

## 云南汉族人群 *CD6* 基因多态性与非小细胞肺癌的相关性<sup>\*</sup>

洪超, 周紫云, 张新文, 刘舒媛, 陈珺, 李传印, 熊增平, 吴克川, 李珣,  
梁燕<sup>\*\*</sup>

(中国医学科学院 & 北京协和医学院医学生物学研究所, 云南省重大传染病疫苗研发重点实验室, 云南昆明 650118)

**[摘要]** 目的: 探究云南地区汉族人群中细胞表面糖蛋白 *CD6* 基因多态性与非小细胞肺癌(NSCLC)的相关性。方法: 选取云南汉族 NSCLC 患者 467 例作为病例组, 选取同期健康正常人群 490 例作为对照组, 采用 Taq-Man 探针基因分型方法对 *CD6* 基因的 SNP 位点(rs12360861、rs11230563 及 rs17824933)进行基因分型, 分析 3 个 SNP 位点等位基因和基因型在病例组和对照组间的分布差异, 分析 3 个 SNP 位点在肺癌不同临床分期的中分布。结果: *CD6* 基因的 SNP 位点(rs12360861、rs11230563 及 rs17824933)的等位基因、基因型和构建的单倍型在病例组与对照组中的分布频率比较, 差异无统计学意义( $P > 0.05$ ); *CD6* 基因的 3 个 SNP 位点的等位基因、基因型在肺癌 I + II 期组和 III + IV 期组间的分布频率比较, 差异无统计学意义( $P > 0.05$ )。结论: *CD6* 基因的 rs12360861、rs11230563 及 rs17824933 多态性位点可能与云南汉族人群 NSCLC 的发生发展无相关性。

**[关键词]** 肺肿瘤; 多态性, 单核苷酸; 云南; 汉族; 细胞表面糖蛋白 *CD6* 基因

**[中图分类号]** R734.2; RZ274 **[文献标识码]** A **[文章编号]** 1000-2707(2018)03-0254-05

**DOI:**10.19367/j.cnki.1000-2707.2018.03.002

## Association between *CD6* Gene Polymorphisms and Non-small-cell Lung Cancer of Han Population in Yunnan Province

HONG Chao, ZHOU Ziyun, ZHANG Xinwen, LIU Shuyuan, CHEN Jun,

LI Chuanyin, XIONG Zengping, WU Kechuan, LI Yu, LIANG Yan

(Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Yunnan key Laboratory of Vaccine Development for Major Infectious Diseases, Kunming 650118, Yunnan, China)

**[Abstract]** **Objective:** To investigate the association between *CD6* gene polymorphisms and non-small cell lung cancer (NSCLC) of Han population in Yunnan Province. **Methods:** 467 NSCLC patients were selected as case group, and 490 healthy persons were selected as control group. TaqMan probe genotyping was applied to detect genotype of the three single nucleotide polymorphisms (SNPs), which were rs12360861, rs11230563 and rs17824933 of *CD6* gene. Then the distributional differences of alleles and genotypes of the three SNPs in case group and control group were analyzed, as well as their distributions in different clinical stages of lung cancer. **Results:** Differences of the distribution frequencies of alleles, genotypes and constructed haplotype of the three SNPs (rs12360861, rs11230563 and rs17824933) in the case group and the control group were not statistically significant ( $P > 0.05$ ). Similarly, there was no significant difference of the SNP distribution between different pathologic stages of NSCLC (I + II VS III + IV) ( $P > 0.05$ ). **Conclusion:** The rs12360861, rs11230563 and rs17824933 in the *CD6* gene might not be associated with NSCLC in the Yunnan Han population.

<sup>\*</sup>[基金项目] 国家自然科学基金项目(81573206); 云南省应用基础研究重点项目(2016FA034)

<sup>\*\*</sup>通信作者 E-mail: liangyan@imbcams.com.cn

网络出版时间: 2018-03-20 网络出版地址: <http://kns.cnki.net/kcms/detail/52.1164.R.20180320.1721.015.html>

[Key words] lung cancer; polymorphisms, single nucleotide; Yunnan; Han population; cell surface glycoprotein CD6 gene

肺癌是世界范围内发病率最高的恶性肿瘤,在我国已成为发病率及死亡率最高的恶性肿瘤<sup>[1]</sup>。肺癌可分为非小细胞肺癌(NSCLC)和小细胞肺癌(SCLC)2种类型,其中约85%为NSCLC。研究证实肺癌的发生不仅与环境因素高度相关,而且与患者遗传因素关系密切<sup>[2]</sup>,尤其是宿主基因中存在的遗传多态性在肺癌的发生发展过程中发挥着重要作用<sup>[3-4]</sup>。CD6基因位于人类染色体11q13,编码分子量为100~130 kDa的细胞表面蛋白CD6,CD6属于SRCRSF成员,主要表达于外周血T细胞和髓质的胸腺细胞表面,也表达于部分NK细胞、B细胞及脑细胞表面<sup>[5-7]</sup>。CD6基因在未成熟CD4<sup>+</sup>和CD8<sup>+</sup>T细胞表面的表达具有显著差异,提示CD6可能与早期细胞免疫应答相关<sup>[8]</sup>。研究还发现,CD6基因中的多态性位点与多种疾病具有相关性<sup>[9-11]</sup>,而目前关于CD6基因多态性与NSCLC的研究报道较少,因此,本研究以云南汉族NSCLC患者作为研究对象,选取CD6基因的3个单核苷酸多态性(single nucleotide polymorphism, SNP)位点rs12360861、rs11230563及rs17824933进行分析,探讨其与NSCLC的相关性。

## 1 对象与方法

### 1.1 对象

随机选取2012年10月~2014年5月确诊为NSCLC患者467例作为病例组,根据世界卫生组织NSCLC组织类型划分标准进行病理分型,包括鳞癌、腺癌、鳞腺癌及其他;根据国际NSCLC临床分期系统进行临床分期,包括I、II、III及IV期;排除手术前接受放疗及化疗等抗肿瘤治疗患者,排除患有其他恶性肿瘤、合并心血管疾病、糖尿病、肝炎、肾病等疾病的患者及资料不全者;选取同期的正常体检者490例作为对照组。所有被检者均来自云南地区彼此无亲缘关系的汉族人群,年龄30~75岁。

### 1.2 方法

采集受试者空腹静脉血5 mL,用EDTA·K<sub>2</sub>或肝素抗凝,使用全血基因组DNA提取试剂盒(QIAamp DNA Blood Mini Kit, QIAGEN公司,货号51106)提取DNA,并对所提取的DNA进行浓度和纯度测定。采用TaqMan探针荧光定量PCR方法

对CD6基因rs12360861、rs11230563和rs17824933位点进行基因分型,TaqMan探针、引物(Taqman assays)和SNPs分型试剂(TaqMan Genotyping Master Mix)均购自美国ABI公司。Taqman assay ID分别为C\_\_25922320\_10、C\_\_31727142\_10和C\_\_33967506\_10,用Applied Biosystems Quant Studio6(美国Applied Biosystems公司)实时荧光定量PCR仪检测SNP位点基因型。PCR反应体积为20  $\mu$ L,反应条件为95  $^{\circ}$ C预变性10 min,92  $^{\circ}$ C变性15 s,60  $^{\circ}$ C退火1 min,共40个循环,40  $^{\circ}$ C长延伸5 min。用去离子水代替样本模板DNA作为阴性对照,分型工作完成后挑选每个位点不同基因型样本进行测序,验证Taqman基因分型的准确性。

### 1.3 统计学分析

各组间的年龄差异用Student *t*检验进行检测,各组间性别分布差异用 $\chi^2$ 检验,用Hardy-Weinberg平衡检验样品的代表性。CD6基因中SNP位点的等位基因、基因型及构建的单倍型在各SNP位点中分布频率的差异用 $\chi^2$ 检验进行比较, $P < 0.05$ 为差异有统计学意义。用在线软件Shesis分析3个位点间的连锁不平衡关系<sup>[12]</sup>,根据连锁不平衡结果构建CD6基因的3个SNP位点的单倍型,并分析所构建的单倍型在各组中的分布差异,多重比较时采用Bonferroni校正。

## 2 结果

### 2.1 基本信息

对照组共纳入490个健康个体。病例组共有肺癌患者467例,其中鳞癌患者165例、腺癌患者276例、鳞腺癌患者8例以及其他病理类型患者18例;临床分期为I期+II期151例,III期+IV期316例。病例组和对照组人群的平均年龄差异及性别分布差异均无统计学意义( $P > 0.05$ )。本研究中的3个SNPs位点的基因型在病例组和对照组中的分布均符合Hardy-Weinberg平衡( $P > 0.05$ ),表明本研究所选人群是具有代表性的。

### 2.2 CD6基因3个SNPs与NSCLC的相关性

CD6基因的rs12360861、rs11230563及rs17824933位点等位基因、等位基因型分布频率比较,差异无统计学意义( $P > 0.05$ )。见表1。

表 1 *CD6* 基因 3 个 SNP 位点等位基因及基因型在 2 组被检者中的频率分布

Tab. 1 The allelic and genotypic frequencies of three SNPs between case group and control group

| SNP 位点     | 等位基因 - 基因型 | 癌症组( <i>n</i> ,%) | 对照组( <i>n</i> ,%) | <i>P</i> | <i>OR</i> (95% <i>CI</i> ) |
|------------|------------|-------------------|-------------------|----------|----------------------------|
| rs12360861 | A          | 1 (0.1)           | 2 (0.2)           | 0.167    | 0.262(0.029 ~ 2.344)       |
|            | G          | 933(99.9)         | 978(99.8)         |          |                            |
|            | A/A        | 0 (0.0)           | 0 (0.0)           | 0.591    |                            |
|            | A/G        | 1 (0.2)           | 2 (0.4)           |          |                            |
|            | G/G        | 466(99.8)         | 488(99.6)         |          |                            |
| rs11230563 | C          | 776(83.1)         | 822(83.9)         | 0.640    | 0.940(0.74 ~ 1.20)         |
|            | T          | 158(16.9)         | 158(16.1)         |          |                            |
|            | C/C        | 328(70.2)         | 349(71.2)         | 0.871    |                            |
|            | C/T        | 120(25.7)         | 124(25.3)         |          |                            |
|            | T/T        | 19 (4.1)          | 17 (3.5)          |          |                            |
| rs17824933 | C          | 912(97.6)         | 950(96.9)         | 0.342    | 1.31(0.75 ~ 2.29)          |
|            | G          | 22 (2.4)          | 30 (3.1)          |          |                            |
|            | C/C        | 445(95.3)         | 460(93.9)         | 0.336    |                            |
|            | C/G        | 22 (4.7)          | 30 (6.1)          |          |                            |
|            | G/G        | 0 (0.0)           | 0 (0.0)           |          |                            |

2.3 *CD6* 基因 3 个 SNP 位点与 NSCLC 临床分期的相关性

本研究进一步对 *CD6* 基因中的 3 个 SNP 位点等位基因和基因型在肺癌 I + II 期患者和 III + IV 期患者的分布差异进行了分析。结果显示,*CD6* 基

因中的 3 个 SNPs 位点的等位基因和基因型在肺癌 I + II 期患者和 III + IV 期患者中的分布频率比较,差异无统计学意义( $P > 0.05$ ),见表 2。提示 *CD6* 基因中的 3 个 SNP 位点可能与肺癌的严重程度无相关性。

表 2 *CD6* 基因 3 个 SNPs 位点的等位基因及基因型在各临床分期肺癌患者中的分布频率

Tab. 2 The allelic and genotypic frequencies of the three SNPs in I + II and III + IV groups

| SNP 位点     | 等位基因 - 基因型 | I + II 期( <i>n</i> ,%) | III + IV 期( <i>n</i> ,%) | <i>P</i> | <i>OR</i> (95% <i>CI</i> ) |
|------------|------------|------------------------|--------------------------|----------|----------------------------|
| rs12360861 | A          | 1 (0.3)                | 0 (0.0)                  | 0.148    | -                          |
|            | G          | 301(100.0)             | 632(100.0)               |          |                            |
|            | A/A        | 0 (0.0)                | 0 (0.0)                  | 0.148    |                            |
|            | A/G        | 0 (0.0)                | 0 (0.0)                  |          |                            |
|            | G/G        | 151(100.0)             | 316(100.0)               |          |                            |
| rs11230563 | C          | 245 (81.1)             | 531 (84.0)               | 0.263    | 0.83(0.59 ~ 1.15)          |
|            | T          | 57 (18.9)              | 101 (16.0)               |          |                            |
|            | C/C        | 100 (66.2)             | 228 (72.2)               | 0.372    |                            |
|            | C/T        | 45 (29.8)              | 75 (23.7)                |          |                            |
|            | T/T        | 6 (4.0)                | 13 (4.1)                 |          |                            |
| rs17824933 | C          | 293 (97.0)             | 619 (97.9)               | 0.943    | 1.03(0.48 ~ 2.19)          |
|            | G          | 9 (3.0)                | 13 (2.1)                 |          |                            |
|            | C/C        | 142 (94.0)             | 303 (95.9)               | 0.378    |                            |
|            | C/G        | 9 (6.0)                | 13 (4.1)                 |          |                            |
|            | G/G        | 0 (0.0)                | 0 (0.0)                  |          |                            |

注:“-”为未做

2.4 CD6 基因 3 个 SNP 构建的单倍型与 NSCLC 的相关性

CD6 基因 SNP 位点 rs12360861、rs11230563 及 rs17824933 的连锁不平衡分析结果显示, rs12360861、rs11230563 和 rs17824933 位点的  $D' > 0.9$ ,提示 3 个 SNP 位点在病例组和对照组中存

在强连锁。根据连锁不平衡结果构建 3 个 SNP 位点的单倍型,并对在病例组和对照组中分布频率  $> 3\%$  的单倍型进行分析,结果显示单倍型 G - C - C、G - C - G 和 G - T - C 在病例组和对照组中分布频率比较,差异无统计学意义( $P > 0.05$ ),见表 3。

表 3 CD6 基因中的 3 个 SNP 位点构建的单倍型在 2 组被检者中的分布

Tab. 3 The haplotype analysis of the three SNPs in CD6 gene between case group and control group

| rs12360861 | rs11230563 | rs17824933 | 病例组(n,%)      | 对照组(n,%)      | P     | 校正 P     | OR(95% CI)           |
|------------|------------|------------|---------------|---------------|-------|----------|----------------------|
| G          | C          | C          | 263.94(0.800) | 792.07(0.808) | 0.755 | $> 0.05$ | 0.950(0.694 ~ 1.303) |
| G          | C          | G          | 10.06(0.030)  | 29.93(0.031)  | 0.998 | $> 0.05$ | 0.999(0.484 ~ 2.064) |
| G          | T          | C          | 55.06(0.167)  | 155.92(0.159) | 0.737 | $> 0.05$ | 1.059(0.757 ~ 1.483) |

3 讨论

肿瘤是危害人类健康的重大疾病之一,在我国,肺癌的发病率及死亡率已居所有的恶性肿瘤之首。这主要是因为肺癌检出多为晚期,患者预后不良,因此寻找与肺癌发生相关的分子标记对于肺癌的早期诊断及治疗非常重要。CD6 是一种表达于细胞表面的 I 型糖蛋白,包含 3 个清道夫受体半胱氨酸富集(SRCR)结构域和活化的白细胞黏附分子(ALCAM)的结合位点。表达于所有 T 细胞,天然杀伤细胞和 B1a 细胞的子集以及其他的一些免疫细胞表面<sup>[13]</sup>。CD6 分子与其配体 ALCAM 的相互作用在淋巴细胞增殖和激活过程中发挥重要作用<sup>[14]</sup>。研究发现,在抗原提呈细胞刺激后 CD6<sup>+</sup>T 细胞的增殖明显强于 CD6<sup>-</sup>T 细胞<sup>[15]</sup>,这说明 CD6T 细胞可能在 T 细胞免疫应答过程中发挥重要作用。研究表明,CD6 基因多态性可能会影响 CD6 的表达<sup>[16]</sup>,从而影响其功能,改变机体免疫应答水平。Wagner 等<sup>[16]</sup>研究发现 CD6 基因的 rs12360861 位点与多发性硬化的发生存在相关性,这可能是由于该位点位于外显子剪接增强子的顺式作用元件中<sup>[17]</sup>,该位点碱基的改变可能会导致 mRNA 剪接异常从而影响 CD6 基因的功能<sup>[18]</sup>。然而本研究结果显示 rs12360861 可能与云南汉族人群 NSCLC 无相关性( $P > 0.05$ ),这与本课题组前期研究的该位点与云南汉族人群 HCV 慢性感染相关性结果一致<sup>[19]</sup>。2017 年,Consuegra 等<sup>[20]</sup>的研究显示 rs11230563 等位基因和基因型均与欧洲人群银屑病相关,此外在中国汉族人群中的研究显示 rs11230563 位点与白塞氏具有相关性<sup>[11]</sup>。而本研

究结果显示该位点与云南汉族人群 NSCLC 的发生无相关性,这可能是因为不同疾病的发病机制不同,所以相同的 SNP 位点在不同疾病中发挥的作用可能不同。Kofler 等<sup>[21]</sup>研究发现多发性硬化风险相关位点 rs17824933 与 CD4<sup>+</sup>T 细胞的增殖相关,随后有研究进一步证实了该位点可能与西班牙人群多发性硬化相关<sup>[8]</sup>。而本研究结果显示 rs17824933 可能与云南汉族人群 NSCLC 的发生无相关性。另外本研究还分析了 CD6 基因中的这 3 个 SNP 位点与 NSCLC 临床分期的相关性,结果显示该 3 个 SNP 位点可能与云南汉族 NSCLC 临床分期无相关性,3 个位点构建的单倍型也可能与云南汉族 NSCLC 的发生无相关性。

综上所述,CD6 分子可能通过免疫 T 细胞的增殖和活化在免疫应答过程发挥作用<sup>[22]</sup>,本研究选择了 CD6 基因中的 3 个 SNP 位点 rs12360861、rs11230563 及 rs17824933 探讨这些位点与云南汉族人群 NSCLC 及其临床分期的相关性,结果显示 CD6 基因中的 3 个 SNP 位点可能与云南汉族 NSCLC 及其临床分期无相关性。要进一步确定 CD6 基因中的多态性位点在 NSCLC 发生发展过程中发挥的作用,需要增加样本量以及多态性位点数量,综合不同人群的研究数据进行分析,评估 CD6 基因中的多态性位点是否有潜力成为 NSCLC 的诊断及治疗靶标。

4 参考文献

[1] 陈万青, 郑荣寿, 张思维, 等. 2013 年中国恶性肿瘤发病和死亡分析[J]. 中国肿瘤, 2017, 26 (1): 1 - 7.  
[2] LIND H, ZIENOLDDINY S, RYBERG D, et al. Inter-

- leukin 1 receptor antagonist gene polymorphism and risk of lung cancer: a possible interaction with polymorphisms in the interleukin 1 beta gene[J]. *Lung Cancer* (Amsterdam, Netherlands), 2005, 50(3): 285 – 290.
- [3] HOSGOOD H D, SONG M, HSIUNG C A, et al. Interactions between household air pollution and GWAS-identified lung cancer susceptibility markers in the Female Lung Cancer Consortium in Asia (FLCCA)[J]. *Human Genetics*, 2015, 134(3): 333 – 341.
- [4] GOVINDAN R, DING L, GRIFFITH M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers[J]. *Cell*, 2012, 150(6): 1121 – 1134.
- [5] BRAUN M, MULLER B, TER MEER D, et al. The CD6 scavenger receptor is differentially expressed on a CD56 natural killer cell subpopulation and contributes to natural killer-derived cytokine and chemokine secretion [J]. *Journal of Innate Immunity*, 2011, 3(4): 420 – 434.
- [6] BODIAN D L, SKONIER J E, BOWEN M A, et al. Identification of residues in CD6 which are critical for ligand binding[J]. *Biochemistry*, 1997, 36(9): 2637 – 2641.
- [7] CHAPPELL P E, GARNER L I, YAN J, et al. Structures of CD6 and Its Ligand CD166 Give Insight into Their Interaction[J]. *Structure*, 2015, 23(8): 1426 – 1436.
- [8] SWAMINATHAN B, CUAPIO A, ALLOZA I, et al. Fine mapping and functional analysis of the multiple sclerosis risk gene CD6[J]. *PloS One*, 2013, 8(4): e62376.
- [9] PARK T J, KIM H J, KIM J H, et al. Associations of CD6, TNFRSF1A and IRF8 polymorphisms with risk of inflammatory demyelinating diseases[J]. *Neuropathology and Applied Neurobiology*, 2013, 39(5): 519 – 530.
- [10] LEPPA V S I, TIENARI P J, ELOVAARA I, et al. The genetic association of variants in CD6, TNFRSF1A and IRF8 to multiple sclerosis: a multicenter case-control study[J]. *PloS One*, 2011, 6(4): e18813.
- [11] ZHENG M, ZHANG L, YU H, et al. Genetic polymorphisms of cell adhesion molecules in Behcet's disease in a Chinese Han population[J]. *Scientific Reports*, 2016, 6:24974.
- [12] SHI Y Y, HE L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci [J]. *Cell Research*, 2005, 15(2): 97 – 98.
- [13] SARUKHAN A, MARTINEZ-FLORENSA M, ESCODAFERRAN C, et al. Pattern Recognition by CD6: A Scavenger-Like Lymphocyte Receptor[J]. *Current Drug Targets*, 2016, 17(6): 640 – 650.
- [14] IBANEZ A, SARRIAS M R, FARNOS M, et al. Mitogen-activated protein kinase pathway activation by the CD6 lymphocyte surface receptor[J]. *Journal of Immunology*, 2006, 177(2): 1152 – 1159.
- [15] RASMUSSEN R A, COUNTS S L, DALEY J F, et al. Isolation and characterization of CD6-T cells from peripheral blood[J]. *Journal of Immunology* (Baltimore, Md : 1950), 1994, 152(2): 527 – 536.
- [16] WAGNER M, BILINSKA M, POKRYSZKO – DRAGAN A, et al. ALCAM and CD6-multiple sclerosis risk factors [J]. *Journal of Neuroimmunology*, 2014, 276(1 – 2): 98 – 103.
- [17] DESMET F O, HAMROUN D, LALANDE M, et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals[J]. *Nucleic Acids Research*, 2009, 37(9): e67.
- [18] BARALLE D, BARALLE M. Splicing in action: assessing disease causing sequence changes [J]. *Journal of Medical Genetics*, 2005, 42(10): 737 – 748.
- [19] 刘城秀, 沈云松, 李桃意, 等. 云南汉族人群 CD6 基因多态性与 HCV 慢性感染的关系[J]. *贵阳医学院学报*, 2016, 41(5): 527 – 531.
- [20] CONSUEGRA-FERNANDEZ M, JULIA M, MARTINEZ-FLORENSA M, et al. Genetic and experimental evidence for the involvement of the CD6 lymphocyte receptor in psoriasis[J]. *Cellular & Molecular Immunology*, 2017, doi: 10.1038/cmi.2017.119.
- [21] KOFLER D M, SEVERSON C A, MOUSSISSIAN N, et al. The CD6 multiple sclerosis susceptibility allele is associated with alterations in CD4 + T cell proliferation [J]. *Journal of immunology*, 2011, 187(6): 3286 – 3291.
- [22] ORTA-MASCARO M, CONSUEGRA-FERNANDEZ M, CARRERAS E. CD6 modulates thymocyte selection and peripheral T cell homeostasis[J]. *J Exp Med*, 2016, 213(8): 1387 – 1397.

(2018-01-02 收稿, 2018-03-01 修回)  
中文编辑: 吴昌学; 英文编辑: 雷 妍