

# 恶性疟原虫 FCC1/HN 株新抗原表达 序列标记位 (ESTs) 的获得

阎宗合<sup>1</sup> 谢毅<sup>2</sup> 李明<sup>1</sup> 王萍<sup>1</sup>  
王燕妮<sup>1</sup> 毕惠祥<sup>1</sup> 李英杰<sup>1</sup>

1 第一军医大学热带医学研究所热带病研究室 广州 510515  
2 复旦大学遗传工程国家重点实验室 上海 200433

**摘要 目的:** 以筛选恶性疟原虫 FCC1/HN 株  $\lambda$ gt11 cDNA 表达文库所获得的强阳性克隆作基础, 对上述强阳性克隆的 cDNA 插入片段进行 DNA 序列测定, 阐明相对应的新表达序列标签 (ESTs), 作为发现新抗原基因的线索。**方法:** 以 cDNA 表达文库接头的较长链作 PCR 引物, 扩增 cDNA 插入片段, 将扩增产物克隆入 M13mp18 测序载体, 进行部分 DNA 序列测定、编辑, 将之在 GenBank 中进行 DNA 序列同源性搜索比较和分析。**结果:** 获得 1 个 C03 序列为已知恶性疟原虫热休克蛋白 70-2 基因片段, 发现 5 个新的具有抗原意义的恶性疟原虫表达序列标记位 (ESTs)。**结论:** 这 5 个新的恶性疟原虫表达序列标记位为发现新的恶性疟原虫抗原基因奠定了基础。

**关键词** 恶性疟原虫 抗原 cDNA 表达序列标记位

1994 年, 人类基因组第一代物理图谱研究小组负责人 Daniel Cohen 着手筹集资金启动恶性疟原虫基因组的研究<sup>[1]</sup>。1995 年, 英国 Wellcome Trust 基金会资助的国际协作组织成功地构建了恶性疟原虫 3D7 株的 YAC 文库, 标志恶性疟原虫基因组计划开始具体实施<sup>[2]</sup>。

在基因组研究中, 新表达序列标记位 (expressed sequence tags, ESTs) 的挖掘具有重要意义, 它们是从 cDNA 文库中获取克隆, 进行 cDNA 插入片段序列分析后而确定的标记序列, 可用于新基因的发现、基因组的作图以及鉴别基因组序列中的编码序列等。这种快速的 cDNA 定性策略为简化整个基因组序列测定中最有价值部分基因的标记和定位提供新的遗传标志, 成为生物研究领域广泛的源泉。1994 年, Chakrabarti 等报道了恶性疟原虫随机 cDNA 克隆的部分序列分析, 从中发现若干新恶性疟原虫 ESTs, 为恶性疟原虫基因组研究提供了宝贵资源<sup>[3]</sup>。恶性疟原虫基因组计划的提出启发了我们应在原有疟疾免疫工作基础上, 具有特色地进行新的抗原表达序列标记位的发掘。所以我们用特有的患者抗血清、免疫血清及单克隆抗体筛选恶性疟原虫 FCC1/HN 株  $\lambda$ gt11 cDNA 表达文库所获得的强阳性克隆作基础, 对上述强阳性克隆的 cDNA 插入片段进行 DNA 序列测定, 阐明相对应的新表达序列标记位 (ESTs), 作为发现新抗原基因的线索。

## 材料与方 法

### 材料

用兔免疫血清、患者抗血清及单克隆抗体筛选

恶性疟原虫 FCC1/HN 株  $\lambda$ gt11 cDNA 表达文库获得的强阳性克隆见参考文献。本文涉及的 6 个 cDNA 表达文库强阳性克隆 (编号按 C01-C06 顺序依次为: B109, E1230, H1225, L1230 II, D0109, B0104) 是经兔免疫血清初筛及复筛后, 与恶性疟患者混合血清、保护性或诊断性单抗呈强阳性反应的克隆<sup>[4]</sup>。

测序载体 M13mp18 由复旦大学遗传工程国家重点实验室保存, 大肠杆菌 TG1 (SupE hsd $\Delta$ 5 thi $\Delta$  (lac<sup>-</sup>proAB)F [traD36 proAB<sup>+</sup> lacIq lacZ $\Delta$ M15]) 购自 Boehringer Mannheim 公司, Y1090 菌种 (SupF hsdR araD139  $\Delta$ bn  $\Delta$ lacU169 rpl $\Delta$  trpC22::Tn10 (tet<sup>r</sup>) pMC9) 购自 GBCO BRL 公司, 各种限制性内切酶和修饰酶为德国 Boehringer Mannheim 公司产品或美国 Promega 公司产品, PCR 全套试剂购自复旦大学遗传所, Dye primer DNA 测序试剂盒为美国 ABI 公司产品。

### 方法

抗体筛选 cDNA 文库所获强阳性克隆 cDNA 插入片段的 PCR 扩增 从 cDNA 文库中用抗体筛选的强阳性克隆<sup>[4]</sup>  $\lambda$  嗜菌体培养上清经梯度稀释及铺板 (Y1090 铺平板菌), 挑取数个嗜菌斑作为 PCR 模板。PCR 引物为构建  $\lambda$ gt11 cDNA 文库时衔接  $\lambda$ gt11 和 cDNA 的 EcoR I (Not I) 接头之较长链 (5'-AATTCGCGGCCGCGTTCGAC-3'), 由复旦大学遗传工程国家重点实验室合成。在 PE9600 PCR 扩增仪上进行 DNA 扩增反应。

抗体筛选 cDNA 文库所获强阳性克隆 cDNA 插入片段的克隆测序 cDNA 片段扩增产物经末端磷酸化及补平, 与经限制性内切酶 Sam I 切割及处

理的M 13 mp18 测序载体连接, 转化, 挑取白斑, 鉴定, 保存目的上清, 抽单链, 在ABI 373A 型自动DNA 序列分析仪上进行DNA 序列测定。

所获恶性疟原虫 DNA 片段初步序列分析及 ESTs 的确定 从ABI 373A 型自动DNA 序列分析仪上拷贝所测 DNA 片段序列文件, 在PCGENE 软件上将其转换成PCGENE 核酸序列文件格式, 去掉接头序列和第一个“测序仪无法辨别的碱基”及其以后序列, 编辑成所需DNA 片段序列。将它们与EMBL/GenBank release 43 DNA 序列数据库在PCGENE 6.8 上进行DNA 同源性比较, 并通过Internet 网在美国NCBI (National Center for Biotechnology Information) GenBank 中的STS 数据

库 (STS Division)、EST 数据库 (EST Division) 以及 GenBank+ EMBL + DDBJ + PDB 核酸序列数据库用BLAST (Basic Local Alignment Search Tool) 命令进行DNA 序列同源性搜索、比较与分析, 确定所获得的抗原 ESTs 性质。

### 结 果

所获 DNA 片段的部分序列 以下列出的 DNA 序列为测序结果中介于接头序列和第一个‘测序仪无法辨别的碱基’之间的有效序列, 其为相应所筛 DNA 表达文库强阳性克隆插入片段的部分DNA 序列, 即上文提及的表达序列标记位 (EST) 序列。

<b>C01 序列</b>	<b>BASE</b>	<b>COUNT</b>	<b>94 A</b>	<b>50 C</b>	<b>25 G</b>	<b>137 T</b>
1	TTTA TATTTA	AAA TA TCCA T	TTTCA TCAAA	A TAAAAA GA T	TC T T C A C A T T	T T A A A G A G T C
61	A T C A C G T T T T	T G T T C T T T T A	T G T C A T G A G C	A T T A T A T C C A	C T T A A A A T A T	T A T T T A C C A A
121	C A A A G A A T T A	T T A T A T A T T G	G T A C A A A C T T	A T T C A T A T G T	T C A T T G T C T T	T A T T T A T T T G
181	T G T T C C C C C T	T T T A A G A G A T	G C A A A A T T T T	T T T T T C A A A C	C A G T A G G A T T	T G C C C A T T C A
241	T T A G T A T T C A	C A T C C T T C T T	T A A T A T A T C T	T C A T C T T T C T	G T A T T T C C A A	A A C A T T T G T T
301	A C T T T T					
<b>C02 序列</b>	<b>BASE</b>	<b>COUNT</b>	<b>117 A</b>	<b>68 C</b>	<b>59 G</b>	<b>119 T</b>
1	G T C A A G A T C T	A T A G A G A C A T	G C A G G A T G T T	T C A A A T T C T A	T C T T A A A A T T	T G A C C A A A T C
61	T T T G G T C A A A	T A A T A A T A C T	G T T A C T C A A A	T A T T A G C T T T	T T A C C A T T T A	G A G C A A A T T A
121	C A A A T C A G G T A	A C A T T C A A A A	T T C T A A C C C T	C C C T T C C C A A	C A T G G G A A T A	T C T A T T A A G T
181	A G G A A A T G A G	A G A A C C C C C T	A A T T T C C C A T	T T C A G T G A T A	A T A C T G A T T G	A G T A T C T C T T
241	A T C T G A N G T G	C T T G G G A C C A	G A A G T G T T T T	G G A T C T G G G T	T T T T C T C A G A	T T T T G A A A T A
301	T T T G C A T T A T	A C C A G T T G A G	T A T C C A A A A A	T C A A A A C T N G	G A A T C T G G A A	T G A T C C C A T G
361	G C C C G					
<b>C03 序列</b>	<b>BASE</b>	<b>COUNT</b>	<b>180 A</b>	<b>50 C</b>	<b>88 G</b>	<b>102 T</b>
1	A G T T T T T G A A	G G A G A A A G A G	C A T T A A C C A A	A G A T A A T C A C	C T T T T A G G A A	A G T T T G A A T T
61	A T C T G G T A T T	C C A C C A G C A C	A A A G A G G A G T	A C C C A A A A T T	G A A G T T A C C T	T T A C C G T A G A
121	C A A A A A T G G T	A T C T T A C A T G	T T G A A G C T G A	A G A C A A A G G T	A C A G G T A A A A	G T A G A G G T A T
181	A A C T A T T A C T	A A T G A C A A A G	G T A G A T T A T C	G A A A G A A C A A	A T C A A A A A A T	G A T T A A T G A T
241	G C A G A A A A A T	T C G C A G A T G A	A G A T A A A A A C	T T A A G A G A A A	A A G T T G A A G C	C A A A A A T A A C
301	C T T G A T A A T T	A T A T A C A G A G	T A T G A A A G C A	A C T G T T G A A G	G T A A A G A T T A	A A T T A G C T G A
361	T A A A A T C G G A	G A A G A A G A T A	A A A A T A C T A T	C C T T T C A G G C	T T G T A A A T G	G G G G G G G G G G
<b>C04 序列</b>	<b>BASE</b>	<b>COUNT</b>	<b>132 A</b>	<b>35 C</b>	<b>65 G</b>	<b>105 T</b>
1	T A A C T C T T T T	C A A A T A T A T G	A A G A A A A A A A	A A A A A A A G T A	A C A A A T G T T T	T G G A A A T A C A
61	G A A A G A T G A A	G A T A T A T T A A	A G A G G G A T G T	G A A T A C T A A T	G A T T G G G C A A	A T C C T A C T G G
121	T T T G A A A A A A	A A A T T T T G C A	T C T C T T A A A A	G G G G G A C C A C	A A A T A A T T A A	A G C C A A T G A C
181	C A T A T G A T T A	G G T T G T A C C	A A T A T A T A A T	A A T T C T T T G T	T G G T A A A T A A	T A T T T T A A G T
241	G G G T A T A A T G	C T C A T G G C A T	A A A A G G C C C A	A A C C G G G G T G	G C C C T T T A A A	A T G T G A A G G G
301	T C T T T T T A T T	T T G G T G G A A T	G G G T A T T T T T	A A A T T A A		
<b>C05 序列</b>	<b>BASE</b>	<b>COUNT</b>	<b>140 A</b>	<b>81 C</b>	<b>88 G</b>	<b>105 T</b>
1	T T T T T T T T T T	T T G G C A G C A A	A A C A T G T A T T	T T A A C G A C T C	T G A G T C A C A	G C A C A T T G G A
61	A A G G T A A A T N	A C C C T C T T C C	C T C T G A T G A G	T T C T G A A A G C	T A G G A C A G C A	A A G C T T C C C T
121	A A C A G G A C A G	A C G G G A A G T C	C A G G A C C C C A	C A C A A G T T C T	A G A A G G G T G T	G G A C C A A C T
181	G A C G T T G A A G	T A G A T G T A G A	T C T A T T T C A A	A G T A T T T C T G	A C T A G G G T C T	A T G T T A A T A A
241	A C A G G G C T G A	T A G G A T T T A T	A G G T C A A G A G	A A A T G G A G T C	T A A A A G A T G	A A A A C T A A A C
301	T A C A G G T C G A	G C A T C C C A A A	T C C A A A A A T C	C A A A A T C T G A	A G C G C T C C A A	A A T C C A A A A C
361	T T T T T G A G C A	C T T A C A T G G C	A C T C A A G G G A	A C G G G C A T G G	G G T C A T T T C A	G A T T T
<b>C06 序列</b>	<b>BASE</b>	<b>COUNT</b>	<b>90 A</b>	<b>88 C</b>	<b>68 G</b>	<b>165 T</b>
1	A C A A A G C T G G	A T G C T A T T A T	T T T T C A A C A C	A T C T G A G A G A	T T G T T C T T T A	A T C T A T T T A C
61	C A A G A A G A C T	G G G G A T T T T T	T T T G T C C C C T	C A C C T T A C A T	T T T G G T T C G A	T T T T T T T T T C
121	T T C A C T C T A C	T T T T A T C C T C	T A G C G G T T T A	G A A G T T A C G T	A T G T A C A T T T	C C T T C T G C T A
181	T T G G T C A C C T	T G T A T T T T C C	A C A T A T A T A T	T T A A A A T T G A	A T T G T C A C G T	A G T T A A T C A G
241	T G T C T A C A G C	C T C C T T A G C A	G G G G G C C A G A	G A C T T C C T G C	A T T T T T G C T T	T C T T C A T T T C
301	C A T C C T C C A G	A C T T C C C C T C	T T G T T G A A G A	T A T C T G G G A G	T T C A G C T T G C	T T G A C T C T T T
361	T A T T A T A A T C	G A A G A A T T T T	T A G T G G A C C C	C T G A G C T T T C	T A G G G C T A T G	C

恶性疟原虫 DNA 片段部分序列

Partial sequence of obtained DNA fragment of *P. falciparum*

表 1 所获 cDNA 序列与数据库序列同源性比较结果  
Table 1 Results of comparison of homology of obtained cDNA sequences with database sequences

序列编号 Sequence No.	最高同源计分 Highest score	最高同源性 (%) Highest homology	所比序列长度 Compared sequ length	同源序列名称 Significant homology with
C01	418	58.9	326/306	
C02	452	57.1	403/365	
C03	1586	98.3	409/420	PHSP70-2
C04	402	55.4	368/337	
C05	448	55.1	445/415	
C06	456	55.3	434/411	

**cDNA 序列同源性的搜索、比较及分析** 所获表达序列标记位 (EST) 序列用 BLAST 命令在 DNA 序列数据库中进行同源性搜索、比较及分析结果见表 1。其中“最高同源计分”、“最高同源性”为比较搜索出的众多序列中计分最高者,即同源性最强者;“所比序列长度”为计算机实际比较序列长度。

**所获抗原 cDNA 序列 (ESTs) 性质确定** 由上述同源性比较结果可知, C03 序列为已知恶性疟原虫抗原热休克蛋白 70-2 基因片段 (*P. falciparum* antigenic heat shock protein 70-2 mRNA), 其余 5 个抗原 cDNA 片段为未知序列, 即我们所发现的恶性疟原虫 FCC1/HN 株新抗原表达序列标记位 (ESTs)。

## 讨 论

序列标记位 (STS) 正在成为基因组物理图谱分析中的标准标记, 这些来源于基因组克隆且较短的序列, 独自地代表了物理图谱中的特定位置。ESTs 可以与随机测序基因组 DNA 而获得的 STS 达到同样目的, 并且额外指出其是表达基因。有关恶性疟原虫基因组编码基因的数量, 估计约为 3 500 或 5 000 个, 基因组大小为  $3 \times 10^7$  bp, 平均转录长度约 2 kb, 而恶性疟原虫基因组 DNA 编码序列几乎占去其 40%, 比起其它物种基因组 DNA 通常只有 2% 左右转录而言, 就显得很高。因此在恶性疟原虫基因组研究中, 发掘 ESTs 的意义尤为突出<sup>[3]</sup>。

恶性疟原虫具有复杂的生活史, 而通常构建的 cDNA 文库均是针对红内期而言, 因为红内期恶性

疟原虫的培养较为容易, 也就是说, 所发掘的 ESTs 均局限于红内期的编码基因, 其它期的序列标记位可以通过构建基因组文库来获得<sup>[5]</sup>。另外, EST 往往仅是基因片段, 以 EST 为线索在基因组文库中筛选全序列也是我们正在进行的工作之一<sup>[6]</sup>。

值得强调的是, 我们所分析的 cDNA 克隆均是经过免疫学筛选所得, 大多具有与之相对应的单抗, 无疑对以后与 cDNA 序列相关的编码蛋白研究提供了方便。当然, 要进一步鉴定这些 ESTs 的存在还有多方面工作要做, 包括对应基因全序列的测定及表达、染色体图谱分析和定位以及组织分布及免疫学特性等。

## 参 考 文 献

- 1 Gavaghan H. Tunisian institute to tackle secrets of malaria genome Nature 1994; 371 732
- 2 Foster J, Thompson J. The *Plasmodium falciparum* Genome Project: A Resource for Researchers Parasitol Today 1995; 11 1-5
- 3 Chakrabarti D, Reddy GR, Dame JB, et al. Analysis of expressed sequence tags from *Plasmodium falciparum*. Mol Biochem Parasitol 1994; 66(1) 97-104
- 4 王燕妮, 谢毅, 肖谷田, 等. 免疫筛选恶性疟原虫 cDNA 克隆. 中国寄生虫学与寄生虫病杂志 1997; 15 193-197
- 5 Reddy GR, Chakrabarti D, Schuster SM, et al. Gene sequence tags from *Plasmodium falciparum* genomic DNA fragments prepared by the genease activity ofmung bean nuclease Proc Natl Acad Sci USA 1993; 90 9867-9870
- 6 阎宗合, 谢毅, 毕惠祥, 等. 恶性疟原虫基因组文库的构建及新抗原 cDNA 片段的筛选. 第一军医大学学报 1997; 17(2) 94-97  
1999 年 3 月 15 日收稿 1999 年 7 月 16 日修回  
(编辑: 富秀兰)

## FINDING OF NEW FCC1/HN ANTIGENIC EXPRESSED SEQUENCE TAG(ESTs) OF PLASMODIUM FALCIPARUM

YAN Zhonghe<sup>1</sup>, XIE Yi<sup>2</sup>, LIM ing<sup>1</sup>, WANG Ping<sup>1</sup>, WANG Yanni<sup>1</sup>,  
BI Huixiang<sup>1</sup>, LI Yingjie<sup>1</sup>

<sup>1</sup> Department of Tropical Diseases, Institute of Tropical Medicine, The First Military Medical University, Guangzhou, 510515

<sup>2</sup> State Key Laboratory of Genetic Engineering, Fudan University, Shanghai, 200433

### ABSTRACT

**AM:** To sequence the strong positive clones obtained by immuno-screening of *Plasmodium falciparum* FCC1/HN  $\lambda$ gt11 cDNA expression library, and to elucidate the antigenic expressed sequence tags through sequencing the cDNA insert of these positive clones, and new antigenic ESTs could serve as a resource to pursue their corresponding antigen genes. **METHODS:** cDNA inserts of positive  $\lambda$ gt11 phage clones were amplified by PCR. The PCR products, after purification, were cloned into the M13mp18 sequencing vector. Single-stranded M13 DNA was prepared and sequenced. Then the acquired sequences were compared in homologies with EMBL/GenBank database on the PC/GENE software system and searched in NCBI (National Center for Biotechnology Information) GenBank using BLAST (Basic Local Alignment Search Tool) command. **RESULTS:** Sequence C03 was part of the known *P. falciparum* antigenic heat shock protein 70 (Pfhsp70) gene, while the other 5 sequences were new *P. falciparum* antigenic expressed sequence tags (ESTs). **CONCLUSION:** The 5 new antigenic ESTs generated could serve as the breaking through points in our efforts to find out new *P. falciparum* antigen genes.

**Key words:** *Plasmodium falciparum*, antigen, cDNA, expressed sequence tags

## 上海市寄生虫学会1999年会

上海市寄生虫学会1999年会于12月17日至18日在上海市青浦区召开。年会主题是：上海市寄生虫病流行现状分析；寄生虫防治经验和防治研究进展。中国预防医学科学院寄生虫病研究所、上海医科大学、上海第二医科大学、第二军医大学、上海中医药大学、华山医院、上海市疾病预防控制中心及青浦、嘉定、浦东新区、金山、奉贤等区县卫生防疫站等共14个单位40余位科研人员和疾病防治第一线的卫生工作人员出席了会议，学会理事长冯正研究员主持会议。

会议邀请许隆祺、薛纯良、蔡黎和石尧忠四位教授分别作了题为“我国食源性寄生虫病流行态势概述”、“孕期弓形虫感染的诊断和处理”、“上海市寄生虫病流行现状与展望”和“寄生虫病10例报告”四篇学术报告，大会共交流学术论文和工作报告26篇。来自各有关院校、上海市疾病预防控制中心和各区县卫生防疫部门的会员分别作了“蠕虫感染对儿童智商的影响”、“寄生虫感染与卫生习惯的关系”、“青浦区寄生虫病防治和监测工作报告”、“金山区人群寄生虫感染十年消长分析”、“嘉定区人体肠道线虫感染十年变化的研究”、“上海市流动人口弓形虫感染的血清学监测分析”、“唾液作为日本血吸虫病检测标本的实验和引用研究”等工作报告，充分反映了我国及上海市近十年来寄生虫病流行状况及防治研究成果。对于近年来出现的新情况、新变化以及有关的寄生虫病诊断方法也进行了交流和讨论。

随着上海市经济建设的迅速发展，人民生活水平的不断提高，寄生虫病防治科研人员多年来坚持不懈地努力工作，使原有的几种主要寄生虫病（如血吸虫病、疟疾、丝虫病）的防治和研究取得了巨大成就，达到了基本消灭的标准。同时，随着人们生活习惯及饮食方式的变化，城市建设和旅游业的

发展、流动人口急剧上升，食源性寄生虫病（如肺吸虫病、囊尾蚴病、姜片吸虫病等）、免疫缺陷和免疫力低下人群机会性寄生虫感染（如弓形虫、卡氏肺囊虫和隐孢子虫等）和输入性寄生虫病等有不断升高的趋势。对此，与会者认为应该引起寄生虫病防治研究及有关领导部门的足够重视。

本次会议为上海地区寄生虫学科的专家教授和寄生虫病防治人员提供了互相交流互相学习的机会。大家进行了认真热烈的研讨，一致认为，上海作为国际化的大都市，不仅在经济和城市建设方面高速发展，同时在疾病防治研究、提高市民健康水平及生活质量方面也应尽快向国际大都市应有的水平迈进，并学习各省（市、区）的防治经验，为全国寄生虫病防治作出贡献。会议提出，在新世纪来临之际，上海市应加强各科研院所、医科院校和基层防疫部门的合作，积极开展寄生虫病流行状况调查及科研防治工作，通过各种媒体大力开展健康教育宣传工作，提高自我保健意识，加强卫生工作人员教育培训，为上海市有关部门决策献计献策；要继续努力降低上海市土源性寄生虫感染率，做好食源性寄生虫病、机会性寄生虫病和输入性寄生虫病的控制和防治工作，争取在“十五”期间大幅度降低本市寄生虫感染率，进一步减轻寄生虫病的危害，为保护市民健康，提高生活质量作贡献。

会议期间，冯正理事长主持召开了学会第四次理事会议，总结1999年学会工作，讨论2000年工作计划，并向大会作了学会工作报告和大会总结发言。与会者一致认为大会开得十分成功，达到了预期目的，同时大家充满信心地迎接新世纪的到来。

(钟万剑)