

# 新疆维吾尔自治区妇女宫颈感染人乳头瘤病毒 16 型 E4 和 L2 的遗传变异分析

程豪政<sup>1</sup>, 妥静<sup>1</sup>, 董杨柳<sup>1</sup>, 王乐<sup>1</sup>, 者湘漪<sup>2</sup>, 李洪涛<sup>2</sup>, 李冬妹<sup>1</sup>, 潘泽民<sup>1</sup>

**摘要** 目的 探讨人乳头瘤病毒 16 (HPV16) E4、L2 基因变异及氨基酸变化,分析 HPV16 的进化特征。方法 从医院收集 40 例 HPV16 感染阳性的宫颈脱落细胞样本和 40 例宫颈组织细胞样本,提取样本的病毒 DNA,对 40 例宫颈脱落细胞的 DNA 进行 HPV16 E4 和 L2 Sanger 测序及对 40 例宫颈组织细胞 DNA 高通量测序,构建 HPV16 E4 和 L2 基因系统进化树,分析 HPV16 E4 和 L2 基因的变异特征。结果 有 72 个 HPV16 E4 变异样本,10 个核苷酸变异位点(4 种错义变异和 7 种同义变异),74 个 HPV16 L2 变异样本,40 个核苷酸变异位点(23 种错义变异和 18 种同义变异)。宫颈癌(CC)中 T4177C、A4288C、A4654C 变异频率明显高于非宫颈癌,差异有统计学意义( $P < 0.05$ )。结论 ① 新疆维吾尔自治区主要流行病毒株为欧洲株,少数为亚洲株。② HPV16 L2 基因中 T4177C、A4288C、A4654C 在宫颈癌中的变异频率高于非宫颈癌,且 G4181A 与亚洲株相关。

**关键词** 人乳头瘤病毒 16; 基因变异; 宫颈癌; E4 和 L2 基因; 单核苷酸多态性

**中图分类号** R 373.9

**文献标志码** A **文章编号** 1000-1492(2023)11-1884-06  
doi:10.19405/j.cnki.issn1000-1492.2023.11.014

2023-07-18 接收

基金项目:国家自然科学基金(编号:82060518、U1503125);新疆生产建设兵团国际科技合作计划项目(编号:2019BC007)

作者单位:<sup>1</sup> 石河子大学医学院生物化学教研室,新疆地方与民族高发疾病教育部重点实验室,石河子 832002

<sup>2</sup> 石河子大学医学院人体解剖学与组织胚胎学教研室,石河子 832002

作者简介:程豪政,男,硕士研究生;

潘泽民,男,博士,教授,博士生导师,责任作者,E-mail: panteacher89@sina.com

宫颈癌(cervical cancer, CC)是妇女恶性肿瘤死亡的第四大因素<sup>[1]</sup>。据报道<sup>[2]</sup>,中国新疆维吾尔自治区宫颈癌发病率与病死率很高,该地区人乳头瘤病毒(Human Papillomavirus, HPV)阳性率为 14.02%,其中多为 HPV52、HPV53、HPV16 和 HPV18。

研究<sup>[3]</sup>表明,感染 HPV16 具有更强的致病性和致癌性。HPV16 分为 A 谱系、B 谱系、C 谱系和 D 谱系 4 种变异谱系<sup>[4]</sup>,中国新疆维吾尔自治区女性感染 HPV16 以欧洲病毒株为主<sup>[5]</sup>。HPV16 E4 基因产物 E4 蛋白质,是 HPV 表达量最高的蛋白质,其氨基酸主要来源于 E4 开放阅读框(open reading frame, ORF),且 E4 ORF 包含在 E2 ORF 内<sup>[6-7]</sup>。E4 蛋白质可增强病毒复制和病毒粒子的排出<sup>[8-10]</sup>。HPV 衣壳蛋白质在病毒进入细胞与细胞核的过程中起关键作用<sup>[11-12]</sup>。L2 促进了病毒基因组的逆行运输,促进细胞间期将病毒基因组整合到宿主基因<sup>[13-14]</sup>。新疆维吾尔自治区 HPV16 E4 和 L2 基因的变异及分布特征尚不清楚,因此,研究 HPV16 E4 和 L2 基因变异,有助于发现 HPV16 基因变异位点与宫颈癌的发生和发展的关系。

## 1 材料与方法

**1.1 样本的采集** 在伊犁州友谊医院、喀什地区人民医院和石河子大学附属医院共收集了 80 例患者的 HPV16 感染阳性的宫颈细胞样本(脱落细胞 40 例,组织细胞 40 例);样本收集取得了所有患者的知情同意。所有患者都无外地的长期旅居史,收集

GAP1 group significantly decreased (all  $P < 0.01$ ). circ-RACGAP1 could target and inhibit the expression of miR-4324 ( $P < 0.01$ ). Compared with the control group, the expression level of RACGAP1 protein in T24 cells in the sh-circ-RACGAP1 group decreased ( $P < 0.01$ ), and the expression levels of PI3K/AKT signaling pathway proteins phosphatidylinositol-3-kinase (p-PI3K), phosphorylated protein kinase B (p-AKT), nuclear factor kappa B (NF- $\kappa$ B) decreased (all  $P < 0.01$ ). **Conclusion** circ-RACGAP1 is highly expressed in bladder cancer tissues and cell lines, knocking down circ-RACGAP1 can inhibit the malignant biological behavior of T24 cells, and circ-RACGAP1 plays a role by inhibiting the expression of miR-4324 and activating the PI3K/AKT signaling pathway.

**Key words** bladder cancer; circular RNA; circ-RACGAP1; miR-4324

样本存放于  $-80\text{ }^{\circ}\text{C}$  低温冰箱。

**1.2 样本的 DNA 提取与 PCR 扩增** DNA 提取试剂盒(北京天根生化有限公司)提取 DNA, DNA 储存在  $-20\text{ }^{\circ}\text{C}$  的冰箱。混合反应液(40  $\mu\text{l}$ )包括  $2\times$  Taq PCR SuperMix 20  $\mu\text{l}$ , 引物各 1  $\mu\text{l}$  及 2  $\mu\text{l}$  样本 DNA。反应程序温度控制为  $94\text{ }^{\circ}\text{C}$ , 持续 5 min; 34 个循环( $94\text{ }^{\circ}\text{C}$  30 s,  $52\text{ }^{\circ}\text{C}$  30 s,  $72\text{ }^{\circ}\text{C}$  1 min);  $72\text{ }^{\circ}\text{C}$  5 min。存放于  $-20\text{ }^{\circ}\text{C}$  冰箱。引物的信息, 见表 1。

表 1 引物信息

引物名称	引物序列(5'-3')
P1	F: GAAGCATCAGTAACTGTGGTAGAGG R: TGTAACAATTGCACCTTTTATGTTTT
P2	F: CATAACACATGCACGCTTTTT R: TGTTGGAGGCTGCAATACAGA
P3	F: GTACCTTCCATTCCCCAGA R: GCACCTATAGATTTCCACTACGA
P4	F: AATATAGCTCCAGATCCTGACTTTT R: TGCCTGCAACATATTCATCC

**1.3 测序** 上海天昊遗传分析中心对 40 例脱落细胞 DNA 样品进行基因测序检测, 用虾碱性磷酸酶(北京生物技术有限公司)和核酸外切酶 I(北京生物技术有限公司)纯化 PCR 产物。使用 Big-Dye Terminator v3.1 Cycle Sequencing Kit(美国 ABI 公司), 酒精纯化产物后, DNA 分析仪(ABI3130XL)测样品基因序列。40 例组织细胞 DNA 样本通过 Illumina HiSeq 平台, 以  $2\times 150\text{ bp}$  双端测序模式进行高通量测序, 获得 FastQ 数据。

**1.4 HPV16 变异株进化分析** 分析单核苷酸多态性(Polyphred 软件), 与欧洲原型病毒株(GenBank: NC\_001526.4)的序列比对, 同时与其他 HPV16 变异病毒株比对: 欧洲株 A1-A3: A1 (HQ644283.1, HQ644268.1, HQ644280.1, HQ644282.1), A2: (AF536179.1), A3: (HQ644236.1); 亚洲株 A4 [HQ644235.1 (A4), HQ644248.1, AF534061.1, HQ644251.1]; 非洲株 AF: B (HQ644238.1, HQ644240.1, HQ644290.1, HQ644298.1); C (AF472509.1, HQ644239.1, HQ644249.1, HQ644237.1); 北美 NA: D1 (HQ644257.1); 亚美型 AA: D2 (HQ644263.1, HQ644277.1, HQ644279.1, HQ644281.1); D3 (HQ644247.1, HQ644255.1, HQ644253.1, AF 402678.1)。MEGA11 构建系统发育树。

**1.5 统计学处理** 直接计数 HPV16 E4、L2 基因变异位点的变异频数, SPSS 26.0 软件分析统计结果,

使用  $\chi^2$  检验的统计方法, 分析 HPV16 E4、L2 单核苷酸变异与宫颈癌的相关性。  $P < 0.05$  为差异有统计学意义。

## 2 结果

**2.1 HPV16 E4、L2 基因的变异分析** HPV16 E4 和 L2 基因变异结果: 80 例 DNA 样本基因测序结果, 其多态性位点见表 2、3。

表 2 HPV16 E4 基因变异与氨基酸的改变

变异位点	AA	As	Af	变异样本数	变异比例	氨基酸变化
				(n=72)	(%)	
G2504A	G	G	G	1	1.39	K12K
T2520C	T	C	T	28	38.89	L18L
A2541C	A	A	A	1	1.39	T25P
C2546T	T	T	T	70	97.22	T26T
C2546A	T	T	T	1	1.39	T26T
G2585A	A	A	A	25	34.72	P39P
G2621C	G	G	G	1	1.39	Q51H
A2636C	A	A	A	3	4.17	E56D
T2660C	T	C	T	28	38.89	C64C
T2672G	T	T	T	5	6.94	T68T
T2703G	T	T	T	1	1.39	L79V

AA: 亚美株; As: 亚洲株; Af: 非洲株

在 80 例 HPV16 阳性样本中, 72 例样本存在 HPV16 E4 基因变异, 10 个位点存在核苷酸变异, 包括 4 种错义变异和 7 种同义变异, 其中错义变异分别为: A2541C (1/72, 1.39%)、G2621C (1/72, 1.39%)、A2636C (3/72, 4.17%) 和 T2703G (1/72, 1.39%)。氨基酸改变分别为: 苏氨酸替换为脯氨酸 (T25P)、谷氨酰胺替换为组氨酸 (Q51H)、谷氨酸替换为天冬氨酸 (E56D)、亮氨酸替换为缬氨酸 (L79V)。常见同义变异为 nt2520 (T-C) (28/72, 38.89%)、nt2546 (C-T/A) (70/72, 97.22%; 1/72, 1.39%)、nt2585 (G-A) (25/72, 34.72%)、nt2660 (T-C) (28/72, 38.89%)。

在 80 例 HPV16 阳性样本中, 74 例样本存在 HPV16 L2 变异, 40 个位点存在核苷酸变异, 包括 23 种错义变异和 18 种同义变异。如表 3 所显示, HPV16 L2 最常见的错义变异核苷酸变异位点为: T4177C、A4362C、A4362T 和 A4654C, 其氨基酸变化为丝氨酸转变为脯氨酸 (S269P)、亮氨酸转变为苯丙氨酸 (L330F) 和异亮氨酸转变为亮氨酸 (I428L)。17 种同义变异中变异频率较高的位点为: nt4074 (G-A) (73/74, 98.65%) 和 nt4602 (A-G) (15/74, 20.27%), G4074A 普遍存在于变异样本中。

表3 HPV16 L2 基因变异和氨基酸改变

变异位点	AA	As	Af	变异样本数 (n=74)	变异率 (%)	氨基酸 变化
A3396C	A	A	A	1	1.35	K8N
C3457A	C	C	C	1	1.35	P29T
T3534C	T	T	T	2	2.70	F54F
C3592T	C	C	C	1	1.35	P74S
A3594C	A	A	A	1	1.35	P74P
T3633C	T	T	T	1	1.35	A87A
G3699T	G	G	G	9	12.16	V109V
A3715G	A	A	A	2	2.70	I115V
A3732T	A	A	A	6	8.11	P120P
C3861T	T	T	T	4	5.41	F163F
G3879A	G	G	G	1	1.35	L169L
T3912C	T	T	T	1	1.35	H180H
A3957G	A	A	A	1	1.35	E195E
C3984T	C	C	C	1	1.35	S204S
C4032T	C	C	C	1	1.35	R220R
G4074A	A	A	A	73	98.65	Q234Q
G4099A	G	G	G	1	1.35	V243I
G4144A	G	G	G	1	1.35	E258K
T4177G	C	C	C	2	2.70	S269A
T4177C	C	C	C	33	44.59	S269P
C4178A	C	C	C	1	1.35	S269Y
A4180G	A	A	A	1	1.35	S270G
G4181A	G	A	G	4	5.41	S270N
G4186A	G	G	G	4	5.41	D272N
A4192C	A	A	A	1	1.35	S274R
C4205A	C	C	C	1	1.35	A278D
T4275A	T	T	T	1	1.35	I301I
A4288C	A	A	A	4	5.41	I306L
A4362C	A	C	A	33	44.59	L330F
A4362T	A	C	A	27	36.49	L330F
A4362G	A	C	A	5	6.76	L330L
G4390A	G	G	G	1	1.35	E340K
A4414G	A	A	A	1	1.35	T348A
A4437G	A	A	A	1	1.35	A355A
T4441A	T	T	T	1	1.35	S357T
A4464G	A	A	A	1	1.35	G364G
T4494C	T	T	T	1	1.35	I374I
C4505T	T	T	C	6	8.11	S378F
A4518T	A	A	A	1	1.35	V382V
A4602G	A	A	A	15	20.27	L410L
C4623T	T	C	C	1	1.35	P417P
A4628C	A	A	A	1	1.35	N419T
A4654C	A	C	A	23	31.08	I428L

AA:亚美株;As:亚洲株;Af:非洲株

**2.2 HPV16 E4 和 L2 核苷酸序列系统进化树** 使用 N-J 法, Bootstrap method (1000 replication) 和 Kimura 2-parameter model 构建 HPV16 E4、L2 基因系统进化树,只显示大于 50% 的 Bootstrap 值,其中 Bootstrap 值 >70% 提示可靠性很好,系统发育树见图 1。通过 HPV16 E4、L2 的系统进化树可以发现 76 个样本为欧洲变体,4 个样本为亚洲变体。没有

发现非洲变体、美洲变体和亚美变体。亚洲变体 X18、X20、Z91 和 CN-148 中的变异都与 G4181A 变异相关。

**2.3 病例组与对照组中 HPV16 E4、L2 的基因变异**

**2.3.1 病例组和对照组中 HPV16 E4 基因组的基因变异** 如表 4 所示,通过对病例组(宫颈癌组)与对照组(非宫颈癌组)中 E4 基因变异进行统计,发现对照组有 1 种错义变异,病例组有 3 种错义变异,统计结果显示,这些错义变异在对照组与病例组之间没有显著差异( $P > 0.05$ )。

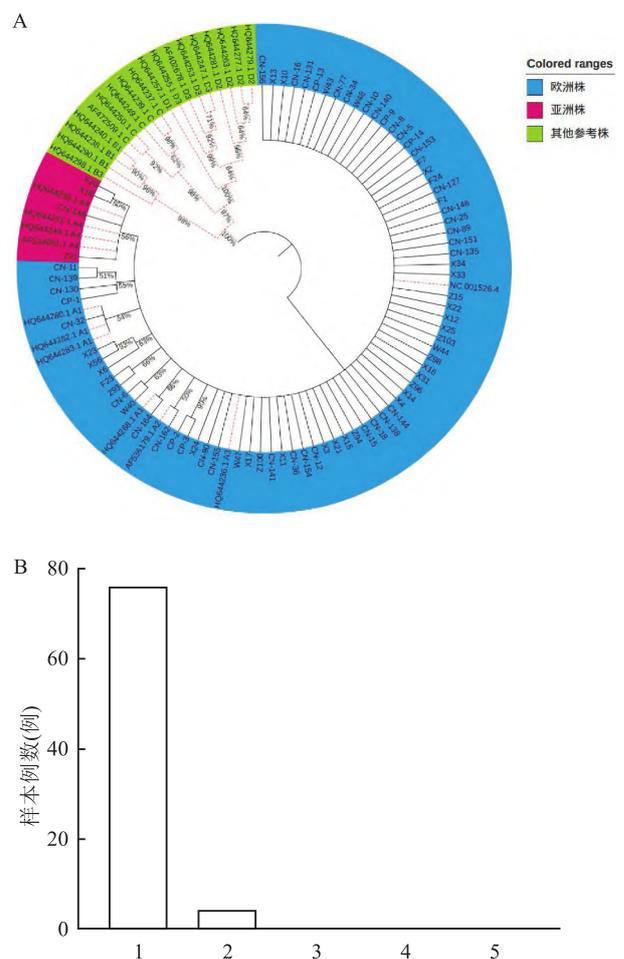


图1 HPV16 病毒 E4 和 L2 基因进化分析及病毒株统计

A: HPV16 病毒 E4 和 L2 基因进化分析,红色虚线连接的为参考病毒株; B: HPV16 病毒 E4 和 L2 基因进化病毒株统计;1:欧洲株;2:亚洲株;3:非洲株;4:北美株;5:亚美株

**2.3.2 病例组与对照组基因组 HPV16 L2 的遗传变异** 如表 5 显示的对照组和病例组中 L2 基因的错义变异分布。对照组有 17 种错义变异,病例组有 14 种错义变异,其中错义变异 T4177C( $P = 0.012$ ),

表4 病例组与对照组中 HPV16 E4 的基因变异

变异位点	对照组 (n=40)		病例组 (n=40)		$\chi^2$ 值	P 值	氨基酸变化
	变异样本数	变异频率 (%)	变异样本数	变异频率 (%)			
A2541C	0	0	1	2.5	1.013	0.314	T25P
G2621C	1	2.5	0	0	1.013	0.314	Q51H
A2636C	0	0	3	7.5	3.117	0.077	E56D
T2703G	0	0	1	2.5	1.013	0.314	L79V

HPV16 E4 参考序列:NC\_001526.4

表5 病例组与对照组中 HPV16 L2 基因组的基因变化

变异位点	对照组 (n=40)		病例组 (n=40)		$\chi^2$ 值	P 值	氨基酸变化
	变异样本数	变异频率 (%)	变异样本数	变异频率 (%)			
A3396C	1	2.5	0	0	1.013	0.314	K8N
C3457A	0	0	1	2.5	1.013	0.314	P29T
C3592T	0	0	1	2.5	1.013	0.314	P74S
A3715G	0	0	2	5.0	2.051	0.152	H115V
G4099A	1	2.5	0	0	1.013	0.314	V243I
G4144A	1	2.5	0	0	1.013	0.314	E258K
T4177G	2	5.0	0	0	2.051	0.152	S269A
T4177C <sup>a</sup>	11	27.5	22	55.0	6.241	0.012	S269P <sup>a</sup>
C4178A	1	2.5	0	0	1.013	0.314	S269Y
A4180G	1	2.5	0	0	1.013	0.314	S270G
G4181A	1	2.5	3	7.5	1.053	0.305	S270N
G4186A	3	7.5	1	2.5	1.053	0.305	D272N
A4192C	1	2.5	0	0	1.013	0.314	S274R
C4205A	0	0	1	2.5	1.013	0.314	A278D
A4288C <sup>a</sup>	0	0	4	10.0	4.211	0.04	I306L <sup>a</sup>
A4362C	18	45.0	15	37.5	0.464	0.496	L330F
A4362T	12	30.0	15	37.5	0.503	0.478	L330F
G4390A	1	2.5	0	0	1.013	0.314	E340K
A4414G	0	0	1	2.5	1.013	0.314	T348A
T4441A	1	2.5	0	0	1.013	0.314	S357T
A4464G	1	2.5	0	0	1.013	0.314	G364G
C4505T	4	10.0	2	5.0	0.721	0.396	S378F
A4628C	0	0	1	2.5	1.013	0.314	N419T
A4654C <sup>a</sup>	7	17.5	16	40.0	4.943	0.026	I428L <sup>a</sup>

HPV16 L2 参考序列:NC\_001526.4;<sup>a</sup>:有显著差异的位点

A4288C ( $P=0.04$ ) 和 A4654C ( $P=0.026$ ) 在病例组变异频率明显高于对照组,且差异有统计学意义 ( $P<0.05$ ),氨基酸变化分别为丝氨酸转变为脯氨酸 (S269P)、异亮氨酸转变为亮氨酸 (I306L) 和异亮氨酸转变为亮氨酸 (I428L)。差异有统计学意义的错义变异位点测序结果,见图 2。

### 3 讨论

HPV 基因变异的研究大多聚焦在病毒 E6 和 E7 基因上,而 E6 和 E7 的变异会对宫颈癌的发生与发展产生影响。HPV16 E4 基因变异和缺失会影响病毒粒子在细胞间的传播,L2 基因的变异和缺失会导致病毒基因组整合到细胞核染色体过程的改

变<sup>[8-13]</sup>。

通过 HPV16 E4、L2 的系统进化树显示,有 76 个样本为欧洲变异株,4 个样本为亚洲变异株。这与已有报道<sup>[15]</sup>的新疆维吾尔自治区主要流行株为欧洲株,少数为亚洲株的情况类似,然而 4 例亚洲变异株均与 L2 基因错义变异 G4181A 有关,这一发现尚未有相关文献报道。在 HPV16 E4 上常见的错义变异位点发生频率并不高,且 HPV16 E4 基因位于 E2 基因的中心位置,该区域编码 E2 蛋白质的铰链结构域<sup>[6-7]</sup>。E4 基因最常见的同义变异 C2546T (97.22%) 和 G2585A (34.72%) 在 HPV16 E2 基因的表达中为错义变异,且位于 E2 基因铰链区间接导致 E2 蛋白质的氨基酸改变,最终导致 E6、E7 表

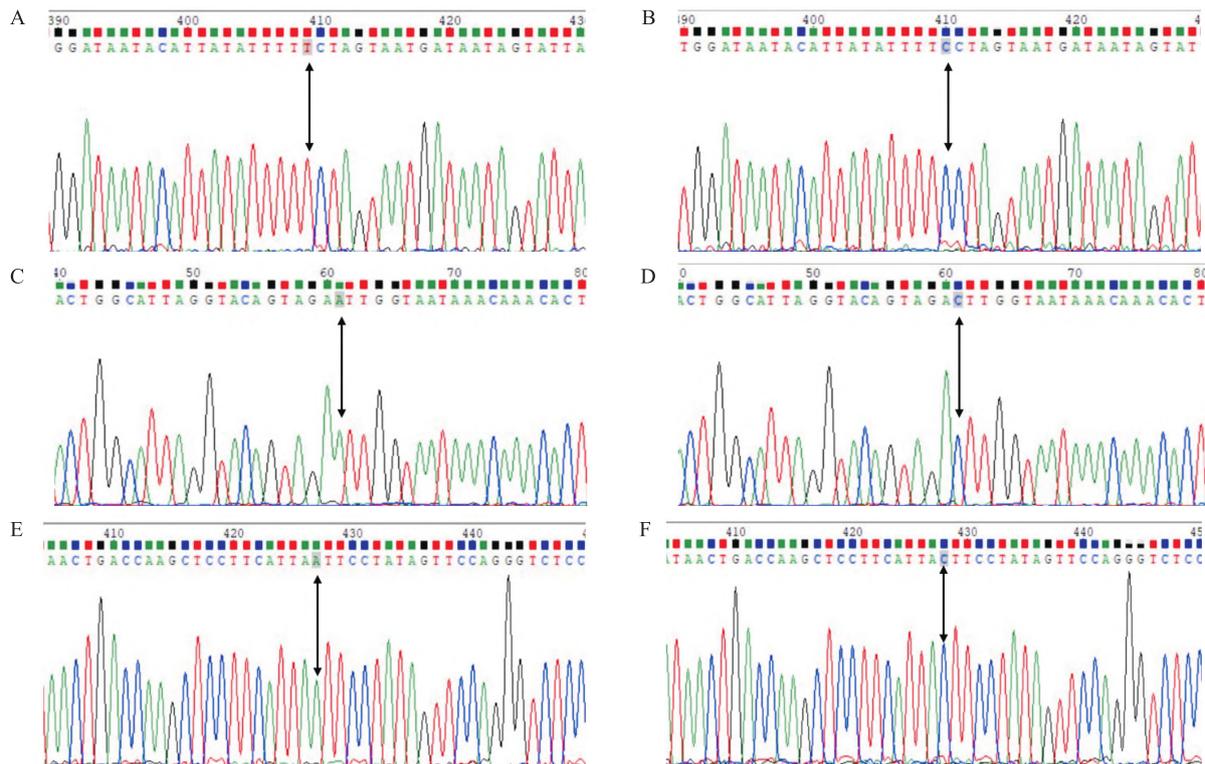


图2 HPV16 L2 基因错义变异具有显著差异位点的测序结果

A: 4 177 位点原型;B: 4 177 位点变异型;C: 4 288 位点原型;D: 4 288 位点变异型;E: 4 654 位点原型;F:4 654 位点变异型;黑色箭头指向变异的位点

达的变化,表 2 显示的 E4 基因变异 G2504A、T2520C、C2546T、G2585A、T2660C 和 T2703G 与已有的关于 E4 的报道结果一致<sup>[16]</sup>,而错义变异 A2541C、G2621C、A2636C 和同义变异 T2672G 尚未见文献报道。

在 HPV16 L2 基因上常见的错义变异位点分别为 T4177C、A4362C/T 和 A4654C,其氨基酸变化分别为 S249P、L330F 和 I428L,且同义变异 G4074A 普遍存在于样本中。研究表明 L2 基因的错义变异 T4177C、A4288C 和 A4654C 在病例组中高于对照组,差异有统计学意义( $P < 0.05$ )。L2 蛋白质对病毒基因组进入细胞核有促进作用,可以表明该变异可能会促进病毒基因组的感染和整合,这仍需要深入研究给予证明。由于 HPV16 E4 和 L2 在病毒感染和传播方面的作用,E4 基因的改变会间接影响 E2 基因的变化,最终导致 E6 和 E7 表达的改变,因此,可以通过 E4 基因的变异间接了解病毒的复制及传播能力。此外,还可以结合细胞实验研究 HPV16 L2 基因 T4177C、A4288C 和 A4654C 的变异对病毒基因组侵入细胞与整合到细胞基因组的作用。

综上所述,新疆维吾尔自治区主要流行病毒株为欧洲株,少数为亚洲株;HPV16 L2 基因中 T4177C、A4288C、A4654C 在宫颈癌中的变异频率高于非宫颈癌,且 G4181A 与亚洲株相关。

### 参考文献

- [1] Sung H, Ferlay J, Siegel R, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021,71(3): 209-49.
- [2] Wang J, Tang D, Wang K, et al. HPV genotype prevalence and distribution during 2009-2018 in Xinjiang, China: baseline surveys prior to mass HPV vaccination[J]. BMC Womens Health, 2019,19(1):90.
- [3] Zhou Z, Yang H, Yang L, et al. Human papillomavirus type 16 E6 and E7 gene variations associated with cervical cancer in a Han Chinese population[J]. Infect Genet Evol, 2019,73:13-20.
- [4] Burk R D, Harari A, Chen Z. Human papillomavirus genome variants[J]. Virology, 2013,445(1-2):232-43.
- [5] Pan Z, Song Y, Zhe X, et al. Screening for HPV infection in exfoliated cervical cells of women from different ethnic groups in Yili, Xinjiang, China[J]. Sci Rep, 2019,9(1):3468.
- [6] Doorbar J. The E4 protein; structure, function and patterns of expression[J]. Virology, 2013,445(1-2):80-98.

- [7] Tsakogiannis D, Ruether I, Kyriakopoulou Z, et al. Molecular and phylogenetic analysis of the HPV 16 E4 gene in cervical lesions from women in Greece[J]. *Arch Virol*, 2012,157(9):1729-39.
- [8] Leeman A, Jenkins D, Marra E, et al. Grading immunohistochemical markers p16 and HPV E4 identifies productive and transforming lesions caused by low- and high-risk HPV within high-grade anal squamous intraepithelial lesions[J]. *Br J Dermatol*, 2020, 182(4):1026-33.
- [9] Stevenson A, Kavanagh K, Pan J, et al. Risk stratification of cervical disease using detection of human papillomavirus (HPV) E4 protein and cellular MCM protein in clinical liquid based cytology samples[J]. *J Clin Virol*, 2018,108:19-25.
- [10] Kajitani N, Schwartz S. Role of viral ribonucleoproteins in human papillomavirus Type 16 gene expression[J]. *Viruses*, 2020, 12(10):1110.
- [11] Namvar A, Bolhassani A, Hashemi M. HPV16 L2 improves HPV16 L1 gene delivery as an important approach for vaccine design against cervical cancer[J]. *Bratisl Lek Listy*, 2016,117(3):179-84.
- [12] Wüstenhagen E, Hampe L, Boukhallouk F, et al. The cytoskeletal adaptor obscurin-like 1 interacts with the Human papillomavirus 16 (HPV16) capsid protein L2 and is required for HPV16 endocytosis[J]. *J Virol*, 2016,90(23):10629-41.
- [13] Li S, Bronnimann M, Williams S, et al. Glutathione contributes to efficient post-Golgi trafficking of incoming HPV16 genome[J]. *PLoS One*, 2019,14(11):e0225496.
- [14] Pouyanfar S, Spagnoli G, Bulli L, et al. Minor capsid protein L2 polytope induces broad protection against oncogenic and mucosal Human papillomaviruses[J]. *J Virol*, 2018,92(4):e01930-17.
- [15] Zhe X, Xin H, Pan Z, et al. Genetic variations in E6, E7 and the long control region of human papillomavirus type 16 among patients with cervical lesions in Xinjiang, China[J]. *Cancer Cell Int*, 2019,19:65.
- [16] 赵雨笛, 姚月婷, 周紫云, 等. 人乳头瘤病毒16型E4基因突变与宫颈癌发生发展的相关性研究[J]. *病毒学报*, 2019,35(4):584-91.

## Genetic variation analysis of *Human Papillomavirus 16 E4* and *L2* in women with cervical infection in Xinjiang Uygur Autonomous Region

Cheng Haozheng<sup>1</sup>, Tuo Jing<sup>1</sup>, Dong Yangliu<sup>1</sup>, Wang Le<sup>1</sup>, Zhe Xiangyi<sup>2</sup>,  
Li Hongtao<sup>2</sup>, Li Dongmei<sup>1</sup>, Pan Zemin<sup>1</sup>

(<sup>1</sup>Dept of Biochemistry and Molecular Biology, Shihezi University School of Medicine, Key Laboratory of Xinjiang Endemic and Ethnic Diseases, Shihezi 832002; <sup>2</sup>Dept of Anatomy and Histology and Embryology, Shihezi University School of Medicine, Shihezi 832002)

**Abstract Objective** To explore the nucleotide variation and protein amino acid changes of E4 and L2 genes of *Human Papillomavirus 16* (HPV16), and to analyze the evolutionary characteristics of HPV16 virus. **Methods** 40 HPV16 infection-positive cervical exfoliated cells samples and tissue cell samples were collected from hospital, viral DNA was extracted, Sanger sequencing perform in cervical exfoliated cells DNA and high-throughput sequencing technology sequenced in cervical tissues DNA for E4 and L2 genes of HPV16, HPV16 E4 and L2 gene phylogenetic evolution trees were constructed, and variation of HPV16 E4 and L2 genes were analyzed. **Results** There were 72 HPV16 E4 variant samples with nucleotide variants (4 missense mutations and 7 synonymous mutations) at 10 sites, HPV16 L2 gene variants in 74 samples, and nucleotide variants (23 missense mutations and 18 synonymous mutations) at 40 sites. The variation frequency of T4177C, A4288C and A4654C in cervical cancer was significantly higher than that in non-cervical cancer, and the difference was statistically significant ( $P < 0.05$ ). **Conclusion** ① The main HPV16 virus strains in Xinjiang are European strains, and a few are Asian strains. ② The mutation frequency of T4177C, A4288C and A4654C in HPV16 L2 gene is higher than that in non-cervical cancer, and G4181A is related to the Asian strain.

**Key words** *Human Papillomavirus 16*; gene variation; cervical cancer; E4 and L2 genes; single nucleotide polymorphisms