

冰核细菌 *Erwinia ananas* 110 冰核基因克隆和序列分析*

唐朝荣 孙福在 赵廷昌 李瑞峰**

(中国农业科学院植物保护研究所 北京 100094)

摘要: 从作者自行分离的冰核细菌 (*Erwinia ananas* 110) 中克隆到我国第 1 个细菌冰核基因, 并完成其序列测定和分析。所克隆基因编码区全长 3921bp, 编码 1306aa, 氨基酸序列明显分为 3 个区即 N-端 (161aa)、C-端 (41aa) 的单一序列区和中部的高度重复序列 R 区 (1104aa), 以 16 氨基酸为重复单元的 R 区占整个编码序列的 84.5%。序列分析表明我们所克隆的基因为一个新冰核基因, 将其命名为 *iceA*, 该基因已在 GenBank 上登录, 登录号为: AF387802。

关键词: 冰核基因克隆, 序列测定和分析, 新冰核基因 *iceA*

中图分类号: Q93 **文献标识码:** A **文章编号:** 0253-2654 (2002) 04-0045-07

* 国家自然科学基金资助项目 (No. 30170624)

Project Granted by Chinese National Natural Science Fund (No. 30170624)

青年 863 资助项目

** 东北农业大学生命科学学院 97 级实习生

收稿日期: 2001-06-07, 修回日期: 2001-08-15

CLONING AND CHARACTERIZATION OF A NEW ICE NUCLEATION ACTIVE GENE FROM ICE NUCLEATION ACTIVE BACTERIUM *ERWINIA ANANAS* 110

TANG Chao-Rong SUN Fu-Zai ZHAO Ting-Chang LI Rui-Feng*

(Institute of Plant Protection, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100094)

Abstract: Using the chromosomal DNA of an ice nucleation active bacterium *Erwinia ananas* 110 as template, an ice nucleation active (*ina*) gene was amplified by PCR with *Taq* plus DNA polymerase. After sequencing and compared with reported *ina* genes, the cloned gene was identified as a new *ina* gene and was registered in GenBank at the accession number of AF387802. The new *ina* gene, named as *iceA*, has 3921 bp for its coding region, which encodes 1306 amino acids consisting of repetitive segment (R-domain, 1104aa), which is flanked by N- and C-terminal sequences, with 161aa and 41aa, respectively.

Key words: Cloning of an ice nucleation active (*ina*) gene, Sequencing and Characterization, New *ina* gene *iceA*

冰核细菌具有降低水过冷却能力, 促进过冷却水结冰的独特特性(即冰核活性), 现已成为一种重要的生物资源应用于人工造雪、制冰、食品冷冻浓缩、促冻杀虫等领域^[1]。细菌冰核基因赋予其冰核活性表型, 国外已克隆7个细菌冰核基因^[1], 细菌冰核基因的应用研究现已成为生物冰核领域的研究热点, 涉及多个领域: 构建细菌细胞表面展示系统(Bacterial Cell Surface Display)^[2~4]、用作报告基因^[5~6]、促冻杀虫^[7]和病原微生物高敏检测^[8]等, 显示良好应用前景。

我国从1986年起由中国农科院植保所孙福在研究员率先开展冰核细菌研究, 取得不少成果^[9], 但在冰核细菌分子生物学研究方面同国外还有较大差距, 迄今为止还未得到细菌冰核基因, 制约了我国生物冰核学科的发展。本文报道了我国第一个细菌冰核基因的克隆和序列分析, 该研究将会加速细菌冰核基因在我国的研究应用, 为我国生物冰核学科的发展提供了新的生长点。

1 材料与方法

1.1 材料

1.1.1 菌株和质粒: 实验所用菌株和质粒见表1。

1.1.2 培养基: LB培养基: 蛋白胨 10g, 酵母浸汁 5g, NaCl 10g, 10mol/L NaOH 0.1mL, 蒸馏水定容 1000mL。

KB培养基: 蛋白胨 20g, 甘油 10g, MgSO₄ 1.5g, K₂HPO₄ 1.5g, 蒸馏水定容 1000mL, pH值调至 7.2。固体 LB 和 KB 培养基: 在相应液体培养基中加 1.4% ~ 1.6% 琼脂。

1.1.3 抗生素: 氨苄青霉素, 购自上海生工生物工程公司。

1.1.4 各种工具酶及其它药品: 各种限制性内切酶、*Taq* plus DNA polymerase、T4 DNA ligase, Exonuclease III、S1 Nuclease、DNA pol I Klenow、X-gal、IPTG、dNTP 等分别购自上海生工、华美、Promega、MBI、Sigma、Pharmacia 公司, PCR 产物回收试剂盒和 pMD18-T 载体购自 TaKaRa Biotech 公司。

表 1 细菌菌株和质粒

菌株/质粒	相关特征	用 途	来 源
菌株 <i>Escherichia coli</i> JM109	<i>recA</i> , <i>endA1</i> , <i>hsdR17</i> <i>lacZ</i> Δ M15	质粒转化宿主菌	中国科学院微生物 研究所孙卉提供
<i>Erwinia ananas</i> = <i>Paratubercula ananas</i>			
94	Ice Nucleation Active (INA ⁺)	染色体 DNA 用做 PCR 扩增	孙福在 ^[10]
110	INA ⁺	染色体 DNA 用做 PCR 扩增	孙福在 ^[10]
ICMP 1850	INA ⁺	染色体 DNA 用做 PCR 扩增	ICMP 模式菌株
M232A	INA ⁺	染色体 DNA 用做 PCR 扩增	USA 引进, 为防霜 生防菌
质粒 pMD18-T	pUC18 的衍生质粒, 为一种 T 载体	克隆 PCR 产物	TaKaRa Biotech 公司

1.2 方法

1.2.1 细菌冰核基因的 PCR 扩增: 以 *Erwinia ananas* 94、110、1850 和 M232A 染色体 DNA 为模板, 根据国外已克隆冰核基因^[11-12]设计引物, 引物由上海生工生物工程公司合成。用高保真 *Taq* plusI DNA polymerase 进行 PCR 扩增冰核基因。

Primer1 (+) (正向引物): 5' CCCTAATGAAAT GAAGG TTTAGACATG3' 27nt

Primer1 (-) (反向引物): 5' CCTCTGTTATGCGGATTATTCTTCGGG3' 27nt

PCR 扩增条件: 95℃ 预变性 5min; 94℃ 40sec、51℃ 1min、72℃ 5min, 25 个循环; 72℃ 延伸 10min。

1.2.2 PCR 产物的回收、连接、转化和阳性克隆鉴定: PCR 扩增产物的回收、连接采用 TaKaRa 回收试剂盒和 T-载体使用说明书进行。连接产物转化 JM109 感受态细胞, 通过蓝白斑筛选, 挑取复印白斑, 用 Vali 小滴冻结法测定冰核活性^[13], 并提取质粒进行酶切鉴定。

1.2.3 克隆基因的序列测定: 采用 Sanger 双脱氧链终止法对经酶切鉴定和冰核活性测定为阳性的重组质粒进行序列测定。测序由上海博亚生物技术有限公司完成, 采用 M13-R 和 M13-F 从克隆基因的两头测序, 在完成基因两端各 3 个反应后, 因基因中间为高度重复序列, 公司无法设计引物对剩余部分 (约 600bp) 进行序列测定, 作者采用 Exonuclease III (Exo III) 处理重组质粒生成嵌套缺失突变体后进行定向测序, 实验操作参照《分子克隆实验指南》(第二版)^[14]和 MBI 的 Exo III 使用说明书进行。

1.2.4 序列分析: 采用 DNASIS 5.0 和 Wdnasis 软件进行有关序列分析。

2 结果与分析

2.1 细菌冰核基因的 PCR 扩增、回收、连接和转化

以不同菌株染色体 DNA 为模板, 用 primer1 (+) 和 primer1 (-) 为引物, PCR 扩增结果见图 1。3 个冰核活性菌株都扩增出约 4kb 的 1 条特异性条带, 与国外发表的同源冰核基因大小一致^[11-12]。此外, 从没有冰核活性的 M232A 菌株上也扩增到 4kb 的特异性条带, 说明该菌株具有冰核基因同源基因, 只是该基因已丧失正常功能而成为假

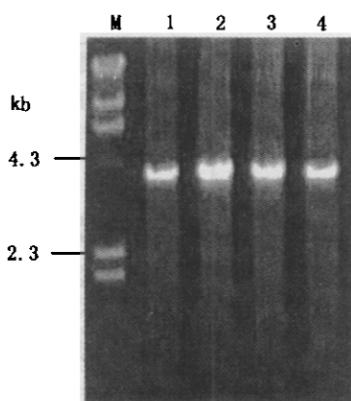


图1 冰核基因的 PCR 扩增

M λ DNA/*Hind*III, 1 菌株 M232A,
2 菌株 110, 3 菌株 94, 4 菌株 1850

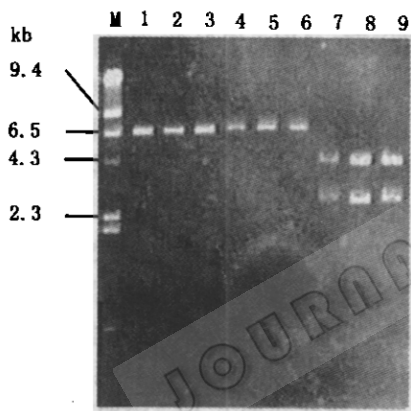


图2 不同转化子重组质粒的酶切鉴定

M λ DNA/*Hind*III, 1~3 1~3 号转化子/*Eco*RI,
4~6, 1~3 号转化子/*Sal*I,
7~9, 1~3 号转化子/*Eco*RI + *Sal*I

基因。实验所采用的菌株 110 为我们自行分离的冰核细菌菌株,冰核活性强且稳定,此后的实验仅利用该菌株的扩增产物。

用 TaKaRa PCR 回收试剂盒回收用菌株 110 扩增的特异性条带,经与 pMD18-T 载体通过 TA 互补连接后转化 *E. coli* JM109,蓝白斑筛选到大量白斑。

2.2 转化子冰核活性测定及重组质粒鉴定

在引物设计时,作者在 primer1 (+) 中包含了基因的起始密码子 ATG (下划线标出) 和其前的 SD 序列 (见方框内序列), primer1 (-) 中包含了基因的终止密码子 TTA (下划线标出,非编码链序列),因此扩增产物包含基因的完整读码框和 SD 序列,如果扩增基因正向插入 (实际连接时有 2 种可能,正向、反向各 50% 几率) 到载体 pMD18-T *lacZ* 启动子下游时,所扩增的基因应能表达冰核活性。实验中随机挑取 24 个转化子 (白斑), -4℃ 测定显示其中 13 个转化子有冰核活性 (作者注:冰核菌在 -5℃ 以上具冰核活性,而非冰核菌在 -10℃ 以下才能催化水结冰),占总转化子的 54%,符合预期结果。随机挑取 3 个转化子 (2 个 INA⁺, 1 个 INA⁻),利用载体上插入片段两端多克隆位点上的 2 个限制性内切酶位点 (*Eco*RI 和 *Sal*I) 对这些转化子的重组质粒进行酶切鉴定,结果见图 2。单酶切均产生 1 个约 6.7kb (载体和扩增产物长度之和) 的片段,双酶切均产生约 2.6kb (载体) 和 4kb (扩增产物) 的 2 个片段。冰核活性测定和酶切鉴定表明已得到插有目的片段 (冰核基因) 的阳性克隆,将 3 号转化子所带的重组质粒命名为 pINA105。

2.3 克隆基因的序列测定

博亚公司完成基因 N-端 1.5kb 和 C-端 1.9kb 的序列测定,因该基因序列中间为高度重复序列 (详见后面的序列分析),公司无法设计引物完成剩余部分约 600bp 的序列测定,作者采用核酸外切酶 III (*Exo*III) 对重组质粒 pINA105 定向切割,对剩余部分序列进行定向测序。这里需要指出的是,对上步实验中所挑取的转化子 (白斑) 测序结果表明,它们都携带插有所扩增冰核基因的重组质粒,但具冰核活性的转化子其重组质粒 (包括 pINA105) 的冰核基因均反向插入在载体 *lacZ* 启动子下游,说明冰核基因表达没有利用 *lacZ* 启动子,但在靠近冰核基因上游的载体上并未报道有其它启动子,同时,在无冰核活性的转化子中外源片段则正向插入在 *lacZ* 启动子的下游,造成这种现象的原因还有待于进一步研究。

酶切鉴定表明 *Bam*HI 和 *Sac*I 仅在重组质粒 pINA105 插入基因下游的多克隆位点上有切点,而在基因内部无切点, *Sac*I 位点比 *Bam*HI 位点更远离插入基因, *Sac*I 消化后

产生 4 核苷酸的 3' 突出端、*Bam*HI 产生 3' 凹端, *pINA105* 经 *Bam*HI 和 *Sac*I 完全消化后, *Exo*III 就可向基因内部纵深单向切割, 而对 3' 凹端则不起作用。研究表明 *Exo*III 的外切酶反应速率十分同步, 单位时间切除的核苷酸长度基本一致, *pINA105* 经 *Exo*III 切割不同时间后取样, 再经 *S1* 核酸酶作用后电泳检测片段大小, 合并可产生 4.5 ~ 5.0kb 片段 (可涵盖未测出的区域) 的样品, 经 DNA polymerase I Klenow 作用、T4 DNA 连接酶连接后转化 *E. coli* JM109, 随机挑取 50 个转化子提取质粒并用 *Eco*RI 和 *Hind*III (距克隆基因 C-端约 0.4kb 外有 1 个切点, 在经 *Exo*III 作用后该位点已被切除) 双酶切鉴定筛选符合要求的缺失突变体, 其中 6 个重组子携带的重组质粒可产生约 2.6kb (载体) 和 1.6 ~ 2.3Kb (基因经 *Exo*III 切割后剩余片段) 2 个片段 (图 3), 以这 6 个重组质粒为模板, 用 Primer1 (+) 和 M13R (-48) 为引物进行 PCR 扩增鉴定 (图 4), 除 5 号重组子未得到特异性扩增片段外, 其余 5 个重组子均产生预期大小的扩增片段。对经酶切鉴定和 PCR 扩增鉴定均符合要求的 5 个重组子送博亚公司测序, 最终完成所克隆冰核基因的完整序列测定。

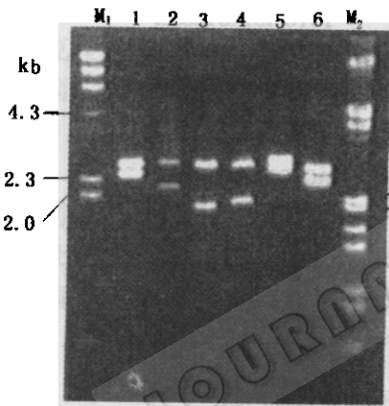


图 3 缺失突变体的酶切鉴定

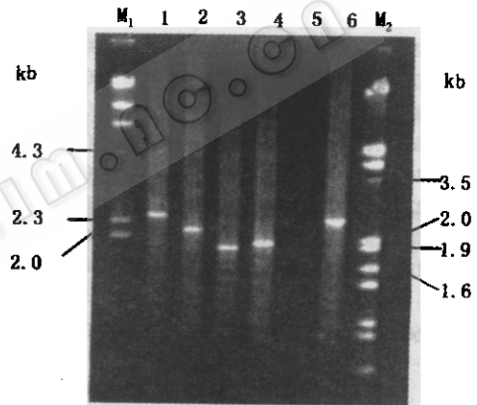
M1 DNA/*Hind*III, M2 λ DNA/*Eco*RI + *Hind*III1~6 不同缺失突变体/*Eco*RI + *Hind*III

图 4 缺失突变体的 PCR 鉴定

M1 DNA/*Hind*III, M2 λ DNA/*Eco*RI + *Hind*III1~6 不同缺失突变体/*Eco*RI + *Hind*III

所克隆基因编码区全长 3921bp, 由此推导的氨基酸序列全长 1306aa, 将所克隆的冰核基因同国外所报道的冰核基因进行同源性比较, 结果表明该基因同国外报道的冰核基因 *inaA*^[14] 的同源性最高, 核苷酸和氨基酸的同源性均为 94%, 二者最大的差别是作者所克隆的冰核基因比 *inaA* 少 1 个 16 氨基酸重复单元 (详见后)。所克隆冰核基因的氨基酸序列可明显分为 3 个区即 N-端 (161aa)、C-端 (41aa) 的单一序列区和中部的高度重复序列 R 区 (1104aa), 以 16 氨基酸为重复单元的 R 区占整个编码序列的 84.5% (图 5)。

序列比较表明所克隆的基因为一个新冰核基因, 基因序列已在 GenBank 上登录, 登录号为: AF387802。作者将所克隆的基因命名为 *iceA*。

1	MKEDKVLILRTCANNMADHGGIHWPLSGIVECKYWKPVKGFENGLTGLIWGKGSDSLSL	60		
61	HADARWVVAEVDADAECIAIETHGWIKFPAEVLHVCTKTSAMQFILHHRADYVASTEMQA	120		
121	GPGAPDVTSEVKAGNRSPLVTDIDATIESGSTQPTQTIEI	161		
162	ATYGSTLSCITHQSQLI	AGYGSTETAGDSSILI	AGYGSTGTAGSDSTLV	209
210	AGYGSTQTAGEESSQM	AGYGSTQTGMKCSDLT	AGYGSTGTAGADSSLI	257
258	AGYGSTQTAGEDSSLT	AGYGSTQTAQKCSDLT	AGYGSTGTAGADSSLI	305
306	AGYGSTQTAGEESTQT	AGYGSTQTAQKCSDLT	AGYGSTGTAGDDSSLI	353
354	AGYGSTQTAGEDSSLT	AGYGSTQTAQKCSDLT	AGYGSTGTAGADSSLI	401
402	AGYGSTQTAGEESTQT	AGYGSTQTAQKCSDLT	AGYGSTGTAGDDSSLI	449
450	AGYGSTQTAGEDSSLT	AGYGSTQTAQKCSDLT	AGYGSTSTAGYESSLI	497
498	SGYGSTQTACYGSTILT	AGYGSTQTAQNESDLI	TCYGSTSTAGANSSLI	545
546	AGYGSTQTASYNSVLT	AGYGSTQTAREGSDLT	AGYGSTGTAGSDSSII	593
594	AGYGSTQTASYHSSLT	AGYGSTQTAREQSVLT	TCYGSTSTAGADSSLI	641
642	AGYGSTQTAGYNSILT	AGYGSTQTAQEGSDLT	AGYGSTSTAGADSSLI	689
690	AGYGSTQTASYHSSLT	AGYGSTQTAREQSVLT	TCYGSTSTAGADSSLI	737
738	AGYGSTQTAGYNSILT	AGYGSTQTAQERSDLT	AGYGSTSTAGADSSLI	785
786	AGYGSTQTACYHSILT	AGYGSTQTAQERSDLT	TCYGSTSTAGSDSSLI	833
834	AGYGSTQTAGYNSILT	AGYGSTQTAQENSDDL	TCYGSTSTAGYDSSLT	881
882	AGYGSTQTACYHSILT	AGYGSTQTAQERSDLT	TCYGSTSTAGPDSSLI	929
930	AGYGSTQTAGYNSILT	AGYGSTQTAQENSDDL	TCYGSTSTQCPDSSLI	977
978	AGYGSTQTASFKSTLM	AGYGSSQAAREQSSLT	AGYGSTSMAGYDSSLI	1025
1026	AGYGSTQTACYQSILT	AGYGSTQTAEHSSILT	AGYGSTATAGADSSLI	1073
1074	AGYGSSLTSGIRSFILT	AGYGSTLISGLRSVLT	AGYGSSLISGRSSILT	1121
1122	AGYGSNQIASHRSSLI	AGPESTQITGNRSMIL	AGKSSQTAGYRSTLI	1169
1170	SGADSVQMAGERGKLI	AGADSTQTAGDRSKLL	AGNNSYLTAGDRSKLT	1217
1218	AGNDCILMACDRSKLT	AGINSILTACGRSKLI	GSNGSTLTAGENSVLI	1265
1266	FRCWDGKRYTNVVAKTGKGIEADMPYQMEDDNNIVNKPEE			1306

图5 冰核基因 (*iceA*) 氨基酸序列的分区排列图

3 结论与讨论

3.1 完成我国第一个细菌冰核基因的克隆

用我们自行分离的冰核细菌 *Erwinia ananas* 110 染色体 DNA 为模板, 根据国外已发表冰核细菌基因序列设计引物, 成功地扩增出约 4kb 的细菌冰核基因。对于 4kb 这样的大片段 PCR 扩增, 国内外有关报道不多。作者认为扩增成功的关键有 2 个, 选择 *Taq* plus DNA 聚合酶和设计合适引物。*Taq* plus DNA 聚合酶由高保真 *Pfu* DNA 聚合酶和 *Taq* DNA 聚合酶混合组成, 在保证扩增产物保真性和 PCR 扩增效率的同时, PCR 产物还可直接用于 TA 互补克隆。最初也尝试用普通的 *Taq* DNA 聚合酶来扩增目标基因, 但无论如何调整反应体系和变换反应条件, 都无法扩增到预期片段。引物设计对长片段 PCR 扩增同样至关重要, 在反复试验的基础上作者设计了理想的特异性引物。此外, DNA 模板浓度、复性温度等调整对于特异性扩增也是必要的。

研究中, 从无冰核活性的 *Erwinia ananas* M232A 中也扩增到约 4kb 的一条特异性条带。根据报道, 几乎所有 *Erwinia ananas* 菌株都是冰核细菌, 冰核活性的测定已成为该种细菌鉴定的一个特征。推测该种细菌所有菌株中都存在冰核基因, 在极少数无冰核活性菌株中该基因仍存在, 只是因发生突变、丧失正常功能而成为一个假基因。实验

中所用的菌株 M232A 是从美国引进, 无冰核活性、用作生物防霜的菌株, 不清楚该菌株冰核活性的丧失是自然突变引起的还是经过人工诱变所致。

3.2 完成所克隆冰核基因的序列测定

对于高度重复序列的测序, 用常规 Walking 法 (即测完一段后重新设计物测邻近序列) 往往不成功。在研究中, 作者采用外切核酸酶 III 生成嵌套缺失突变体法对公司未完成的重复序列进行定向测序。运用该法的关键是在靶 DNA 的一端找到 2 个合适的限制酶切位点, 切点离靶序列近的酶必须产生平端或 3' 凹端, 另一种酶则产生 3 核苷酸或 4 核苷酸的突出端, 2 种酶在重组 DNA 的其它区段上均无切点。作者所选的 2 个酶切位点 (*Bam*HI 和 *Sac*I) 符合这些要求。此外, 单纯用限制酶切分析来选择嵌套缺失突变体有时会发生错选。在本研究中, 作者将酶切鉴定和 PCR 扩增结合起来筛选突变体, PCR 扩增中所用的引物一个是冰核基因 5' 端的特异性引物, 另一个为冰核基因 3' 端载体上的通用引物, 公司测序后表明经双重选择的克隆均符合要求, 为相互嵌套的缺失突变体。

3.3 克隆基因为新基因, 在 GenBank 上登录, 命名为 *iceA*

所克隆的冰核基因编码区全长 3921bp, 编码 1306aa, 氨基酸序列明显分为 3 个区: 单一序列的 N-端 (161aa)、C-端 (41aa) 及以 16 氨基酸为重复单元的高度重复 R-区 (1104aa), 该基因同国外报道的冰核基因 *inaA* 同源性最高, 核苷酸和氨基酸同源性均为 94%。

序列分析表明我们所克隆的基因为 1 个新的冰核基因, 该基因已在 GenBank 上登录, 登录号为: AF387802。至此, 我国获得了第一个细菌冰核基因, 这对加速冰核基因在我国的应用研究及生物冰核学科的发展具有重要意义。

参 考 文 献

- [1] Lee R E, Warren G J, Gusta L V. Biological Ice Nucleation and Its Applications, APS Press, USA. 1995.
- [2] Jung H C. Nat Biotechnol, 1998, 16: 576 ~ 580.
- [3] Bassi A S, Ding D N. Biotechnol Pros, 2000, 16: 557 ~ 563.
- [4] Loper J E, Henkeis M D. Appl Environ Microbiol, 1999, 65 (12): 5357 ~ 5363.
- [5] Kin Y G, Jung H C, Pan J G. Appl Environ Microbiol, 2000, 66 (2): 788 ~ 793.
- [6] Jaeger C H, Lindow S E, Miller W, et al. Appl Environ Microbiol, 1999, 65 (6): 2685 ~ 2690.
- [7] Watanabe K, Abe K, Sata M. Journal of Applied Microbiology, 2000, 88: 90 ~ 97.
- [8] Wolber P K, Green R L. Food Sci Technol, 1990, 1: 80 ~ 82.
- [9] 孙福在. 中国农业科学, 1996, 29 (5): 62 ~ 67.
- [10] 孙福在, 朱 红, 何礼远, 等. 自然科学进展, 1994, 4 (4): 449 ~ 456.
- [11] Warren G, Corotto L. Gene, 1989, 85: 239 ~ 243.
- [12] Abe K, Watabe S, Emori Y, et al. FEBS Lett., 1989, 258: 297 ~ 300.
- [13] Vali G. J Atmos Sci, 1971, 28: 402 ~ 409.
- [14] 金冬雁, 黎孟枫. 分子克隆实验指南 (第二版). 北京: 科学出版社, 2001.