

·信息研究·

## 遗传性多发性骨软骨瘤致病基因 *EXT1* 和 *EXT2* 的多态性研究

房 玮, 裴元元, 黄玮俊, 胡 彬, 王一鸣\*

(中山大学中山医学院医学遗传学教研室, 广东广州 510080)

**摘要:**【目的】研究 *EXT1* 与 *EXT2* 基因在我国南方正常人及致病性突变已明确的遗传性多发性骨软骨瘤病人中的多态性。【方法】选取 50 例中国南方健康汉族个体及 13 例致病突变已明确的遗传性多发性骨软骨瘤病人, 提取基因组 DNA, 对包含 5'UTR 区、编码区、外显子-内含子交界区及 3'UTR 区的 PCR 产物进行直接测序。鉴定基因内的遗传变异, 并将结果和国际数据库中的数据进行对比。【结果】在所有研究对象中共发现 15 个不同的 *EXT1* 基因的单核苷酸多态性(SNPs): 3 个在编码区, 为同义突变, 11 个在内含子区, 1 个在 3'UTR 区; 其中有 3 个 SNPs 为数据库未报道的新发现多态位点。22 个 *EXT2* 基因的 SNPs: 3 个在编码区, 也为同义突变, 16 个在内含子区, 3 个在 3'UTR 区; 其中有 7 个 SNPs 为本研究新发现多态位点。【结论】中国南方汉族人与遗传性多发性骨软骨瘤病人 *EXT1* 和 *EXT2* 基因的多态性与 dbSNP 数据库登记资料不完全一致, 某些 SNPs 可能存在种族差异。本研究不仅有利于了解 *EXT1* 和 *EXT2* 基因结构, 也为其多态性与致病性突变的鉴定提供了依据。

**关键词:** *EXT* 基因; 遗传性多发性骨软骨瘤; 中国人; 基因多态性

**中图分类号:** R394.5      **文献标志码:** A      **文章编号:** 1672-3554(2011)02-0263-07

### Investigation of Polymorphisms in Causal Genes of HME: *EXT1* and *EXT2*

FANG Wei, PEI Yuan-yuan, HUANG Wei-jun, HU Bin, WANG Yi-ming\*

(1. Department of Medical Genetics, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510080, China)

**Abstract:** 【Objective】To investigate the polymorphisms of the *EXT1* and *EXT2* genes in Southern Han Chinese people and in 13 hereditary multiple exostoses (HME) patients whose causal mutations had already been identified. 【Methods】Genomic DNA was extracted from 50 healthy unrelated southern Chinese Han individuals and 13 HME patients with known casual mutations in these two genes. PCR products spanning the coding regions, intron-exon boundaries, 5' and 3'UTR of the genes were sequenced. Variations identified in the sequences were compared with those retrieved from the databases. 【Results】Fifteen genetic variations of *EXT1* were identified, of which 3 are synonymous coding variations, 11 in intronic region and 1 in 3'UTR. 3 of the 15 variations were novel SNPs. Twenty-two variations of *EXT2* were detected, 3 are synonymous coding variations, 16 in intronic region and 3 in 3'UTR. 7 of the 22 genetic variations in *EXT2* were novel. 【Conclusions】Ten novel variations in *EXT1* and *EXT2* gene have been identified in our study. Twenty-seven variations already in the databases were confirmed. Our results provide useful information for the genetic variations of the two genes among Southern Han Chinese. The results also provide important information for differentiation of causal mutations of *EXT1* and *EXT2* genes from innocuous polymorphisms.

**Key words:** *EXT* genes; hereditary multiple exostoses; Chinese; polymorphism

[J SUN Yat-sen Univ(Med Sci), 2011, 32(2): 263-268, 278]

遗传变异广泛存在于人类基因组中, 其不仅是个体鉴定(含亲子鉴定)的重要生物标志物<sup>[1]</sup>, 而且部分遗传变异还和疾病的易感性或表型特征

相关。因此, 挖掘人群的遗传变异有重要的应用价值。遗传性多发性骨软骨瘤(hereditary multiple exostoses, HME)又名遗传性多发性外生骨疣, 是一

收稿日期: 2010-09-25

基金项目: 985 II 期工程资助

作者简介: 房玮, 硕士生, 医学遗传学, E-mail: cnfangwei@163.com; \* 通信作者, 王一鸣, 教授, 博士生导师, E-mail: ywzhong@hotmail.com

种罕见的常染色体显性遗传性骨骼异常疾病。主要表现为累及长骨的骨骼发育异常,四肢长骨骨骺附近常形成大小不等的骨性隆起<sup>[2]</sup>。西方人群中报道的发病率约为 1/5 万,约有 0.5 ~ 2% 的患者会恶变为骨肉瘤或软骨肉瘤<sup>[2]</sup>。研究表明分别位于人类染色体 8q24.11-q24.13 的 *EXT1*<sup>[3]</sup> (MIM ID \* 608177)、11p12-p11 的 *EXT2*<sup>[4]</sup> (MIM ID \* 608210) 和 19p 的 *EXT3*<sup>[5]</sup> (MIM ID \* 600209) 基因的病理性突变可导致 HME。但基因中的遗传变异并非全为病理性变异,且部分变异为中性的,即与疾病无关。鉴定和区分致病性突变及基因多态性对本病的诊断有重要的临床意义。*EXT1* 基因碱基对数目约为 312 kb,含 11 个外显子,编码含 746 个氨基酸的蛋白质;*EXT2* 基因碱基对数目约为 149 kb,含 14 个外显子,编码含 718 个氨基酸的蛋白质。在人类基因突变数据库中 (human genome mutation database, HGMD, <http://www.hgmd.cf.ac.uk/>), 目前发现与 HME 有关的 *EXT1* 基因突变有 174 种, *EXT2* 基因突变 84 种,且各种突变类型均有报道。为了挖掘这两个基因的多态性在我国人的分布情况,我们对 50 个正常汉族个体及 13 个致病性突变已明确<sup>[6]</sup>的 HME 个体进行了 5'UTR 区、编码区、外显子-内含子交界区及 3'UTR 区的测序,并在本文中报道结果。

## 1 材料与方法

### 1.1 研究对象

随机选取 50 例经中山大学附属医院体检中心检查证实为健康,彼此无血缘关系的中国南方汉族个体;13 例经专科医生根据国际诊断标准<sup>[7]</sup>确诊且致病突变已明确的中国南方 HME 病人,在获得参与者的知情同意后,采集外周血。病人基本资料见表 1<sup>[6]</sup>。

### 1.2 DNA 提取

外周血 200  $\mu\text{L}$ 。用天根小提试剂盒 (天根生化,北京)抽提 DNA,严格按试剂盒说明进行操作。

### 1.3 遗传变异的检测

针对 *EXT1* (GeneBank NM\_000127.2) 和 *EXT2* (GeneBank NM\_207122.1) (<http://genome.ucsc.edu/>) 基因的编码区、外显子-内含子交界区、部分内含子及 UTR 区,使用 primer 5 软件进行引物设计,由上海生工公司合成 (表 2),并进行直接测

表 1 13 例 HME 患者的基本资料

Group <sup>1)</sup>	Patients <sup>2)</sup>	Age	Sex
Family			
1	1	12	male
2	2	15	male
3	3	16	female
4	4	31	male
5	5	14	male
6	6	13	male
7	7	12	male
8	8	11	male
9	9	20	male
Sporadic cases			
12	10	14	male
13	11	6	male
14	12	15	male
15	13	28	male

1) Group numbers please refer to reference 6. 2) Patients 1-9 were from nine HME families. Patients 10-13 were sporadic cases. No common ancestry was traced among the families or sporadic cases. Pathogenic mutations identified in these patients please see reference 6

序。PCR 反应体系为 30  $\mu\text{L}$ ,含 1 $\times$ PCR Buffer 3  $\mu\text{L}$ 、2 mmol/L dNTP 2.5  $\mu\text{L}$ 、2.5 mmol/L  $\text{MgCl}_2$  3  $\mu\text{L}$ 、3.2 mmol/L 上下游引物各 1  $\mu\text{L}$ 、1 U/ $\mu\text{L}$  Taq DNA 聚合酶 (Fermentas) 1  $\mu\text{L}$ 。反应在 Biometra PCR 仪 (Biometra German) 上完成。扩增程序为:95  $^\circ\text{C}$  预变性 3 min;95  $^\circ\text{C}$  变性 30 s、引物特异性退火温度 (见表 2) 45 s、72  $^\circ\text{C}$  延伸 45 s、共 38 个循环;最后 72  $^\circ\text{C}$  延伸 10 min。PCR 产物经大肠杆菌核酸外切酶和虾碱性磷酸酶 (Fermentas) 纯化后以 BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) 按标准程序进行测序。产物在 ABI3730 测序仪 (Applied Biosystems) 上分析。应用 Sequence Scanner v1.0 (Applied Biosystems) 对测序结果进行分析。

### 1.4 遗传变异的命名

本文参考人类基因组变异协会 (Human Genome Variation Society, HGVS) 所制定的序列变异命名法则,以起始密码子 ATG 的 A 为第+1 位核苷酸,对测序所得 SNPs 进行命名 (<http://www.hgvs.org/mutnomen/recs-DNA.html>)。

### 1.5 遗传变异与数据库的比对

本文把捕获的遗传变异数据与美国国立生物信息中心 (National Center for Biotechnology

表2 *EXT1*、*EXT2* 基因测序区域及相应引物Table 2 Sequencing regions and primers of *EXT1* and *EXT2* genes

Gene	Region <sup>1)</sup>	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'	Temperature of annealing(°C)
<i>EXT1</i>	E1-1	TCTTCTTTTTCTCCGTCT	GATTCCTACCATTGTTCCA	50
	E1-2	ATCGCCGAAAGTTACCAAA	CAGTTCAGGCTCAAAGGG	56
	E2	GCAACCCAACCTCCTTCCT	TAAATCCTCCCTCCACCCC	58
	E3	GATTTTGTGGTGGCAGC	GGTATTGAAAGGGGTGG	55
	E4	TAGACTAACCTAAGCAAC	ACTGGACCAATCACACAT	51
	E5	CACTCTTTTTCATTTGCTCC	ATGCTCTGCTCTGTTTTTG	52
	E6	TTTTCATCTTGTCTGTCT	GTGCTTGTATCTTCTTTAG	48
	E7	AATGTTCTGAGTTGTGTGG	TTTGAAAGCCTATTGTGGTC	52
	E8	GCAGGTGAGGATGGGAGAA	AACAAAAAGGAGGGCAGG	54
	E9	TTTTTCTTCCTTTCATTTTC	TTTTCACACTTCAGACCAACT	48
	E10	GTTCTTCTTTCTAACTCACC	ACCAATCATACTCTTTTTTC	51
E11	AATCTTTGTTTTGTGTTT	ATTCTTGATAATGGGTAG	52	
<i>EXT2</i>	E2	TTGTTATTTACACACACCAC	GCAGACTACTCTTCACGGGC	59
	E3	GTGCGTTCATTTTTCCCTG	GTGCCCCATTACCCTCTA	57
	E4	CTGGGAAGTAAGGAAAGGG	AGCATAAAGCAGAGGAGCC	52
	E5	GTGGAGGTGAAGACTGGT	CTGGACTGCCTTATGAAT	57
	E6	ACAAGGTAGGCTGAGGTAA	AATGGAAAATAGAGAAAAA	50
	E7	AGTGAAGAAGGGAGGGGAA	GCAGGAAAATGGCAAATGT	53
	E8	TAAACTCTGCCATAACAC	ATTTCCAACAAGACCCAT	50
	E9	TTTTGAGGAGGGGAAGACT	GGAAGAACCAATGGGAACT	50
	E10	TTTTGCTTCTCTGCCTCTCA	CTTACGCACACCTTTTGGAC	55
	E11	CAGCATCTGTCTTTGAGTTT	GTCACACAATATGGCTTTTC	50
	E12	ATGCCTCCTTTACCCCTTC	CATCATAGCCTTTATTCTG	53
	E13	CAGTTACAGAAGGCAAAAAG	TACCTGAAAAATAATCCA	50
	E14	TTCACCGCTTTTTTCCACAC	CAGGTCCCAAGGCATTTATC	56

1)E:exon

Information,NCBI) 的 dbSNP 数据库(<http://www.ncbi.nlm.nih.gov/snp>)进行比对,得到最终结果。

### 1.6 基因型频率及等位基因频率计算

基因型频率及等位基因频率参照以下公式进行计算:基因型频率=某一基因型例数/该组总例数;等位基因频率=(纯合子例数×2+杂合子例数)/(总例数×2)<sup>[8]</sup>。

## 2 结果

### 2.1 健康个体组结果

对 50 例正常人样本进行 *EXT1* 和 *EXT2* 基因部分编码区及邻近内含子区直接测序,所得结果与 NCBI 的 dbSNP 数据库(<http://www.ncbi.nlm.nih.gov/snp>)中已经登录的遗传变异进行比较。*EXT1* 总测序长度约 2.6 kb,得到 7 个遗传变异

(表 3),6 个为已报道变异位点,1 个为新发现多态位点。*EXT1* 的 7 个遗传变异位点中,1 个是缺失型,6 个是替换型;6 个替换型 SNPs 中,4 个为转换,2 个为颠换。所有 7 个位点,2 个位于编码区,5 个位于内含子区。位于编码区的 2 个 SNPs 均为同义突变。c.963-105 C>T(rs76603199),c.1536+7 A>G 和 c.1431 C>T(rs17439693)为未在病人组发现的 SNPs,4 个多态位点(rs10098992,rs36225999,rs17479145 和 rs7837891)与病人组多态位点重复。

*EXT2* 总测序长度约 1.8 kb,得到 2 个遗传变异(表 3),1 个为已报道变异,另一个为新发现多态位点。这些多态性位点中,除 rs36225999 属插入/缺失(deletion/insertion polymorphism,DIP)外,其余均为单核苷酸多态(single nucleotide polymorphisms,SNPs)。*EXT2* 的 2 个 SNPs 均为替

表 3 *EXT* 基因的遗传变异在健康人组中的位置,类型,基因型及等位基因频率Table 3 Characters of *EXT* genetic variations in healthy people

No.	Genetic variation <sup>1)</sup>	Genotype freq <sup>2)</sup>	Allele freq	Identity in dbSNP
<i>EXT1</i>				
cSNP-1	c.1431 C>T	CC/CT 0.933/0.067	C/T 0.967/0.033	rs17439693
cSNP-2	c.1761 G>A	GG/GA/AA 0.236/0.654/0.109	G/A 0.564/0.436	rs7837891
Intron-1	c.963-105 C>T	CT/TT 0.036/0.964	C/T 0.018/0.982	rs76603199
Intron-2	c.1536+59 G>T	GG/GT 0.933/0.067	G/T 0.967/0.033	rs10098992
Intron-3	c.1536+85_93delCTCCCCAGG	WW/W-0.933/0.067	W/- 0.967/0.033	rs36225999
Intron-4	c.1536+7 A>G	AA/AG 0.917/0.083	A/G 0.958/0.042	novel
Intron-5	c.1723-103 C>G	CC/CG 0.704/0.296	C/G 0.852/0.148	rs17479145
<i>EXT2</i>				
cSNP-1	c.805 C>T	CC/CT 0.984/0.016	C/T 0.992/0.008	novel
Intron-1	c.919+316 C>G	CC/CG/GG 0.667/0.217/0.116	C/G 0.775/0.225	rs11037875

dbSNP: database of single nucleotide polymorphisms. 1) Standard nomenclature ([www.hgvs.org/mutnomen/](http://www.hgvs.org/mutnomen/)) was used for describing sequence variations, with +1 corresponding to the A of the ATG translation initiation codon of GeneBank NM\_000127.2 for *EXT1* and NM\_207122.1 for *EXT2*.

2) W: wild type; -: deletion. Amino acid substitution: In *EXT1*, p.Pro477Pro for cSNP-1, p.Glu587Glu for cSNP-2; In *EXT2*, Leu269Leu for cSNP-1

换型, 其中 1 个为转换, 1 个为颠换。一个在编码区, 一个在内含子区, 编码区 SNP 为同义突变。2 个 SNPs (c.805 C>T 和 c.939 +316 C>G (rs11037875)) 均为病人组未发现的 SNPs。

## 2.2 病人组结果

对 13 例 HME 病人样本进行 *EXT1* 和 *EXT2* 基因直接测序, 所得结果与 NCBI 的 dbSNP 数据库 (<http://www.ncbi.nlm.nih.gov/snp>) 中已经登录的多态性位点进行比较。*EXT1* 总测序长度 8.9 kb, 共得到 12 个遗传变异位点(表 4), 10 个为数据库已报道的多态性位点, 2 个为数据库未报道的新发现 SNPs。*EXT1* 的 12 个遗传变异位点中有 2 个是缺失型, 10 个是替换型; 10 个替换型 SNPs 中有 5 个是转换, 5 个是颠换。所有 12 个变异位点中, 2 个位于编码区, 9 个位于内含子区, 1 个位于 3'UTR 区。位于编码区的 2 个 SNPs 经过与数据库对比, 均为同义突变。

*EXT2* 总测序长度 6.1 kb, 共得到 20 个多态性位点(表 4), 其中 14 个为已报道多态位点, 6 个为新发现 SNPs。*EXT2* 的 20 个遗传变异位点中有一个是缺失型, 其余 19 个为替换型; 19 个替换型 SNPs 里, 9 个为转换, 10 个为颠换。所有 20 个多态性位点中, 2 个位于编码区, 15 个位于内含子区域, 3 个位于 3'UTR。位于编码区的 2 个 SNPs 的 2 个 SNPs 经过与数据库对比, 均为同义突变。

## 3 讨论

人类基因组序列存在着很大的多态性, 这是人类基因组组成的重要特征之一, 其中最常见的是单核苷酸多态性 (single nucleotide polymorphisms, SNP)。SNP 指基因组上单个核苷酸的变异引起的 DNA 序列多态性, 包括单个碱基的转换、颠换。SNPs 不仅分布在基因组非编码区域, 而且存在于基因的编码序列中, 称为编码 SNPs (coding SNPs, cSNPs)。据报道, 人类基因组有 25~40 万个 cSNPs, 其中, 20%~30% 引起蛋白质编码序列改变的非同义突变的 SNPs (non-synonymous SNPs, nsSNPs) 会影响蛋白质的功能和某些生物学性状<sup>[9-11]</sup>。发生于基因蛋白质编码区以外的 SNPs, 可通过改变相关基因的调控元件来影响基因的功能, 从而引起人类某些性状的改变和疾病的发生<sup>[12-13]</sup>。因此, 遗传变异可为中性, 也可为疾病相关的或致病性的, 这三者的鉴别具临床意义。

本研究对 50 例正常的中国南方汉族人及 13 例已发现致病突变的 HME 病人进行 *EXT1* 和 *EXT2* 基因多态性分析。所有 63 例测序样品中, 在 *EXT1* 基因 8.9 kb 的测序区域内, 共发现 15 个不同的遗传变异, 3 个在编码区, 11 个在内含子区, 1 个在 3'UTR 区; 与 NCBI 中的 dbSNP 数据库进行比较, 有 3 个 SNPs 数据库中未报道。在 *EXT2* 基

表 4 *EXT* 基因的遗传变异在病人组中的位置, 类型, 基因型及等位基因频率Table 4 Characters of *EXT* genetic variations in HME patients<sup>1)</sup>

No. <sup>2)</sup>	Genetic variation <sup>3)</sup>	Genotype freq <sup>4)</sup>	Allele freq <sup>4)</sup>	Identity in dbSNP
<i>EXT1</i>				
cSNP-1	c.1065 C>T	CC/CT 0.667/0.333	C/T 0.833/0.167	rs11546829
cSNP-2	c.1761 G>A	GG/GA/AA 0.385/0.231/0.385	G/A 0.500/0.500	rs7837891
Intron-1	c.1164+98 A>T	AA/AT 0.923/0.077	A/T 0.961/0.039	novel
Intron-2	c.1165-156 G>T	GG/GT 0.769/0.231	G/T 0.884/0.116	rs9642837
Intron-3	c.1284+66 G>A	GG/GA 0.615/0.385	G/A 0.807/0.193	rs17439812
Intron-4	c.1536+59 G>T	GG/GT 0.615/0.385	G/T 0.807/0.193	rs10098992
Intron-5	c.1536+85_93delCTCCCCAGG	WW/W- 0.583/0.417	W/- 0.792/0.209	rs36225999
Intron-6	c.1537-51 A>G	AA/AG/GG 0.461/0.385/0.154	A/G 0.653/0.347	rs10955837
Intron-7	c.1537-125 G>A	GG/GA/AA 0.833/0.083/0.083	G/A 0.875/0.125	rs4876757
Intron-8	c.1633-26 C>A	CC/CA 0.846/0.154	C/A 0.923/0.077	novel
Intron-9	c.1723-103 C>G	CC/CG 0.769/0.231	C/G 0.885/0.115	rs17479145
3'UTR-1	c.* 144del A	WW/W- 0/1.000	W/- 0.500/0.500	rs71739430
<i>EXT2</i>				
cSNP-1	c.28 C>A	CC/CA 0.583/0.417	C/A 0.792/0.208	rs4755228
cSNP-2	c.1641C>T	CC/CT 0.923/0.077	C/T 0.961/0.039	rs75987184
Intron-1	c.-110 A>C	AA/AC 0.833/0.167	A/C 0.917/0.083	rs75844681
Intron-2	c.-177 A>T	AA/AT/TT 0.667/0.250/0.083	A/T 0.792/0.208	rs11037864
Intron-3	c.744-82 A>G	AG/GG 0.333/0.667	A/G 0.167/0.833	rs10769018
Intron-4	c.1080-18 T>A	TT/TA/AA 0.385/0.385/0.230	T/A 0.576/0.424	rs11037882
Intron-5	c.1080-115 G>T	GG/TT 0.400/0.600	G/T 0.400/0.600	rs11037881
Intron-6	c.1173+103 C>T	CC/CT 0.846/0.154	C/T 0.923/0.077	novel
Intron-7	c.1173+302 C>T	CC/CT 0.923/0.077	C/T 0.961/0.039	novel
Intron-8	c.1306-93 C>T	CC/CT/TT 0.230/0.385/0.385	C/T 0.424/0.576	rs4755233
Intron-9	c.1496-195 C>T	CC/CT 0.231/0.769	C/T 0.615/0.385	rs1027377
Intron-10	c.1496-166 A>C	AA/AC 0.923/0.077	A/C 0.961/0.039	novel
Intron-11	c.1807-51 T>C	TT/TC/CC 0.333/0.416/0.251	T/C 0.541/0.459	rs11037909
Intron-12	c.1936-132_133delGT	WW/W- 0.923/0.077	W/- 0.961/0.039	rs3832748
Intron-13	c.1936-66 G>C	GG/GC 0.923/0.077	G/C 0.039/0.961	novel
Intron-14	c.1936-41 T>C	TT/TC/CC 0.385/0.385/0.230	T/C 0.578/0.422	rs3740878
Intron-15	c.2019-84 T>G	TT/TG 0.923/0.077	T/G 0.961/0.039	rs75740761
3'UTR-1	c.* 56 G>A	GG/GA 0.923/0.077	G/A 0.961/0.039	rs77554103
3'UTR-2	c.* 129 C>G	CC/CG 0.923/0.077	C/G 0.961/0.039	novel
3'UTR-3	c.* 139 C>G	CC/CG 0.923/0.077	C/G 0.961/0.039	novel

1)HME patients: HME patients with known casual mutations; 2)UTR: Untranslated region. 3)Standard nomenclature ([www.hgvs.org/mutnomen/](http://www.hgvs.org/mutnomen/)) was used for describing sequence variations, with +1 corresponding to the A of the ATG translation initiation codon of GeneBank NM\_000127.2 for *EXT1* and NM\_207122.1 for *EXT2*. 4) W: wild type; -: deletion. Amino acid substitution: In *EXT1*, Cys355Cys for HME patients cSNP-1, p.Glu587Glu for cSNP-2; In *EXT2*, p.Arg10Arg for cSNP-1, p.Asp547Asp for cSNP-2

因 6.1kb 的测序区域内, 共发现 22 个不同的遗传变异位点, 3 个在编码区, 16 个在内含子区, 3 个在 3'UTR 区; 7 个为 dbSNP 数据库未报道 SNPs。*EXT1* 基因所发现的 cSNPs: rs17439693 (p.Pro477Pro)、rs7837891 (p.Glu587 Glu)、rs11546829

(p.Cys355 Cys) 以及 *EXT2* 基因的 cSNPs: c.805C>T (p.Leu269Leu, novel)、rs4755228 (p.Arg10Arg)、rs75987184 (p.Asp547Asp) 都未能引起氨基酸改变, 因而对其所编码蛋白质的结构和功能不产生影响。由于所研究的对象为正常个体, 故所发现的

变异应是非致病性变异。对于遗传性多发性骨软骨瘤(HME)患者而言,由于研究对象的致病突变已明确,且所发现的变异均不发生于基因的功能节段,如编码或调控节段,因此本研究所发现的 15 个 *EXT1* 基因的遗传变异位点与 22 个 *EXT2* 基因的变异位点应是非 HME 致病性遗传变异,这为鉴定 HME 致病性突变与多态性提供了第一手的证据。dbSNP 数据库中的数据主要来源于高加索人群,与中国人群在遗传背景上有所不同,该研究所发现的新多态位点可能由于种族差异,为中国人所特有。*EXT* 基因在中国人群的多态性分布还有待于进一步大规模样本的研究来补充。

已有研究表明约 90% HME 与 *EXT1* 或 *EXT2* 基因突变相关<sup>[14]</sup>。有文献报道 HME 病人的致病突变主要为造成相应蛋白质截短的移码及无义突变<sup>[15]</sup>,且在中国 HME 病人当中,*EXT2* 突变的频率高于 *EXT1*<sup>[16]</sup>。本文的另一前期研究发现了 13 个 HME 患者的致病突变,6 个位于 *EXT1* 基因,7 个位于 *EXT2* 基因<sup>[6]</sup>,并且本研究发现 HME 病人 *EXT1* 基因的遗传变异多态位点有 12 个,*EXT2* 基因的遗传变异多态位点有 20 个。因此,就多态性而言,*EXT2* 的多态性远高于 *EXT1* 的多态性。其机理尚不明确。据现有的研究,由于没有这些多态性与两基因编码蛋白功能的任何关系的证据,因此,对于 HME 而言,这些多态性可能为中性的。*EXT1* 和 *EXT2* 基因均为广泛表达。目前的功能研究已证实 *EXT* 基因编码的蛋白具有蛋白多糖中硫酸肝素 (heparan sulphate, HS) 合成酶的功能,参与 HS 链的合成与延伸<sup>[17]</sup>。exostosin-1 (*EXT1*) 和 exostosin-2 (*EXT2*) 这两种基因产物是位于内质网的 II 型跨膜糖蛋白,可形成定位于高尔基体的异聚体,催化硫酸肝素 (heparan sulphate, HS) 的多聚物形成,通过细胞之间和细胞与基质之间的一系列相互作用,参与很多重要的生理过程: 生长因子/受体信号传导、细胞增殖及分化等过程,从而可能影响骨骼的正常生长<sup>[17]</sup>。

本研究不仅有利于了解 *EXT1* 与 *EXT2* 基因结构,且为在中国人中研究 *EXT* 基因的遗传变异的多态性及致病性突变的鉴定提供了证据。此外,当今国际上日趋成熟的外显子组学疾病致病位点研究,以筛除掉存在于 dbSNP 中的多态性位点为重要步骤<sup>[18]</sup>。SNP 研究能完善现有 dbSNP 数据库,使之成为今后疾病外显子组学研究行之有效的工

具。

#### 参考文献:

- [1] Amar A, Brautbar C, Motro U, et al. Genetic variation of three tetrameric tandem repeats in four distinct Israeli ethnic groups[J]. J Forensic Sci, 1999, 44(5):983-986.
- [2] Wuyts W, Van HW. Molecular basis of multiple exostoses: mutations in the *EXT1* and *EXT2* genes [J]. Hum Mutat, 2000, 15(3):220-227.
- [3] Cook A, Raskind W, Blanton SH, et al. Genetic heterogeneity in families with hereditary multiple exostoses[J]. Am J Hum Genet, 1993, 53(1):71-79.
- [4] Wu YQ, Heutink P, Vries BB, et al. Assignment of a second locus for multiple exostoses to the pericentromeric region of chromosome 11 [J]. Hum Mol Genet, 1994, 3(1): 167-171.
- [5] Le MM, Legeai ML, Jeannin PM, et al. A gene for hereditary multiple exostoses maps to chromosome 19p [J]. Hum Mol Genet, 1994, 3(5): 717-722.
- [6] Pei YY, Wang YM, Huang WJ, et al. Novel Mutations of *EXT1* and *EXT2* genes among Families and Sporadic Cases with Multiple Exostoses [J]. Genet Test Mol Biomarkers, 2010, 14(6): 865-872.
- [7] Peterson HA. Multiple hereditary osteochondromata [J]. Clin Orthop, 1989(239):222-230.
- [8] 李璞, 孙开来. 医学遗传学 [M]. 2 版. 北京: 中国协和医科大学出版社, 2006: 157-159.
- [9] Chasman D, Adams RM. Predicting the functional consequences of nonsynonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation [J]. JMB, 2001, 307(2): 683-706.
- [10] Sunyaev S, Ramensky V, Koch I, et al. Prediction of deleterious human alleles. Hum Mol Genet, 2001, 10(6):591-597.
- [11] Ng PC, Henikoff S. Predicting deleterious amino acid substitutions[J]. Genome Res, 2001, 11(5): 863-874.
- [12] Chorley BN, Wang X, Campbell MR, et al. Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies[J]. Mutat Res, 2008, 659(1/2): 147-157.
- [13] Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations [J]. Nat Rev Genet, 2002, 3(5):391-397.
- [14] Bovée JV. Multiple osteochondromas [J]. Orphanet J