtcagtc	cacatgttta	ggctaaggtg	ggcatgaccg	agtggccttc
541	ccccgggaag	accagtgtcg	taactggaat	taagttgtgg	aacataagag	tcaaggctcg
601	tgtctgctgt	gaactggagt	tgagggaatg	tttggggatc	ccacctggta	tcagtaaggg
661	aaccatggcg	acggccagcc	ttgcccatgt	gaggcatttg	ctttgtcagg	ccttctcagt
721	agtggagaag	ggaggaagga	gaatgcagct	ctttcagtgc	tgtctagcag	tgcccaagag
781	ccgtgattgg	gcacctcact	taacgtctaa	cttcagattc	actcttgggc	attcctgtct
841	tcctctccaa	tcttgaatgg	atgttggctt	cgataatgtc	atcctgaagt	ttctttgtcc
901	acacagccct	ggctggttgt	taataagctg	ttaatgcagc	cttgcactca	ggaagccctg
961	atgtttaaag	gaactgtgtc	tttgttcttc	ctctcttccc	tctttctctt	agtcccactt
1021	tatctcttct	ttctcttcct	ctccctcctc	cttttccccc	ctccctccta	ctctcccccc
1081	ttctctcctc	cttctctaag	aaacccagtc	ttaattccac	acaatgcatg	cacatgcagt
1141	agcttctctg	cttgagtgag	ctggtgtgat	taggttttct	aaacatgcac	attggccttg
1201	ctacttgtcc	ttttattccc	ttcccacaga	ccataaacca	agaattattt	ttatttgtat
1261	tattttgatt	tttttaaagt	aaaatattaa	cttttcctct	ttgaAATAAA	ttcccatttg
1321	gaacatcaaa	aaaaaaaaaa	aaaaaaaaaa	aaa		

Fig. 3 NAG73 was sequenced in its entirety

Open reading frame (ORF) analysis revealed that the open reading frame was from nt541 to nt854, the products of translation predicted to be 110 aa. The start code ATG and the stop code TGA is underlined. Polyadenylation signal AATAAA is shown in capital characters. Potential splice acceptor sites(CAG) are underlined. Potential ribosomal recognition site CTG is shown in capital characters. The above structural features supported the possibility that NAG73 might be translated.

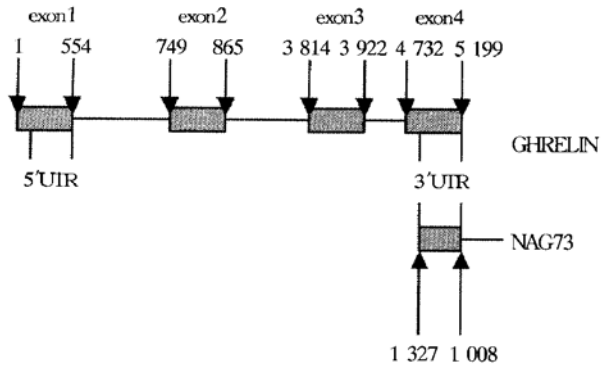

Fig. 4 NAG73 is homologous with the fourth exon and 3′ UTR of GHRELIN gene

It suggests that NAG73 is related to GHRELIN, but defective. The solid blocks represent the exons. The solid lines represent flanking and intervening sequences.

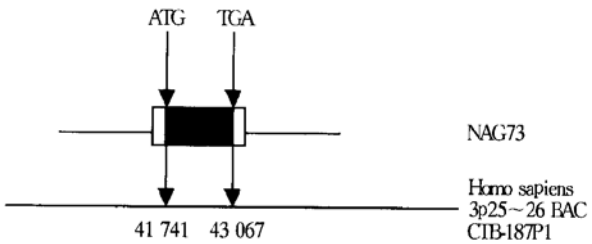

Fig. 5 Alignment of NAG73 and Homo sapiens 3p25~26 BAC CTB-187P1

The ORF of NAG73 is completely homologous with Homo sapiens 3p25~26 BAC CTB-187P1 which lacks the intervening sequence. Coded exons are shown as black box, and 5′ UTR and 3′ UTR are shown as white box.

3 Discussion

A pseudogene is a sequence which is present in

the genome and typically is characterised by close similarities to one or more paralogous functional gene. In general, pseudogene is thought to be non-function. This lack of function is a result of either failure of transcription or translation.

A fundamental feature of pseudogenes is that their nucleotide sequences differ from those of the paralogous function genes at crucial points. Pseudogenes have some common features. First, most of these pseudogenes contains multiple genetic lesions. Second, pseudogenes lack the intervening sequences. Third, many of these pseudogenes have a polyA sequence.

How do pseudogenes occur? Generally, pseudogenes are a consequence of gene duplication which can occur by retrotransposition or by the duplication of genomic DNA. Pseudogenes arising by retrotransposition are occurred after reverse transcription of single-stranded RNA by RNA polymerase II and insertion in the genome as double-stranded sequence.

NAG73 is a full length cDNA sequence located on 3p25~ 26. The structural analysis revealed that NAG73 had only one exon. And NAG73 is homologous with the fourth exon and 3′ UTR of GHRELIN gene. NAG73 has a polyA sequence. These structure features correspond with a pseudogene.

In the human genome, active transposons are rare. The majority of retrotransposed genes are inactivated to processed pseudogenes. There are many repeats such as Alu, SINE distributed in the human genome through transposition. These sequences looks like non-functional. But why the sequences exist after the selective pressure. And it is intensely discussed that if pseudogenes really are non-functional. The structure of some gene is defective, similar to pseudogenes, but they are functional. Retrotransposition can results in functional, intronless genes, or be inactivated to pseudogenes as well. Duplication of DNA is not only the reason of occurrence of pseudogenes, but also explain the generation of gene families from a single ancestral gene. It is difficult to differentiate some pseudogenes and some functional genes. Processed pseudogenes have been found to be part of many different gene families.

Referring to NAG73, Northern blot resulted in active transcription of NAG73 in multiple tissues. Furthermore, there are expression differentia between normal nasopharynx epithelia and nasopharyngeal carcinoma. The result seems to suggest that NAG73 has effect on the tumorigenesis of nasopharyngeal carcinoma. The functional relevance of pseudogene transcripts remains unclear. Sequence analysis indicates possible translational initiation sites. The results of this study suggested that based on hallmark sequence characteristics of processed pseudogenes with NAG73, along with the transcript analysis, NAG73 was a processed pseudogene which is transcrible. The following study will be focused on the functional research of NAG73, making clear that if the intronless product can be detectable.

References

1 Teng D H-F, Hu R, Lin H, *et al*. MMAC1/PTEN Mutations in primary tumor specimens and tumor cell lines. Cancer Research, 1997, **57**(23): 5221~ 5225

2 Hu L F, Eiriksdottir G, Lebedeva T, *et al*. Loss of heterozygosity on chromosome arm 3P in nasopharyngeal carcinoma. Genes, Chromosomes & Cancer, 1996, **17**(2): 118~ 126

3 Huang D P, Lo K W, Choi P H, *et al*. Loss of heterozygosity on the short arm of chromosome 3 in nasopharyngeal carcinoma. Cancer Genet Cytogenet, 1991, **54**(1): 91~ 99

4 Deng L, Jing N, Tan G, *et al*. A common region of allelic loss on chromosome region 3p25.3~ 26.3 in nasopharyngeal carcinoma. Genes, Chromosomes & Cancer, 1998, **23**(1): 21~ 25

5 Vanin E F. Processed pseudogenes: characteristics and evolution. Ann Rev Genet, 1985, **19**: 253~ 272

6 Fujii G H, Morimoto A M, Berson A E, *et al*. Transcriptional analysis of PTEN/MMAC1 pseudogene, ΨPTEN. Oncogene, 1999, **18**(9): 1765~ 1769

一个有假基因特点的鼻咽癌相关基因的分离和克隆*

向娟娟 余 鹰 王洁如 朱诗国 张必成 李忠花 李桂源**

(中南大学湘雅医学院肿瘤研究所, 长沙 410078)

摘要 用定位候选克隆法克隆了一个与鼻咽癌相关的基因, 命名为 NAG73 基因. 结构分析显示 NAG73 的基因序列无内含子, 无典型的启动子序列, poly A 尾. 与人类生长激素促分泌因子基因的第 4 个外显子和 3′ 端非翻译区同源. 说明 NAG73 可能是人类生长激素促分泌因子基因的假基因. 同时, 实验证明 NAG73 在一些细胞和组织中活跃表达. 在正常鼻咽上皮细胞, 鼻咽癌上皮细胞和鼻咽癌活检组织, NAG73 有表达差异. 序列分析显示 NAG73 有编码翻译产物的可能. NAG73 可能

可编码产物. 该产物在鼻咽癌的发病发展中发挥作用.

关键词 假基因, 鼻咽癌, 基因克隆, 染色体 3p

学科分类号 R739.63

* 国家高技术“863”计划资助项目 (102-10-01-05) 和国家重点基础研究发展规划项目 (G1998051008).

** 通讯联系人. Tel: 0731-4805446, E-mail: Ligy@public.cs.hn.cn

收稿日期: 2001-04-26, 接受日期: 2001-06-28

P53 家族新成员——P73 基因

端木德强 李中海 王敬泽

(中国科学院动物研究所, 生物膜与膜工程国家重点实验室, 北京 100080)

P53 基因的突变, 失活及缺失与 50% 的人类肿瘤发生有关. 最近, 又发现一新“P53 样”抑癌基因 P73, 定位于 1p36.2-1p36.3, 由 13 个外显子及 12 个内含子组成, 该基因编码的蛋白质在结构和功能上均与 p53 蛋白相似. P73 基因转录剪切可产生多种同种型 p73 蛋白 α, β, γ 及 δ. p73α, β 低水平表达具广泛性; 在外周血淋巴细胞、原始角质细胞及不同肿瘤细胞株和白血病中可检测到 p73γ 及 δ. 正常人类组织中 p73 低表达, 很多肿瘤细胞却高表达 p73, 研究表明对于癌变细胞最普遍的是 p73 的过表达而不是丢失: p73 蛋白过表达发生在胸腺癌、成神经细胞瘤、肺癌等多种癌变组织中.

p73 蛋白可受多种蛋白质调节作用, 参与多种细胞生理过程. a. 腺病毒 (Ad) 早期蛋白对 p73 的作用: E1A 可通过抑制 MDM2 的表达增强 p73 蛋白的稳定性; p300 对 MDM2 介导的蛋白质降解也有一定作用, E1A 也能抑制 p300 参与的降解途径, 使 p73 蛋白稳定性增强. E4orf6 蛋白不影响 p73β 的转录激活功能, 而更有效地结合和抑制 p73α. E1B55Kda 蛋白和 SV40T 抗原不能和 p73 作用: E1B55Kda 蛋白结合在 p53 氨基端 TAD 区, T 抗原结合在 DBD 区. 这些结合位点不保守, 或在 p73 对应结构上不能和病毒癌蛋白接触. p53 蛋白和 E1B55Kda 相作用的两个重要氨基酸 Lys-24, Pro-27 在 p73 上不存在. 总之, Ad 感染通过不同机制使 p53, p73 蛋白失活. b. p73 和 E2F-1 介导 TCR-AICD: T 细胞表面受体 (TCR) 激活, 通过 TCR-AICD 过程 (TCR-activation-induced cell death) 诱导周边 T 细胞凋亡, 该过程不受 p53 作用; E2F-1 和 p73 功能无效的 T 细胞不能使 TCR 介导的凋亡发生, 可见 TCR-AICD 依赖于 p73 及 E2F-1 两者. c. p73 参与骨髓白血病细胞分化: 与正常粒细胞和淋巴细胞相比, 急/慢性骨髓白血病易表达较低分子质量 p73 蛋白 β, γ, δ, ε, ζ. p73 同种型能强烈地互相作用, 较短的 p73 产物可作为 p73α 的结合抑制剂, 在急性白血病中表达较短 p73 剪切产物能降低 P73 作为肿瘤抑制基因的活性. d. p73 蛋白各结构单元所受调节: YAP (Yes 关联蛋白) WW 区和 p73 蛋白 PPPPY 基元相互作用, 能把胞外信号传送到核, 而 YAP 和 p53 家族成员的不同作用可能导致其在信号转导中的不同功能. 人 p73 C 端形成一五螺旋结构和 Eph 受体酪氨酸激酶 SAM 结构类似, 但两者同源性很低 (15%~17%). p73C 端 SAM 结构不能相互作用, 可能和其他尚未鉴定的信号蛋白作用并参与发育的调节. p73 蛋白在其 DNA 结合区及寡聚化区之间包含一 PXXP 基元, 可以和 c-Abl 激酶 SH3 结构作用. p73 作为激酶的底物; γ 射线可增强 c-Abl 磷酸化 p73 的能力, 离子辐射也导致 c-Abl 在体内外都磷酸化 p73 蛋白酪氨酸残基: 这些发现定义了一包含 p73 和 c-Abl 的前体凋亡途径. e. p73 和 p53 蛋白的相互作用: p53DNA 结合区足以和 p73 相作用, 很多肿瘤细胞和正常组织相比过表达 p73, 肿瘤细胞忍受高水平 p73 有赖于突变 p53 和 p73 结合的能力: p53DNA 结合区伸展状态的改变是其与 p73 结合的关键. p53 和 p73 的结合还需 p73 序列特异性 DNA 结合区和寡聚化区以及 p73 酪氨酸的磷酸化. Marin 等发现突变 p53 和 p73α 及 β 相互作用, 至少部分依赖于 p53 蛋白 72 位氨基酸的多态性, 72 位氨基酸是精氨酸时易于同 p73 结合. p53 和 p73 的结合最终都导致 p73 结合 DNA 的能力下降, 突变 p53 蛋白下调 p73 在某些情况下能促进肿瘤发展.

近年又发现几种新的 p53 样蛋白: p73L, p63 与 p51, 有关研究仍在进行中. 通过以上研究, 也许会对肿瘤生物学有新的认识, 并对肿瘤的基因治疗提供新的思路.