

CD133 基因 3-UTR 区域变异位点与云南汉族人群非小细胞肺癌的相关性*

洪超¹, 赵雨笛¹, 谭芳², 李盈甫², 马千里³, 刘楠楠¹, 刘舒媛¹, 史荔¹, 李传印^{1**}

(1. 中国医学科学院 & 北京协和医学院医学生物学研究所, 云南 昆明 650118; 2. 昆明医科大学第一附属医院, 云南 昆明 650032; 3. 昆明医科大学第三附属医院, 云南 昆明 650118)

[摘要] 目的: 研究 CD133 基因 3-UTR 区域变异位点(rs2240688 和 rs3130)单核苷酸多态性位点(SNPs)与云南汉族人群非小细胞肺癌的相关性。方法: 选取 420 例云南汉族非小细胞肺癌患者为病例组, 588 例云南汉族健康个体作为对照组; 采用 TaqMan 探针基因分型方法对 2 组受试者 CD133 基因中 rs2240688 和 rs3130 位点进行基因分型, 并采用 χ^2 检验分析上述两 SNPs 位点等位基因、基因型及构建的单倍型分布频率在病例组与对照组中的差异。结果: 病例组与对照组 rs2240688 位点等位基因和基因型分布频率比较差异有统计学意义($P = 0.013, 0.033$), 而两组 rs3130 等位基因和基因型分布频率比较差异无统计学意义($P > 0.05$), rs2240688 位点等位基因 G 是云南汉族非小细胞肺癌的风险性因素($OR = 1.301, 95\% CI$ 为 $1.056 \sim 1.601$), rs2240688 及 rs3130 间存在强连锁($D' > 0.8$); 构建的两个 SNPs 位点的单倍型中 rs2240688G-rs3130C 与肺癌风险升高有关($OR = 1.238, 95\% CI$ 为 $1.003 \sim 1.528$), 而单倍型 rs2240688T-rs3130C 可能是肺癌发生的保护性因素($OR = 0.828, 95\% CI$ 为 $0.690 \sim 0.994$)。结论: CD133 基因 3-UTR 区域的 SNP 位点 rs2240688 与云南汉族人群非小细胞肺癌相关, 等位基因 G 可能是云南汉族人群非小细胞肺癌的风险性因素。

[关键词] 癌, 非小细胞肺; 基因; CD133; 3-UTR; 多态性, 单核苷酸; 云南; 汉族

[中图分类号] R734.2 **[文献标识码]** A **[文章编号]** 1000-2707(2019)01-0026-05

DOI: 10.19367/j.cnki.1000-2707.2019.01.005

The Association of Single Nucleotide Polymorphisms in 3-UTR of CD133 Gene with Non-small Cell Lung Cancer in A Chinese Han Population in Yunnan Province

HONG Chao¹, ZHAO Yudi¹, TAN Fang², LI Yingfu², MA Qianli³,
LIU Nannan¹, LIU Shuyuan¹, SHI Li¹, LI Chuanyin¹

(1. Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming 650118, Yunnan, China; 2. The 1st Affiliated Hospital of Kunming Medical University, Kunming 650032, Yunnan, China; 3. The 3rd Affiliated Hospital of Kunming Medical University, Kunming 650118, Yunnan, China)

[Abstract] **Objective:** To study the association of single nucleotide polymorphisms (SNPs, rs2240688 and rs3130) of CD133 gene in 3-UTR region mutation sites with non-small cell lung cancer in Yunnan Han population. **Methods:** A total of 420 patients with non-small cell lung cancer in Yunnan Han population were enrolled as the case group, and 588 healthy individuals from Yunnan Han population as the control group. The TaqMan probe genotyping method was performed to detect

*[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); 中国医学科学院重大协同创新项目 2016-12M-2-001; 云南省应用基础研究计划-昆明医科大学联合专项课题 2017FE468(-193)、2018FE001(-219)

**通信作者 E-mail: chuanyinli@imbcams.com.cn

网络出版时间: 2019-01-18 网络出版地址: <http://kns.cnki.net/kcms/detail/52.1164.R.20190118.1118.005.html>

rs2240688 and rs3130 in the CD133 gene of the two groups. The χ^2 test was used to analyze the difference between the two SNPs loci alleles, genotypes and haplotype distribution frequencies in the case group and the control group. **Results:** There were significant differences in the frequencies of alleles and genotypes between the case group and the control group at rs2240688 ($P = 0.013, 0.033$), whereas the differences in the frequency of rs3130 alleles and genotypes between the two groups were not significant ($P > 0.05$). Allele G in rs2240688 was a risk factor for non-small cell lung cancer in Han population in Yunnan ($OR = 1.301, 95\% CI 1.056 \sim 1.601$). There is a strong linkage between rs2240688 and rs3130 ($D' > 0.8$). rs2240688G-rs3130C is associated with increased risk of lung cancer ($OR = 1.238, 95\% CI 1.003 \text{ to } 1.528$), whereas haplotype rs2240688T-rs3130C may be protective for lung cancer. Sexual factors ($OR = 0.828, 95\% CI 0.690 \sim 0.994$). Conclusion: The rs2240688 in the 3-UTR region of CD133 gene is associated with non-small cell lung cancer in Yunnan Han population. Allele G may be a risk factor for non-small cell lung cancer in Yunnan Han population.

[**Key words**] cancer, non-small cell lung; gene; CD133; 3-UTR; polymorphisms, single nucleotide; Yunnan; Han population

肺癌是现在最主要的导致死亡的恶性肿瘤,据统计,全球每年约有 180 万新发病例及 160 万死亡病例^[1],这其中近 30% 发生在中国^[2]。虽然吸烟是肺癌发生的主要风险因素,但研究发现,过去十年肺癌患者中非吸烟人群的比例正逐渐升高^[3],提示除吸烟外其他因素在肺癌发生过程中也发挥重要作用,这其中比较重要的是遗传因素^[4-5]。CD133 分子(又被称为 prominin-1)由 *PROM1* 基因表达,是跨膜糖蛋白家族成员之一,具有 5 次跨膜结构,分子量约为 120 kD^[6]。CD133 分子最早被发现表达于造血干细胞表面^[7]。近年来研究发现,CD133 分子可能是多种肿瘤(肺癌、结直肠癌、肝癌等)干细胞的标志分子^[8-10],并且与细胞的生长、增殖、分化、迁移等过程密切相关^[11-13]。在肺癌的研究中发现,CD133 分子的表达水平与肺癌的发展呈正相关,且与肺癌预后不良相关^[14]。位于 CD133 基因 3-UTR 区域的单核苷酸多态性(single nucleotide polymorphisms, SNPs)位点可能会通过影响与 microRNA 的相互作用而影响基因的表达调控,从而与疾病的发生发展具有相关性^[15-16]。因此,本研究选取了表达水平与肺癌发生发展具有相关性的 CD133 基因 3-UTR 区域中的两个 SNPs 位点(rs2240688 和 rs3130),研究其与云南汉族人群非小细胞肺癌的相关性。

1 材料与方法

1.1 样本来源

本研究选取 2014 年 5 月~2017 年 10 月在昆

明医科大学第一附属医院被确诊为非小细胞肺癌患者 420 例作为病例组,选取同期健康体检的正常健康个体 588 例作为对照组。排除接收过放化疗等肿瘤治疗的患者,排除其他恶性肿瘤及慢性疾病的患者。所有参加者均为居住在云南地区,彼此无血缘关系的汉族个体。

1.2 方法

1.2.1 样本基因组 DNA 提取 采用 QIAGEN 全血基因组 DNA 提取试剂盒(QIAamp DNA Blood Mini Kit,货号 51106)提取研究对象基因组 DNA,测定 DNA 的浓度和纯度后置于 $-20\text{ }^{\circ}\text{C}$ 储存备用。

1.2.2 SNPs 位点的基因分型 SNPs 位点的基因分型采用 TaqMan 探针基因分型方法,CD133 基因中的两个 SNPs 位点 rs2240688 和 rs3130 基因分型试剂盒购于美国 ABI(Applied Biosystems)公司,编号分别为 C_____89927_1_和 C_____89925_20。TaqMan 探针荧光定量 PCR 使用美国 ABI 公司的 Quant Studio 6 实时荧光定量 PCR 仪,PCR 反应体系为 $20\text{ }\mu\text{L}$,反应程序为 $95\text{ }^{\circ}\text{C}$ 预变性 10 min, $92\text{ }^{\circ}\text{C}$ 变性 15 s, $60\text{ }^{\circ}\text{C}$ 退火 1 min(40 个循环), $40\text{ }^{\circ}\text{C}$ 最后延伸 5 min,PCR 过程设阴性对照(以水代替 DNA 模板)。

1.3 统计学分析

采用 χ^2 检验分析病例组和对照组样本性别差异,独立样本 t 检验分析病例组和对照组年龄差异,Hardy-Weinberg 平衡检验分析本研究纳入样本的群体代表性;使用在线软件 SHEsis^[17-18] 分析 CD133 基因 3-UTR 区域的 rs2240688 和 rs3130 位

点的连锁关系,并根据连锁不平衡结果构建两个 SNPs 位点的单倍型,两个 SNPs 位点与非小细胞肺癌的相关性分析采用 χ^2 检验; $P < 0.05$ 说明不同组间的数据比较差异具有统计学意义。

2 结果

2.1 研究样本基本信息及群体代表性检验

本研究共纳入非小细胞肺癌患者 420 例,正常健康个体 588 例;病例组平均(55.75 ± 10.57)岁,男 289 例、女 131 例;对照组平均(55.41 ± 8.13)岁,男 399 例、女 189 例。病例组和对照组间年龄、性别比较,差异无统计学意义($P > 0.05$)。*Hardy-Weinberg* 平衡检验分析结果显示, *CD133* 基因 3-UTR 区域的两个 SNPs 位点的基因型在病例组(P

=0.28, $P = 0.95$)、对照组($P = 0.65, P = 0.49$)均符合 *Hardy-Weinberg* 平衡定律($P > 0.05$)表明本研究所选取的样本具有群体代表性。

2.2 rs2240688 及 rs3130 位点与非小细胞肺癌的相关性分析

CD133 基因 3-UTR 区域的 rs2240688 及 rs3130 等位基因和基因型在非小细胞肺癌病例组和对照组中的分布特征见表 1。结果显示,rs2240688 位点等位基因和基因型分布频率在病例组和对照组比较,差异有统计学意义($P = 0.013$ 、0.033);而 rs3130 等位基因和基因型分布频率在病例组和对照组比较,差异无统计学意义($P > 0.05$)。进一步分析的结果显示,rs2240688 位点等位基因 G 是云南汉族非小细胞肺癌的风险性因素($OR = 1.301, 95\% CI$ 为 1.056 ~ 1.601)。

表 1 两组被检者 *CD133* 基因 rs2240688 和 rs3130 位点等位基因、基因型频率分布
Tab.1 The frequency distributions of alleles and genotypes of rs22406882 and rs3130 in the 3-UTR of *CD133* gene between case and control groups

SNPs 位点	等位基因/基因型	病例组(<i>n</i> ,%)	对照组(<i>n</i> ,%)	<i>P</i>	<i>OR</i> (95% <i>CI</i>)
rs2240688	G	219 (26.1)	250 (21.3)	0.013	1.301 (1.056 ~ 1.601)
	T	621 (73.9)	922 (78.7)		
	G/G	29 (6.9)	23 (3.9)		
	G/T	161 (38.3)	204 (34.8)		
	T/T	230 (54.8)	359 (61.3)		
rs3130	C	519 (61.9)	744 (63.4)	0.510	0.940 (0.783 ~ 1.129)
	T	319 (38.1)	430 (36.6)		
	C/C	161 (38.4)	238 (40.5)		
	C/T	197 (47.0)	268 (45.7)		
	T/T	61 (14.6)	81 (13.8)		

2.3 rs2240688 和 rs3130 位点的连锁不平衡分析及单倍型构建

结果显示,rs2240688 及 rs3130 SNPs 位点间存在强连锁($D' > 0.8$)。构建两个 SNPs 位点的单倍型,并分析分布频率 >3% 的单倍型在病例组和对照组中分布频率差异,结果显示两组单倍型 rs2240688G-

rs3130C 及 rs2240688T-rs3130C 分布频率比较,差异有统计学意义($P = 0.046, 0.043$)。进一步分析结果显示,单倍型 rs2240688G-rs3130C 与肺癌风险升高有关($OR = 1.238, 95\% CI$ 为 1.003 ~ 1.528),而单倍型 rs2240688T-rs3130C 可能是肺癌发生的保护性因素($OR = 0.828, 95\% CI$ 为 0.690 ~ 0.994)。见表 2。

表 2 两组被检者 *CD133* 基因 rs22406882 和 rs3130 位点单倍型分布频率
Tab.2 The haplotype distributions of rs22406882 and rs3130 in the 3-UTR of *CD133* gene

rs2240688	rs3130	病例组(<i>n</i> ,%)	对照组(<i>n</i> ,%)	<i>P</i>	<i>OR</i> (95% <i>CI</i>)
G	C	208.55 (24.9)	249.95 (21.4)	0.046	1.238 (1.003 ~ 1.528)
T	C	310.45 (37.0)	491.05 (42.0)	0.043	0.828 (0.690 ~ 0.994)
T	T	309.55 (36.9)	428.95 (36.7)	0.751	1.030 (0.857 ~ 1.239)

3 讨论

肺癌是全球最常见的恶性肿瘤,发病率及死亡率均较高,非小细胞肺癌是最主要的肺癌类型,约占肺癌病例的 90%。虽然肺癌治疗手段越来越多,具有较好的治疗效果,但由于患者检出时多为晚期,治疗难度较大,且预后不良。所以寻找与肺癌发生风险相关的分子标记,早期诊断对于降低肺癌的死亡率具有重要的意义。

CD133 分子是众多肿瘤干细胞的标记分子^[19-21],并且在肿瘤患者预后判断中发挥重要作用^[22-23],在肺癌的研究中发现 *CD133* 分子的表达水平与肺癌的发展及预后具有相关性^[14]。这说明肿瘤细胞 *CD133* 分子的表达水平不仅与肺癌的发生发展具有相关性,而且还可作为肺癌预后判断的分子标记。因此,推测影响 *CD133* 基因表达的因素可能也会影响肺癌的发生发展。而且已有研究表明位于 *CD133* 基因 3-UTR 区域的 SNPs 位点(例如 rs2240688 和 rs3130)与多种人类肿瘤具有相关性。因此本研究选取了位于 *CD133* 基因中的 SNPs 位点 rs2240688 和 rs3130,研究其与云南汉族人群非小细胞肺癌发生风险的相关性。结果显示,rs2240688 位点等位基因 G 在病例组中的分布频率显著高于对照组,可能是云南汉族人群非小细胞肺癌发生的风险因素 ($OR = 1.301$; 95% CI 为 $1.056 \sim 1.601$),而 rs3130 可能与云南汉族人群非小细胞肺癌的发生风险无关。该结果与 Liu 等和 Mei 等^[24-25]在肺癌中的研究结果一致。因此,位于 *CD133* 基因 3-UTR 区域的 SNPs 位点 rs2240688 可能是通过影响 microRNA 的靶向结合来影响 *CD133* 基因的表达,从而与云南汉族人群非小细胞肺癌的发生风险具有相关性。

另外,本研究构建了 *CD133* 基因 3-UTR 区域 2 个 SNP 位点的单倍型,并分析了构建的单倍型与云南汉族人群非小细胞肺癌的相关性。结果显示,单倍型 rs2240688G-rs3130C 可能与云南汉族人群非小细胞肺癌发生风险升高相关,相反,单倍型 rs2240688T-rs3130C 可能会使云南汉族人群非小细胞肺癌发生风险降低。这也与等位基因分析结果 rs2240688 等位基因 G 是云南汉族人群非小细胞肺癌的风险性因素相符。

综上所述,本研究发现 rs2240688 与云南汉族人群非小细胞肺癌的发生风险相关,其等位基因 G

可能是云南汉族人群非小细胞肺癌发生的风险因素。同时,本研究还发现,本次研究的结果与其他在中国人群的研究结果相符,为 *CD133* 基因 3-UTR 区域 SNPs 位点作为非小细胞肺癌诊断标记提供了云南汉族人群的研究结果。

4 参考文献

- [1] TORRE L A, BRAY F, SIEGEL R L, et al. Global cancer statistics, 2012 [J]. CA: A Cancer Journal for Clinicians, 2015, 65(2): 87-108.
- [2] CHEN W, ZHENG R, BAAD P D, et al. Cancer statistics in China, 2015 [J]. CA: A Cancer Journal for Clinicians, 2016, 66(2): 115-132.
- [3] SWANTON C, GOVINDAN R. Clinical implications of genomic discoveries in lung cancer [J]. The New England Journal of Medicine, 2016, 374 (19): 1864-1873.
- [4] DE GROOT P, MUNDEN R F. Lung cancer epidemiology, risk factors, and prevention [J]. Radiologic clinics of North America, 2012, 50(5): 863-876.
- [5] TENG Y, DING Y, ZHANG M, et al. Genome-wide haplotype association study identifies risk genes for non-small cell lung cancer [J]. Journal of Theoretical Biology, 2018, 456:84-90.
- [6] ELENA I, GIUSEPPE P. *CD133*: to be or not to be, is this the real question [J]. American Journal of Translational Research, 2013, 5(6): 563-581.
- [7] A H Y, S M, E D Z, et al. *AC133*, a novel marker for human hematopoietic stem and progenitor cells [J]. Blood, 1997, 90(12): 5002-5012.
- [8] CHUNXIA W, JINGPING X, JIASONG G, et al. Evaluation of *CD44* and *CD133* as cancer stem cell markers for colorectal cancer [J]. Oncology Reports, 2012, 28(4): 1301-1308.
- [9] LIN S H, LIU T, MING X, et al. Regulatory role of hexosamine biosynthetic pathway on hepatic cancer stem cell marker *CD133* under low glucose conditions [J]. Scientific Reports, 2016, 6:21184.
- [10] HUANG M, ZHU H, FENG J, et al. High *CD133* expression in the nucleus and cytoplasm predicts poor prognosis in non-small cell lung cancer [J]. Disease Markers, 2015, 2015(1): 986095.
- [11] YUANYAN W, YIZHOU J, FEI Z, et al. Activation of PI3K/Akt pathway by *CD133*-p85 interaction promotes tumorigenic capacity of glioma stem cells [J]. Proceedings of the National Academy of Sciences of the United

- States of America, 2013, 110(17): 6829 – 6834.
- [12] JANG J W, SONG Y, KIM S H, et al. CD133 confers cancer stem-like cell properties by stabilizing EGFR-AKT signaling in hepatocellular carcinoma [J]. *Cancer Letters*, 2017, 389: 1 – 10.
- [13] LIU C, LI Y, XING Y, et al. The interaction between cancer stem cell marker cd133 and src protein promotes focal adhesion kinase (FAK) phosphorylation and cell migration [J]. *Journal of Biological Chemistry*, 2016, 291(30): 15540 – 50.
- [14] WANG D, WEN G M, HOU W, et al. The roles of CD133 expression in the patients with non-small cell lung cancer [J]. *Cancer Biomarkers*, 2018, 22(3): 1 – 10.
- [15] HU J L, HU X L, LU C X, et al. Variants in the 3'-untranslated region of CUL3 is associated with risk of esophageal squamous cell carcinoma [J]. *Journal of Cancer*, 2018, 9(20): 3647 – 3650.
- [16] DEVANNA P, VORST M V D, PFUNDT R, et al. Genome-wide investigation of an ID cohort reveals de novo 3' UTR variants affecting gene expression [J]. *Human Genetics*, 2018, 137(9): 717 – 721.
- [17] 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.
- [18] LI Z, ZHANG Z, HE Z, et al. A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers; update of the SHEsis (<http://analysis.bio-x.cn>) [J]. *Cell Research*, 2009, 19(4): 519 – 523.
- [19] OBRIEN CA, AARON P, STEVEN G, et al. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice [J]. *Nature*, 2007, 445(7123): 106 – 110.
- [20] ERAMO A, LOTTI F, SETTE G, et al. Identification and expansion of the tumorigenic lung cancer stem cell population [J]. *Cell Death & Differentiation*, 2008, 15(3): 504 – 514.
- [21] SINGH S K, CLARKE I D, TERASAKI M, et al. Identification of a cancer stem cell in human brain tumors [J]. *Cancer Research*, 2003, 63(18): 5821 – 5828.
- [22] HORST D, KRIEGL L, ENGEL J, et al. CD133 expression is an independent prognostic marker for low survival in colorectal cancer [J]. *British Journal of Cancer*, 2008, 99(8): 1285 – 1289.
- [23] FELIX Z, REZVAN A, BENITO C, et al. Stem cell marker CD133 affects clinical outcome in glioma patients [J]. *Clinical Cancer Research*, 2008, 14(1): 123 – 129.
- [24] LIU Q F, ZHANG Z F, HOU G J, et al. Polymorphisms of the stem cell marker gene cd133 and the risk of lung cancer in chinese population [J]. *Lung*, 2016, 194(3): 393 – 400.
- [25] MEI C, LEI Y, RONGRONG Y, et al. A microRNA-135a/b binding polymorphism in CD133 confers decreased risk and favorable prognosis of lung cancer in Chinese by reducing CD133 expression [J]. *Carcinogenesis*, 2013, 34(10): 2292 – 2299.

(2018-11-21 收稿, 2019-01-02 修回)

中文编辑: 吴昌学; 英文编辑: 张启芳