

研究论文

## 口蹄疫病毒分型诊断芯片探针设计初报

相 磊<sup>1,2</sup>, 陈小玲<sup>2\*</sup>, 李伍举<sup>4</sup>, 梁之昶<sup>3</sup>, 章振华<sup>2</sup>, 王学文<sup>1</sup>, 胡国良<sup>1</sup>, 徐福洲<sup>2</sup>, 石 岗<sup>2</sup>

(1. 江西农业大学动物科技学院, 江西南昌 330045; 2. 北京市农林科学院畜牧兽医研究所, 北京 100097;

3. 新疆农业大学动物科技学院, 新疆乌鲁木齐 830052; 4. 军事医学科学院基础医学研究所, 北京 100850)

**摘 要:**为建立口蹄疫病毒(Foot-and-mouth disease virus, FMDV)不同血清型与基因型的基因芯片检测方法,设计针对 O 型 8 个基因型、A 型 3 个基因型和亚洲 1 型的特异性探针。从美国 GenBank 与英国世界口蹄疫参考实验室基因库下载了 O 型、A 型和亚洲 1 型 FMDV 的 VP1 基因序列 547 条。对每一血清型序列用 DNA Star 软件 ClustalW 程序进行多重比对,做系统发育分析并进行基因分型。用生物学软件 Bio Sun 2.0 建立基因型数据库,设计每一基因型的特异性探针。共设计出 104 条候选探针,通过芯片试验筛选出 12 条特异性探针。以各型特异性探针所对应的靶序列模板做 10 倍系列稀释进行 PCR 扩增,扩增产物与探针杂交,验证各探针的灵敏度。对 O 型 SEA、Euro-SA、ME-SA、WA 4 个基因型的各条探针的灵敏度进行了检验,结果这些探针能够检测到 10<sup>2</sup> 数量级拷贝数的阳性靶标。

**关键词:**口蹄疫病毒;血清型;基因型;VP1 基因;基因芯片;寡核苷酸探针

中图分类号:S852.659.6

文献标识码:A

文章编号:1007-5038(2008)02-0001-05

基因芯片又称 DNA 微阵列,是专门用于核酸检测的生物芯片,也是近年来应用最广泛的微阵列芯片。它是指在固相载体上按照特定的排列方式固定上大量序列已知的 DNA 片段,形成 DNA 微矩阵,固定于固相载体上的 DNA 片段称为探针。将样品基因组 DNA/RNA 通过 PCR/RT-PCR 扩增等技术掺入标记分子后,与位于微阵列上的已知探针序列杂交,用于杂交的样品称为靶序列,通过激光共聚焦荧光检测系统对芯片进行扫描,检测杂交信号强度,经计算机软件进行数据比较和综合分析后,即可获得样品中大量基因序列特征或基因表达特征的信息<sup>[1]</sup>。

口蹄疫病毒(Foot-and-mouth disease virus, FMDV)可侵染猪、牛、羊等畜种及其他 70 多种家养和野生偶蹄动物,是一种烈性传染病。FMDV 属于小 RNA 病毒科(Picornaviridae)、口蹄疫病毒属(Aphthovirus),含有约 8 500 个核苷酸的单股 RNA 病毒。FMDV 有 7 个不同的血清型,即 A、O、C、Asia1、SAT1、SAT2 和 SAT3 型。划分同一血清型的 FMDV 株普遍采用基因型<sup>[2]</sup>。以 15% 的核酸序列差异率为临界值,对 105 株血清 O 型毒株 VP1 基因 3 端核苷酸序列进行了同源性分析,使用不加权配

对组算法(unweighted pair-group method using arithmetic averages,UPGMA)绘制了系统发育树,经过分析,对 O 型 FMDV 确定了 8 个主要基因型<sup>[3-4]</sup>。由于这些基因型与特定的地域有关,故又称为拓扑型,即中国型(Cathay)、中东-南亚型(ME-SA)、东南亚型(SEA)、欧洲-南美型(Euro-SA)、印尼-1 型(ISA-1)、印尼-2 型(ISA-2)、东非型(EA)和西非型(WA)。A 型病毒株可以分为 3 个基因型,即欧洲南美型(Euro-SA)、亚洲(Asia)和非洲型(Africa)。亚洲 1 型核苷酸序列差异比其他血清型小很多,不能将其分为 1 个以上的基因型<sup>[1]</sup>。本试验旨在设计口蹄疫病毒 O 型内 8 个基因型、A 型内 3 个基因型和亚洲 1 型的特异性探针,为口蹄疫病毒分型诊断芯片的研究奠定基础。

### 1 材料与方法

#### 1.1 材料

1.1.1 种毒和样品 O、A、亚洲 1 型 FMDV 灭活细胞培养物各 1 份,由中国农业科学院兰州兽医研究所提供。经赛百盛基因技术公司合成代表 6 个基因型的 6 条核苷酸序列(每条约 100 bp)。

1.1.2 序列来源 从美国 GenBank(<http://www.ncbi.nlm.nih.gov>)与英国世界口蹄疫参考实验室

收稿日期:2007-09-20

基金项目:国家自然科学基金资助项目(30571383);北京市自然科学基金项目(6062009)

作者简介:相 磊(1982-),男,硕士研究生,山东临沂人,主要从事临床兽医学研究。\*通讯作者

(WRLFMD)基因库(<http://www.iah.bbsrc.ac.uk>)下载的547条VP1基因序列。

1.1.3 生物学软件 军事医学科学院开发的生物学软件 Bio Sun 2.0, 生物信息学处理软件 DNA Star, 引物设计软件 Primer Premier 5.0。

1.1.4 仪器与试剂 iCycler PCR 仪为 BIO-Rad 公司产品, GenePix4000B 扫描仪为 Axon 公司产品, PixSys 5000 点样仪为 Cartesian 公司产品, DU640 紫外分光光度计为 Beckman 公司产品, 基因芯片购自 CEL Associates 公司, Cy3 荧光染料购自 Amersham 公司, PCR 试剂、琼脂糖凝胶 DNA 回收试剂盒、普通质粒小提试剂盒、感受态细胞(TOP10)均购自天根生化科技有限公司, Trizol Reagent 购自上海英骏生物技术公司, Takara one step RT-PCR Kit(AMV)和 PMD18-T Vector 购自宝生物工程(大连)有限公司。

## 1.2 方法

### 1.2.1 FMDV 的 VP1 基因区序列系统发育树分型

从美国 GenBank 与英国世界口蹄疫参考实验室基因库下载 FMDV 的 VP1 基因序列 547 条, 其中 O 型 FMDV 210 条, A 型 FMDV 113 条, 亚洲 1 型 FMDV 224 条。对每一个血清型应用 DNA Star 软件 Clustal W 程序进行多重比对, 然后用 Tree View 程序绘制系统发育树, 对所下载序列进行基因型划分<sup>[4-6]</sup>。

1.2.2 FMDV 分型探针设计 应用生物学软件 Bio Sun 2.0 对每个基因型内的序列设计探针<sup>[7]</sup>。将所要进行基因型探针设计的该基因型所有序列转化集成为一个数据库, 导入 Database for Targets 数据库, 将该数据库和除该基因型序列以外的所有其他序列做成另一个数据库, 导入 Database for Targets & non-Targets 数据库, 在相应的对话框内选择探针参数进行探针设计。

通过软件自带的 BLAST 工具搜索该基因型数据库的其他序列, 筛选出能覆盖尽可能多序列的该基因型的通用探针。

1.2.3 探针的筛选 为了减少非特异性杂交, 保证芯片杂交的敏感度和精确度, 探针设计还要遵循以下原则: 探针的  $T_m$  值应该接近整个基因组的平均  $T_m$  值, 上下波动不超过 5; 重复的碱基连续不超

过 6 个; G + C 含量在 40% ~ 60%; BLAST 与其他序列的相似性小于 75%, 若相似性在 75% 以上则可能会出现交叉反应; 探针连续同源的碱基数不应超过 20 个<sup>[3,8-9]</sup>。或在进行探针筛选时, 将同一基因型的探针一起点制芯片, 然后将芯片与靶序列杂交, 从而筛选出特异性最好、杂交信号最强的探针。

1.2.4 病毒阳性质粒的构建 用 Trizol Reagent 提取病毒 RNA, 一步法 RT-PCR 进行目的片段的扩增, PCR 产物经回收纯化、连接、转化、质粒提取和测序, 构建阳性质粒<sup>[10-11]</sup>。

1.2.5 靶序列的扩增 将 Baxi M K 等<sup>[2]</sup>设计的引物修改为, 上游 5'-GCT CGG TAC TAC ACA CAG-3', 下游 5'-GGT TGG ACT CAA TGT CTT-3', 并且在 1.1.1 中合成的核苷酸序列两端也加上该对引物, 扩增片段分别为 1 070 bp 与 100 bp。

PCR 反应体系为, 无菌去离子水 18  $\mu$ L, 10  $\times$  PCR buffer 2.5  $\mu$ L, 2.5 mmol/L dNTP 2  $\mu$ L, 2.5 U/ $\mu$ L Taq DNA 聚合酶 0.5  $\mu$ L, 5  $\mu$ mol 上游引物 0.5  $\mu$ L, 100  $\mu$ mol 下游引物(荧光标记) 0.5  $\mu$ L, 质粒模板 DNA 1  $\mu$ L, 总反应体积 25  $\mu$ L。PCR 程序为 94  $^{\circ}$ C 2 min; 94  $^{\circ}$ C 30 s, 50  $^{\circ}$ C 1 min, 68  $^{\circ}$ C 1 min, 进行 35 个循环; 68  $^{\circ}$ C 5 min, 4  $^{\circ}$ C 保存。

1.2.6 探针的特异性与灵敏度初步验证 将设计出的探针 5' 端氨基酸修饰后合成, 通过点样仪点到醛基化玻片上(每个探针作 2 个重复)。对阳性质粒采用不对称 PCR 进行扩增, PCR 体系中的一条引物采用 Cy3 荧光标记。将 PCR 扩增产物与探针在 42  $^{\circ}$ C 条件下进行杂交, 验证探针的特异性。将阳性质粒作  $10^1 \sim 10^7$  系列稀释, 用稀释后的质粒为模板, PCR 扩增产物与探针杂交验证探针的灵敏度。

## 2 结果

### 2.1 FMDV 基因分型与探针设计

分别应用 DNA Star 软件中的 Clustal W 程序多重比对后, 通过 Tree View 构建系统发育树进行基因分型, 并对每一个基因型用 Bio Sun 2.0 设计特异性探针。对 O 型 8 个基因型共设计出 55 条探针, 对 A 型 3 个基因型共设计出 40 条探针, 亚洲 1 型 FMDV 只有一个基因型, 设计出了 9 条特异性探针。共设计出候选探针 104 条(表 1)。

表 1 O型和A型 FMDV 基因分型序列数和设计出的探针数

Table 1 Type O and A FMDV sequence numbers genotyped and probe numbers designed

血清型 Serumtype	基因型 Genotype	序列数/条 Number of sequences	探针数/条 Number of probes
O	Cathay	74	12
	ME-SA	68	15
	Euro-SA	21	4
	SEA	24	13
	WA	4	3
	EA	4	3
	ISA-	3	3
A	ISA-	2	2
	Asia	35	18
	Euro-SA	24	15
	Africa	20	7

## 2.2 筛选出的 FMDV 探针序列及探针特征

表 2 列出了本试验中筛选出的 12 条基因型探针的序列,这些探针的特性见表 3。

表 2 筛选出的 12 条 FMDV 探针序列

Table 2 Twelve FMDV probe sequences screened

探针编号 Probe	基因型 Genotype	探针序列 Probe sequence
1	Cathay	5 - GCAA GTACGGT GACACCA GCACTAACAACGTGA GA GGC-3
2	ME-SA	5 - T GCAA GTA TGGCA GA GCCCCGT GACCAA TGT-3
3	Euro-SA	5 - TCCAA GGCCTT GCTGGCAA TCCACCCACT GAA GCC-3
4	SEA	5 - CTGGTGGGTGC GCTCCCTCCGTA CTGCCACTTACT-3
5	WA	5 - CCTTCAA GTTCTGCCCCGA GGCGGGCGCCGA TGT-3
6	EA	5 - TCCGCGGCTCTTTTGGCCACCCACCCGA G-3
7	ISA-	5 - AA TGCA GGTACGGTACACACACCACGACCAACGTGA-3
8	ISA-	5 - ACACTGGTA GGA GGGCTC GTGCGGCCGCCACTTA-3
9	Asia	5 - ACAA GTACTCCGCGGCCA GTGA GCGCACACGGGGT-3
10	Euro-SA	5 - ACCCACCAACACGGTCTA GTGGGTACGTTG-3
11	Africa	5 - CGCA GACACCCCA GCACACACTTGTAGGTGCCTTG-3
12	Asia	5 - GCCTTTGCTA GCTCTTGACACCACTCA GGACC GCCG-3

表 3 筛选出的 12 条 FMDV 探针特性

Table 3 The characteristics of 12 FMDV probes screened

探针编号 Probe	基因型 Genotype	探针长度/ bp Probe length	探针 VPI 位置/ bp Probe position in VPI	Tm/	GC/ %	G
1	Cathay	38	401 ~ 438	78.094 6	0.552 63	- 0.08
2	ME-SA	32	402 ~ 433	77.371 88	0.593 75	- 0.48
3	Euro-SA	37	561 ~ 597	81.340 37	0.594 59	- 0.53
4	SEA	34	181 ~ 214	79.688 86	0.617 65	0.63
5	WA	36	441 ~ 476	84.419 85	0.666 67	- 2.72
6	EA	30	561 ~ 590	82.237 47	0.7	- 1.52
7	ISA-	36	401 ~ 436	77.949 19	0.527 78	- 1.07
8	ISA-	35	181 ~ 215	83.444 46	0.657 14	- 1.21
9	Asia	35	401 ~ 435	83.709 61	0.657 14	- 1.21
10	Euro-SA	30	166 ~ 195	75.138 49	0.566 67	- 2.83
11	Africa	35	161 ~ 195	79.271 37	0.6	- 0.24
12	Asia	36	561 ~ 596	79.864 93	0.611 11	- 0.05

将同一基因型的若干条探针与靶序列杂交,筛选出特异性最好,杂交信号最强的探针。以 ME-SA 基因型特异性探针的筛选为例,如图 1 所示选取探针 b 为 ME-SA 型特异性探针。图中 a, b 为 ME-SA 设计出的 2 条型特异性探针, a 为 5'-GCAA GTATG-GCGA GA GCCCCGTGACCAATGTG-3'; b 为 5'-TG-CAA GTATGGCA GA GCCCCGTGACCAATG-3'。

### 2.3 探针的特异性与灵敏度验证

已经验证了部分探针的特异性与灵敏度,以 ME-SA 特异性探针的灵敏度为例(图 2)。

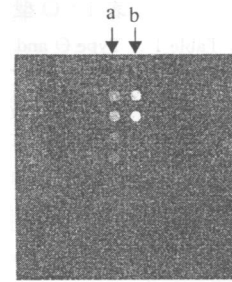
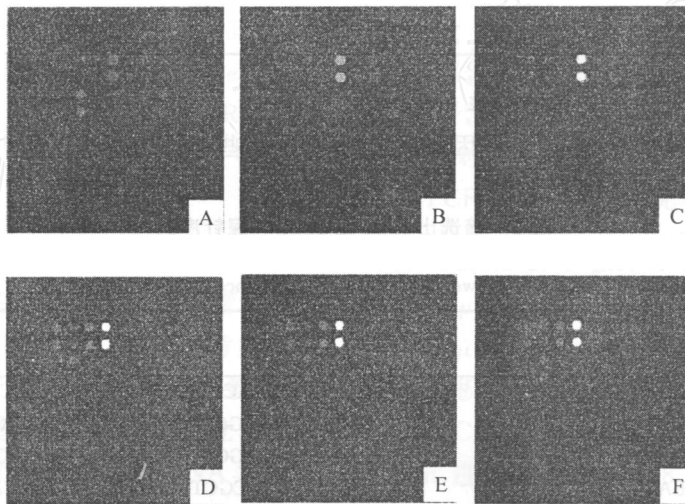


图 1 特异性探针筛选

Fig. 1 Specific probe selection



A. 空白; B. 质粒拷贝数为  $10^2$  数量级; C. 质粒拷贝数为  $10^3$  数量级; D. 质粒拷贝数为  $10^4$  数量级; E. 质粒拷贝数为  $10^5$  数量级; F. 质粒拷贝数为  $10^6$  数量级

A. Blank; B.  $10^2$  plasmid copy numbers; C.  $10^3$  plasmid copy numbers; D.  $10^4$  plasmid copy numbers; E.  $10^5$  plasmid copy numbers; F.  $10^6$  plasmid copy numbers

图 2 ME-SA 特异性探针的灵敏度

Fig. 2 The sensitivity of the ME-SA genotype probe

### 3 讨论

基因芯片技术从产生到现在才不过短短十几年的时间,但它在生命科学各个领域的应用已经展示出了其蓬勃的生命力。它的主要特点是高通量、微型化和自动化。基因芯片上可以高度集成成千上万密集排列的探针,能够在很短的时间内分析大量的生物分子,使人们能够快速准确地获取样品中的生物信息,检测效率是传统检测手段的成百上千倍。因此,以基因芯片为手段的研究将会有巨大的商业价值和良好的发展前景。

FMDV 开放阅读框(ORF)中编码着 4 种结构蛋白,该 4 种结构蛋白的基因区称为 P1 区,依次编码 1A (VP4)、1B (VP2)、1C (VP3) 和 1D (VP1)。FMDV 7 个血清型的结构蛋白氨基酸组成数不同,在 4 种结构蛋白中,差异最大的是 VP1,其次为 VP3 和 VP2<sup>[1]</sup>。根据 FMDV 7 个血清型编码衣壳表面

基因 VP1 的 3 末端序列的系统发育分析,得到了与 7 个血清型相对应的 7 个基因世系,证明了这些基因世系与血清型密切相关,而衣壳编码区以外的基因组任意一端都与血清型无关<sup>[2]</sup>。因此,在进行 FMDV 分型诊断芯片的探针设计过程中,选择用 VP1 基因序列设计探针。

用生物学软件 Bio Sun 2.0 进行探针设计,能够将许多需要进行探针设计的序列组建成数据库,从而针对数据库进行探针设计,这就比单纯对一条序列进行探针设计的软件具有方便、快捷和高通量处理的优点。Bio Sun 2.0 虽然能够对所设计探针的序列根据需要进行组库(如 Cathy 基因型数据库),方便快捷地设计出探针,但是它所设计出的探针只与 Targets 数据库中相应的序列是一一对应的,即特异的。要检验该探针能否探测到其他序列,还要通过软件自带的 BLAST 工具搜索该基因型数

据库的其他序列,筛选出能覆盖尽可能多序列的基因型的通用探针,并尽量减少探针的数目,以便简化试验和降低成本。根据常规的探针设计参数,我们把探针的长度控制在 30 bp ~ 38 bp 之间,探针的  $T_m$  值控制在 75 ~ 85 之间。经过 Bio Sun 2.0 设计出的探针再用其 BLAST 程序对下载的所有口蹄疫病毒序列数据库进行搜索,从理论上排除探针的非特异性杂交。

将设计出来的探针合成后通过点样仪点至醛基化的玻片上,将阳性样品用荧光标记的引物进行 PCR 扩增。PCR 产物与芯片进行杂交,验证探针的特异性和敏感性。已经确定了 12 条探针的基因型特异性,检验了 O 型 SEA、Euro-SA、ME-SA、WA 4 个基因型的各条探针的灵敏度。下一步工作将验证更多探针的特异性与灵敏度,并结合临床样本确定每一条探针的 cut off 值。

#### 参考文献:

- [1] 李瑶. 基因芯片与功能基因组[M]. 北京:化学工业出版社, 2004.
- [2] 谢庆阁. 口蹄疫[M]. 北京:中国农业出版社, 2004:16-50.
- [3] Baxi M K, Baxi S, Clavijo A, et al. Microarray-based detection and typing of foot-and-mouth disease virus [J]. Vet Journal, 2006, 172(3): 473-481.
- [4] Samule, A R, Knowles, N J. Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (topotypes) [J]. J General Virol, 2001, 82: 609-621.
- [5] 相磊, 章振华, 胡国良, 等. O 型口蹄疫病毒 VP1 基因区序列系统发育树分型[J]. 动物医学进展, 2007, 28(3): 8-13.
- [6] 乔纳森·佩夫斯纳. 生物信息学与功能基因组学[M]. 孙之荣, 译. 北京:化学工业出版社, 2006: 37-45.
- [7] Li W J, Fan M. M probe: computer-aided probe design for oligonucleotide microarray experiment [J]. Appl Bioinform, 2002, 1(3): 163-166.
- [8] 吕梁, 马文丽, 孙朝辉, 等. 细小病毒 B19 Oligo 探针设计[J]. 生物技术通讯, 2004, 15(1): 17-19.
- [9] Kane M D, Jatko T A, Stumpf C R, et al. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays [J]. Nucleic Acids Research, 2000, 28(22): 4552-4557.
- [10] 相磊, 梁之昶, 章振华, 等. 三株口蹄疫病毒的克隆和基因型鉴定[J]. 生物技术通讯, 2007, 18(6): 935-938.
- [11] 萨姆布鲁克 J, 拉塞尔 D W. 分子克隆实验指南[M]. 3 版. 黄培堂, 等, 译. 北京:科学出版社, 2002.

## Preliminary Study on Microarray Probe Designing for FMDV Genotyping

XIANG Lei<sup>1,2</sup>, CHEN Xiao-ling<sup>2</sup>, LI Wu-ju<sup>4</sup>, LIANG Zhi-chang<sup>3</sup>,

ZHANG Zhen-hua<sup>2</sup>, WANG Xue-wen<sup>1</sup>, HU Guo-liang<sup>1</sup>, XU Fu-zhou<sup>2</sup>, SHI Gang<sup>2</sup>

(1. College of Animal Science and Technology, Jiangxi Agricultural University, Nanchang, Jiangxi, 330045, China;

2. Institute of Animal Science and Veterinary Medicine, Beijing Academy of Agriculture and Forestry, Beijing, 100097, China;

3. College of Animal Science and Technology, Xinjiang Agricultural University, Urumqi, Xinjiang, 830052, China;

4. Institute of Basic Medical Sciences, Academy of Military Medical Sciences, Beijing, 100850, China)

**Abstract:** To develop a microarray for FMDV genotyping and serotyping, genotype specific oligonucleotide probes for eight genotypes in serotype O FMDV, three genotypes in serotype A and Asia FMDV were generated. A total of 547 complete VP1 gene sequences of serotype O, A and Asia I were downloaded from NCBI and WRLFMD GenBanks. The genotypes of the sequences were determined by Megalignment comparison and phylogenetic analysis of the VP1 gene using DNA Star ClustalW program and TreeView program. BioSun 2.0 software was used for genotype based database setting and oligonucleotide probe designing within each targeted genotype. A total of 104 candidate probes were designed and 12 of them were confirmed to be genotype specific by microarray testing. The PCR amplified target samples from cloned templates and their 10 time serial dilutions of the target samples were hybridized with the candidate probes to determine their sensitivity. The sensitivity test with 1 probe in each of 4 genotypes in type O FMDV showed that the probes were able to detect minimum  $10^2$  copies of positive targets.

**Key words:** FMDV; serotype; genotype; VP1 gene; microarray; oligonucleotide probe