

泛肽相关蛋白基因 UBAP1 单核苷酸多态及与鼻咽癌的相关研究*

熊 炜^{1)**} 曾朝阳^{1)**} 沈守荣²⁾ 李小玲¹⁾ 李伟芳¹⁾
周 鸣¹⁾ 李 江¹⁾ 贺 林³⁾ 冯国鄞⁴⁾ 李桂源^{1)***}

(¹⁾中南大学湘雅医学院肿瘤研究所, 长沙 410078; (²⁾中南大学湘雅三医院消化内科, 长沙 410078;

(³⁾上海交通大学 Bio-X 生命科学研究中心, 上海 200031; (⁴⁾上海市精神卫生中心, 上海 200031)

摘要 UBAP1 基因是在鼻咽癌 9p 最小共同缺失区内新克隆的鼻咽癌候选抑癌基因, 通过采用病例-对照研究方法, 对 105 例鼻咽癌患者和 183 例正常人 UBAP1 基因的 5 个单核苷酸多态 (SNPs) 用测序法进行了分型, 在实验过程中, 偶然发现了一个新的 SNP 位点, 并已登录至 dbSNP (登录号: rs3135929). 经相关分析发现, 位于 UBAP1 基因第 6 外显子中的 SNP rs1049557 与鼻咽癌发病存在显著相关性, 基因型 GG 和 GC 的相对危险度分别为 1.64 和 1.31. 实验结果进一步支持了 UBAP1 基因与鼻咽癌的发生发展可能存在密切关系, SNP rs1049557 由于处在 UBAP1 基因下游 3' 非翻译区, 其多态类型可能在某种程度上影响 UBAP1 基因的表达调控, 从而与鼻咽癌发病相关.

关键词 单核苷酸多态, UBAP1 基因, 鼻咽癌, 相关分析

学科分类号 R739.63

鼻咽癌 (nasopharyngeal carcinoma, NPC) 是东南亚及我国南方各省常见的一种恶性肿瘤^[1,2], 其发病具有明显的地域性和家族聚集性. 流行病学调查发现, 5%~10% 的鼻咽癌患者有家族遗传史, 说明遗传因素在鼻咽癌的发生发展中起重要作用. 鼻咽癌的病因涉及 EB 病毒的感染、化学致癌物作用及多种癌基因与抑癌基因结构和功能的改变^[3~5], 而且存在基因组不稳定的特性. 我室通过全基因组扫描和比较基因组杂交 (comparative genomic hybridization, CGH) 发现, 鼻咽癌基因组存在高频率 3p、9p、11q、13q 和 14q 的等位基因杂合性丢失^[6], 随后采用定位候选克隆策略从 9p21~22 最小共同缺失区内克隆了与鼻咽癌相关的肿瘤抑制候选基因 UBAP1 (泛肽相关蛋白, ubiquitin associated protein, GenBank 登录号: AF222043)^[7]. UBAP1 基因在鼻咽癌活检组织中表达明显下调, 其功能正在深入研究中.

单核苷酸多态 (single nucleotide polymorphisms, SNPs) 是随着人类基因组计划的实施而发展起来的新一代遗传标记, 由于其分布广、密度高而被期望在对诸如癌症、糖尿病等多基因疾病的研究中起重要作用. 为了进一步了解 UBAP1 基因与鼻咽癌的关系, 我们从 dbSNP 数据库中选取定位于 UBAP1 基因内的 4 个 SNPs, 利用测序的方法, 对 UBAP1 基因多态与鼻咽癌发病的相关性进行了研究.

1 材料和方法

1.1 样品准备

本课题采用病例-对照方法. 105 例未经放疗处理的初诊鼻咽癌患者外周血样品取自湖南省肿瘤医院, 所有病例均经组织病理学确证, 除 1 例鼻咽腺样囊性癌和 2 例泡状核细胞癌外, 其余均为鼻咽低分化鳞状细胞癌. 患者平均年龄 47 ± 11 岁, 75.7% 为男性. 正常人外周血样品取自长沙市中心血站健康无偿献血员, 从 2 000 余例中, 依照尽可能与病例组匹配的原则分层随机抽取 183 例, 平均年龄 47 ± 12 岁, 75.4% 为男性. 每份样品均取血 10 ml, 酸性柠檬酸葡萄糖溶液 B (acid citrate dextrose solution B) 抗凝, 依常规酚-氯仿法抽提白细胞基因组 DNA.

1.2 SNP 位点及分型

当 SNP 位于基因的编码序列中时, 称为 cSNP (coding region SNP), 若 cSNP 引起蛋白质重要部位氨基酸的变异, 可导致其功能的改变; 位于基因调控序列中的 SNP 则可能影响基因表达的剂量. 因此, 一般认为这两种 SNP 的生物学意义更为显

* 国家重点基础研究发展规划 (973) (G1998051008) 和国家自然科学基金资助项目 (30100027).

** 熊炜和曾朝阳为并列第一作者.

*** 通讯联系人.

Tel: 0731-4805446, E-mail: Ligy@public.cs.hn.cn

收稿日期: 2002-03-11, 接受日期: 2002-05-28

著, 是基因组中决定人类表型多样性的核心信息^[8]. 在 dbSNP (<http://www.ncbi.nlm.nih.gov/SNP>) 中, UBAP1 基因只有第 6 外显子中有 3 个 cSNP 登录, 这 3 个 SNP rs1049557, rs10939 和 rs1049567 恰好紧挨在 184 bp 范围内, 故用一对引物 (forward primer 1: 5'-agtgtccttcccacttcagc, reverse

primer 1: 5'-gcctgtctgcttaccattc) 就可以同时扩增、测序, 对 3 个 SNP 进行基因分型; 同时我们还随机选取了位于第一内含子中的 1 个 SNP rs1537218 (测序引物为 forward primer 2: 5'-aatgtgatgattccaaagc, reverse primer 2: 5'-agaaatcccaaaccaacaac), 各 SNP 在 UBAP1 基因上的位置见图 1.

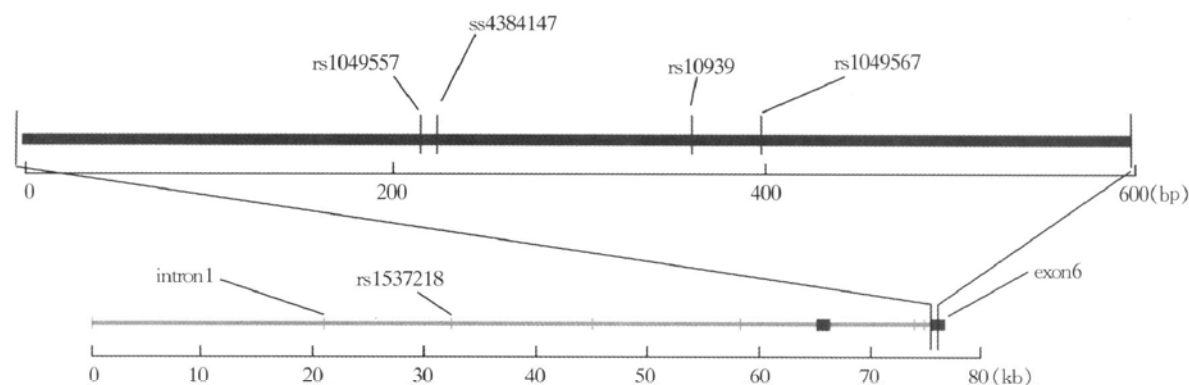


Fig. 1 Locations of SNPs in UBAP1 gene

PCR 反应总体积为 25 μ L, 包括 Tris-HCl (pH 8.3) 10 mmol/L、KCl 50 mmol/L、MgCl₂ 1.5 mmol/L、dNTPs 各 0.2 mmol/L、引物各 0.8 μ mol/L、基因组 DNA 10 ng、Hotstar Taq 酶 (Qiagen 公司) 1 U. PCR 反应在 GeneAmp PCR system 9700 (PE Applied Biosystems) 上进行, 反应循环参数为: 首先 94 $^{\circ}$ C 12 min, 然后 94 $^{\circ}$ C 30 s、58 $^{\circ}$ C 30 s、72 $^{\circ}$ C 30 s (30 个循环), 最后 72 $^{\circ}$ C 7 min.

PCR 产物经 384 Multiscreen filter plates (Millipore 公司) 纯化, 用 ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit 做测序反应, 最后在 ABI 3100 DNA 测序仪 (PE Applied Biosystems) 上测序.

测序结果 (* .ab1 文件) 使用 Phred 软件自动识别各样品碱基序列^[9,10], Phrap 软件将各个样品的测序结果排列对齐, Polyphred 软件自动搜索 SNP 位点^[11], 最后由 Consed 软件展示 SNP 位点及峰型图^[12]. 以上整套 4 种软件均在 Linux 操作系统上运行.

1.3 统计分析

依方积乾等所述方法^[13]进行卡方检验、计算相对危险度和置信区间, 所有统计分析均在 SPSS10.0 (SPSS Inc.) 软件上完成.

2 结 果

用测序法对全部 288 个样品均成功地完成了 4

个 SNP 的分型 (图 2), 在 rs10939 位点我们没有

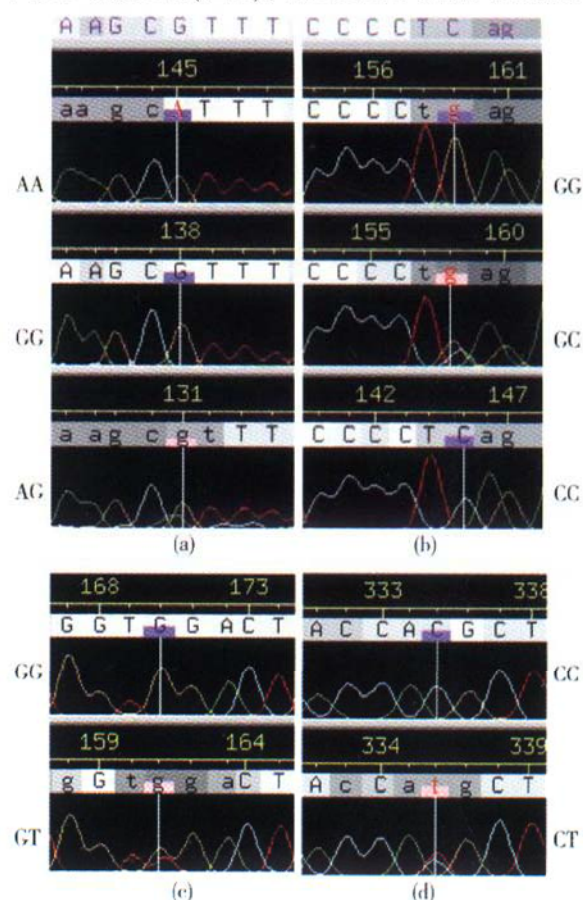


Fig. 2 Trace view of rs1537218A/G (a), rs1049557C/G (b), Novel SNP G/T (c), and rs1049567 C/T (d) in Consed

Genotype "TT" of ss4384147 and rs1049567 are not found in our subjects.

检测到多态, 即未见 GT 或 TT 基因型, 所有样品均为 GG, 但同时我们偶然地在 SNP 位点 rs1049557 下游仅 9 bp 处发现了一个新的 SNP 位点 (已登录至 dbSNP, 登录号: ss4384147). 我们随机抽取了 16 份 (6%) 样品以同一方向测序两次, 另随机抽

取 16 份样品以不同方向各测序一次, 重复实验与第一次结果完全吻合, 说明我们的测序方法重复性好, 结果可靠.

5 个位点 SNPs 基因频率见表 1.

Table 1 Allele frequencies of SNPs in UBAP1 gene

SNPs	rs1537218		rs1049557 ¹⁾		ss4384147		rs10939		rs1049567		%
	A	G	C	G	G	T	G	T	C	T	
Control	14.9	85.1	56.6	43.4	94.3	5.7	100	0	93.7	6.3	
Case	11.6	88.4	43.3	56.7	95.7	4.3	100	0	93.3	6.7	

¹⁾ $\chi^2 = 9.345$, $v = 1$, $P = 0.002$.

从表 1 中可以看出, 除 rs1049557 基因频率在病例和对照中存在显著差异以外 ($P = 0.002$), 其余位点要么未见多态 (rs10939), 要么由于杂合性太小 (次要等位基因频率 $< 15\%$, ss4384147 和 rs1049567 的次要等位基因纯合子 TT 均未检测到), 基因频率在病例和对照中均未见显著差异.

SNP rs1049557 的基因型频率及相对危险度列于表 2, 在病例-对照之间, rs1049557 基因型频率也存在显著差异 ($P < 0.01$), 基因型 GG 和 CG 的相对危险度分别为 1.64 和 1.31 (95% 置信区间分别为 1.07~1.62 和 1.13~2.38). 也就是说, 在所检测人群中, SNP rs1049557 等位基因 GG 和 CG 的携带者患鼻咽癌的可能性分别是 CC 基因型的 1.64 和 1.31 倍.

Table 2 Frequency distribution among cases and controls and relative risks associated with genotypic variants of SNP rs1049557

Genotype	Frequency		RR ¹⁾	95% CI
	Control	Case		
CC	75 (41.0%)	29 (27.6%)	1.00	
CG	57 (31.1%)	33 (31.4%)	1.31	1.07~1.62
GG	51 (27.9%)	43 (41.0%)	1.64	1.13~2.38

¹⁾ $\chi^2 = 6.777$, $v = 1$, $P < 0.01$. RR: relative risk; CI: confidence interval.

3 讨 论

基因组多态性研究是后基因组时代的一个重要研究领域, 而据估计在人类基因组中约 90% 以上的多态形式是 SNP, 截止到 2002 年 2 月 12 日, 在 dbSNP 中登录的 SNP 已达到 4 126 321 个 (<http://www.ncbi.nlm.nih.gov/SNP>), 也就是说, 在人

类基因组中平均每 700 bp 左右就存在一个 SNP. SNP 多态系统由于其极高的密度和便于高通量分析等优点受到越来越多的关注, 人们已寄希望于它来定位肿瘤、精神疾患、心血管疾病、糖尿病等一系列多基因疾病的主基因^[14], 并最终攻克这些现代医学的难题.

UBAP1 基因是本实验室在 9p 鼻咽癌高频缺失区克隆的, 可能与 Ubiquitin pathway 相关的抑癌基因候选者. 已有研究结果表明该基因在鼻咽癌活检组织中表达明显下调或缺失^[7], 对该基因的功能及其在鼻咽癌发生发展过程中所起的作用正在深入研究之中.

本文采用测序这一准确、高效的 SNP 分型方法, 首次对 UBAP1 基因的单核苷酸多态及其与鼻咽癌的相关性进行了研究, 结果表明 SNP 位点 rs1049557 与鼻咽癌存在显著相关 ($P = 0.002$, 基因型 GG 和 CG 的相对危险度分别为 1.64 和 1.31). 本研究进一步支持了 UBAP1 基因与鼻咽癌的发生发展可能存在密切关系. SNP rs1049557 虽然位于 UBAP1 基因第 6 外显子中, 但并不是在 UBAP1 基因开放阅读框 (ORF) 中, 而是位于终止密码子后 412 bp 处, 距加尾信号 555 bp. 真核基因 3' 非翻译区可能与 mRNA 的寿命有关^[15], SNP rs1049557 的多态类型是否在某种程度上影响 UBAP1 基因的表达调控, UBAP1 基因在鼻咽癌中表达下调或缺失是否确实与 rs1049557 存在某种联系值得进一步研究.

SNP 是人类在漫长进化过程中基因组与内外环境交互作用的累积结果, 虽然 SNP 的数目极其繁多、密度很高, 但值得指出的是不同的 SNP 位点由于其在进化过程中产生有先有后、在不同人群

中分布频率不一定一样、杂合程度亦不尽相同,在利用它们进行相关分析或者连锁不平衡分析时,所能提供的信息量也可能存在较大差别^[16],因而在进行有关的研究时,SNP位点的选择显得十分重要。本研究中共使用了5个SNP位点,然而除了rs1049557外,其余4个位点均由于以上原因(杂合度<0.15)而未能提供有效信息,在正常人群和鼻咽癌患者中未见明显差异。这一点,提醒我们要在后续研究中引起重视。

参 考 文 献

- 1 Parkin D M, Laara E, Muir C S. Estimates of the worldwide frequency of sixteen major cancers in 1980. *Int J cancer*, 1988, **41** (2): 184~ 197
- 2 Hildesheim A, Levine P H. Etiology of nasopharyngeal carcinoma: a review. *Epidemiol Rev*, 1993, **15** (2): 466~ 485
- 3 余 鹰, 谢 奕, 李桂源, 等. 一个新鼻咽癌抑瘤候选基因的克隆及其功能初步分析. *生物化学与生物物理进展*, 2000, **27** (3): 319~ 324
Yu Y, Xie Y, Li G Y, *et al.* *Prog Biochem Biophys*, 2000, **27** (3): 319~ 324
- 4 余 鹰, 朱诗国, 李桂源, 等. BRD7 单核苷酸多态性及鼻咽癌易感性分析. *生物化学与生物物理进展*, 2001, **28** (4): 568~ 572
Yu Y, Zhu S G, Li G Y, *et al.* *Prog Biochem Biophys*, 2001, **28** (4): 568~ 572
- 5 向娟娟, 余 鹰, 王洁如, 等. 一个有假基因特点的鼻咽癌相关基因的分离和克隆. *生物化学与生物物理进展*, 2002, **29** (1): 51~ 55
Xiang J J, Yu Y, Li G Y, *et al.* *Prog Biochem Biophys*, 2002, **29** (1): 51~ 55
- 6 李忠花, 王 璐, 李桂源, 等. 比较基因组杂交研究鼻咽癌遗传变异. *中华医学遗传学杂志*, 2001, **18** (5): 338~ 342
Li Z H, Wang L, Li G Y, *et al.* *Chin J Med Genet*, 2001, **18** (5): 338~ 342
- 7 Qian J, Yang J B, Li G Y, *et al.* Isolation and Characterization of a novel cDNA, UBAP1, derived from the tumor suppressor locus in human chromosome 9p21-22. *J Cancer Res Clin Oncol*, 2001, **127** (10): 613~ 618
- 8 Michele C, David A, James I, *et al.* Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat genet*, 1999, **22** (3): 231~ 238
- 9 Ewing B, Hillier L, Wendt M C, *et al.* Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Res*, 1998, **8** (3): 175~ 185
- 10 Ewing B, Green P. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Res*, 1998, **8** (3): 186~ 194
- 11 Nickerson D A, Tobe V O, Taylor S L. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res*, 1997, **25** (14): 2745~ 2751
- 12 Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. *Genome Res*, 1998, **8** (13): 195~ 202
- 13 方积乾. *医学统计学与电脑实验*. 第二版. 上海: 上海科学技术出版社, 2001. 309~ 326
Fang J Q. *Medical Statistics and Computerized Experiment*. 2nd. Shanghai: Shanghai Science and Technology Publishers, 2001. 309~ 326
- 14 Leonid K. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet*, 1999, **22** (2): 139~ 144
- 15 童克中. *基因及其表达*. 北京: 科学出版社, 1998. 19~ 22
Tong K Z. *Gene and its expression*. Beijing: Science Press, 1998. 19~ 22
- 16 Goddard K A, Hopkins P J, Hall J M, *et al.* Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. *Am J Hum Genet*, 2000, **66** (1): 216~ 234

Studies of Single Nucleotide Polymorphisms in UBAP1 Gene and Their Association with Nasopharyngeal Carcinoma*

XIONG Wei^{1)**}, ZENG Zhao-Yang^{1)**}, SHEN Shou-Rong²⁾, LI Xiao-Ling¹⁾, LI Wei-Fang¹⁾,
ZHOU Ming¹⁾, LI Jiang¹⁾, HE Lin³⁾, FENG Guo-Yin⁴⁾, LI Gu-Yuan^{1)***}

⁽¹⁾ Cancer Research Institute, Xiangya Medical School of Central South University, Changsha 410078, China;

⁽²⁾ Department of Gastroenterology, 3rd Xiang-Ya Hospital, Central South University, Changsha 410078, China;

⁽³⁾ Bio-X Life Science Research Center, Shanghai Jiaotong University, Shanghai 200031, China;

⁽⁴⁾ Shanghai Mental Health Center, Shanghai 20031, China)

Abstract Nasopharyngeal carcinoma (NPC) is rare in most parts of the world, but prevalent in south China. Recently, UBAP1 gene, which is located in the region of minimal heterozygosity deletion at 9p21.3~ 22.1 and down-expressed in NPC, has been cloned. The latest results suggest that the UBAP1 gene is the candidate tumor suppressor for NPC. Association study using 5 single nucleotide polymorphisms (SNPs) within UBAP1 gene by means of sequencing was performed in 105 unrelated case subjects and 183 control subjects which matched to the NPC cases on age, sex and residence. Occasionally, a novel SNP has been found, and been submitted to the dbSNP (accession number: ss4384147). Significant result was obtained for one SNP mark (rs1049557), which is resident at 3' non-translation region of UBAP1 gene; the relative risk of this SNP mark is 1.64 (genotype GG) and 1.31 (genotype CG). The result has proved again that UBAP1 gene may play a

certain role in occurrence and development of nasopharyngeal carcinoma. The SNP mark rs1049557, considering it's residence, may influence on the expression of UBAP1 gene.

Key words single nucleotide polymorphisms, UBAP1 gene, nasopharyngeal carcinoma, association analysis

* This work was supported by grants from The Special Funds for Major State Basic Research of China (G1998051008) and The National Natural Sciences Foundation of China (30100027).

** XIONG Wei and ZENG Zhao Yang have contributed equally to this work.

** Corresponding author. Tel: 86-731-4805446, E-mail: Ligy@public.cs.hn.cn

Received: March 11, 2002 Accepted: May 28, 2002

第四届国际转基因动物学术讨论会 4th International Conference on Transgenic Animals Nov 26-29, 2002 Shanghai, China

大会主席: Dr. Wenhao Xu (Biochemistry and Molecular Biology University of Kansas Medical Center)

胡以平教授 (中国人民解放军第二军医大学转基因动物研究所)

大会副主席: 成国祥教授 (上海转基因研究中心主任)

大会主题: The Post-genomic Era of The Transgenic Technology (后基因组时代的转基因技术)

大会议程:

- 1 Overview: "The past and present of the transgenic technology"
开幕词: 基因技术的过去与现在
- 2 Keynote speech: "TG and GT: what are they coded in research, medicine, and biotechnology?"
主题报告: TG 和 GT 在基础研究、医学及生物技术领域中的应用
- 3 Transgenic frontiers
转基因技术前沿
 - (1) Transgenics for fundamental research (cancer, cardiovascular diseases, neuroscience, AIDS)
转基因与基础研究: a. 肿瘤; b. 心血管疾病; c. 神经科学; d. 感染性疾病 (如 AIDS、HBV 等)
 - (2) Transgenics for biotechnology (therapeutics, genetic trait modification, xenotransplantation)
转基因与生物技术: a. 治疗学; b. 遗传性状修饰 (动物和植物, 例如转基因小鼠); c. 异种移植
 - (3) Novel transgenic technology (inducible, viral mediated, tosk)
转基因新技术: a. 诱导型; b. 病毒介导; c. Tosk
- 4 Knockout rounds
基因剔除技术体系
 - (1) Embryonic stem cell technology 胚胎干细胞技术
 - (2) Programmed genetic modification 程序性遗传修饰
 - (3) Novel gene-targeting technology (Cre-loxP, Flp-FRT, chromosomal engineering, gene trap)
基因打靶新技术: a. Cre-loxp, Flp-FRT; b. 染色体工程; c. 基因捕获
 - (4) Human cell knockout 人类基因剔除
- 5 Transgenic banking
转基因建库
 - (1) Cryopreservation (Freeze-dry sperm) 冷藏法 (精液干冻法)
 - (2) IVF (ICSI)
 - (3) Transgenic informatics 转基因信息学
- 6 Closing remark: "The future of the transgenic technology"
闭幕词: 转基因技术的未来

请有意参会代表提交论文到大会组委会 (未提交论文者亦可参会)。

联系人: 刘小姐、沈小姐; 电子信箱: beijing@bilong.com meeting@bilong.com; 传真: 010-62015131

电话: 010-82015225、13910414026、13910414020; 地址: 100088 北京市海淀区北太平庄路 2 号院 2 号楼 1809 室

• 会议地点: 上海国际饭店 (中国上海南京西路 170 号) • 会议时间: 2002 年 11 月 26~29 日 (26 日报到) • 本届大会工作语言为英语 • 会务费 1500 元/每人 (含午餐费, 不包住宿; 需要提供住宿的参会代表收费为 80 元/天), 学生凭学生证减半, 参加过以往一届会议的优惠 5%, 二届的 10%, 三届的 15%

因篇幅有限, 详情请随时通过任何方式联系大会组委会