

质粒pXZ10145核苷酸序列测定和分析

沈天翔 贾盘兴 那淑敏 门大鹏

(中国科学院微生物研究所, 北京 100080)

以Sanger双脱氧中止法为原理, 利用美国ABI公司370A自动核酸序列分析仪, 我们测定了谷氨酸棒杆菌质粒pXZ10145全长4887bp的核苷酸序列, 该质粒上存在包括ApaI等18种限制酶的单一识别点, 其它限制酶切点的数目和位置也被定出, 质粒上有8个可能的阅读框架, 通过序列分析, 我们确定了pXZ10145断裂生成缺失突变体pNAT65的两侧位点, 在两个断裂点上发现存在“ATCTAGC”7个碱基的同向重复序列。

关键词 谷氨酸棒杆菌; 质粒; 核苷酸序列测定; 同向重复序列

棒状类细菌内源质粒筛选与鉴定是构建棒杆菌载体系统, 开展这个重要氨基酸生产菌基因工程研究的基础。至今为止, 世界上多家实验室都为之做出了努力^[1], 其中核苷酸序列测定是一项重要内容。1986年, 从乳糖发酵短杆菌中分离到的质粒pAM330核苷酸全序列得到确定, 并分析了其上的遗传信息^[2]。质粒pXZ10145是上海工业微生物所雷肇祖等人于1981年报道的一株谷氨酸棒杆菌小质粒^[3], 它具有作为棒杆菌基因工程质粒载体的一些特征^[4], 这是目前棒杆菌小质粒中唯一发现的抗药基因。为增加对pXZ10145遗传性状的研究, 并构建棒杆菌克隆载体, 我们测定了该质粒的全部核苷酸序列, 利用计算机确定了一部分遗传信息。我们在核酸序列水平上比较质粒pXZ10145和它的缺失突变体pNAT65间关系及缺失发生的位点, 讨论缺失发生的原因。

材料和方法

(一) 菌种、质粒和噬菌体

大肠杆菌HB101, JM101是用于质粒

和噬菌体DNA转化(或转染)的受体菌及重组噬菌体颗粒制备的宿主。谷氨酸棒杆菌1014-6 T带质粒pXZ10145, 是由浙江大学杜珠还教授赠送^[5]。质粒pNAT65是本实验室分离到的pXZ10145缺失突变体^[6]。实验中使用的克隆, 测序载体, 如pBR322, pACYC177, pUC18, M13mp18/19均是本实验室收藏。

(二) 试剂

限制酶、T4连接酶、核酸外切酶Bal31等购于Boehringer Mannheim公司, 所有操作均参照公司说明书。氯化铯、聚乙二醇6000和琼脂糖购于Sigma公司。

(三) 方法

1. 质粒或噬菌体RF DNA提取: 大肠杆菌质粒和M13噬菌体(包括重组噬菌体)RF型DNA提取参照文献[7]。棒杆菌质粒采用改进的碱法提取: 500ml LB液体发酵液离心收集菌体, 并以100 mmol/L NaCl, 10 mmol/L Tris-HCl

本文于1992年4月23日收到。

本项目得到国家自然科学基金和科学院重点课题资助。

本工作完成过程中得到本所技术室王永力、刘振英同志大力协助, 特此致谢。

(pH8.0), 1mmol/L EDTA溶液100ml洗一遍,收集的菌体悬于20 ml原生质体制备液中(50mmol/L Tris-HCl, pH8.0, 100 mmol/L NaCl, 12.5%蔗糖, 10mg/ml溶菌酶), 37°C作用 3 h, 加入 4 ml 10% SDS溶液后, 置于55°C水浴保温 15min, 待温度降下来后加入预冷的 0.2 mol/L NaOH溶液40ml, 反应 5 min后加入60ml醋酸钾, 振荡混合, 放置 30min, 离心后的上清液内加入 0.6 体积异丙醇, 室温下放置30min, 沉淀物悬于 15ml TE缓冲液中, 5 000 rpm 离心 10min, 取上清液加入CsCl做超离心纯化。

2. 转化及转染: 连接反应混合物转化或转染大肠杆菌宿主细菌按文献〔7〕操作。

3. 核酸序列测定: DNA 序列测定依据Sanger 双脱氧中止法〔8〕, 使用美国ABI 公司 370A 自动序列分析仪。测序反应试剂盒购于ABI公司, 其中包括单, 双脱氧核苷酸底物, Taq DNA 聚合酶, 带荧光标记物的M13通用引物(Universal primer “5'TAAAACGACGGCCAGT-3'”)。序列反应使用单链DNA为模板,

重组噬菌体单链提取参照文献〔7〕。测序全过程, 包括聚合酶延伸反应, 聚丙烯酰胺凝胶制备, 电泳及数据收集分析均按ABI公司操作说明书执行。结果借助DNA 序列分析软件进行处理。

结果与讨论

(一) 质粒全序列测定

为完成全部核苷酸序列的测定, 质粒 pXZ10145酶切片段被克隆到M13mp18或 mp19 的多聚接头区中, 构建出 M13-TX 系列重组噬菌体。这些重组噬菌体内的克隆片段的位置及测序方向标于图 1 中。为便于获得所需的 DNA 片段, 我们首先将 pXZ10145 一些较大酶切片段克隆到大肠杆菌多拷贝质粒上, 如pBR322, pACYC 177, pUC18上, 构建出pTS30, pTS60等重组质粒。除以克隆的酶切片段直接测序外, 我们使用了外切酶 Bal31 处理一些亚克隆 DNA 片段, 以使序列测定获得满意的覆盖。

质粒pXZ10145全长为 4887bp, 核苷酸排列顺序见图 2。我们选择了质粒上单

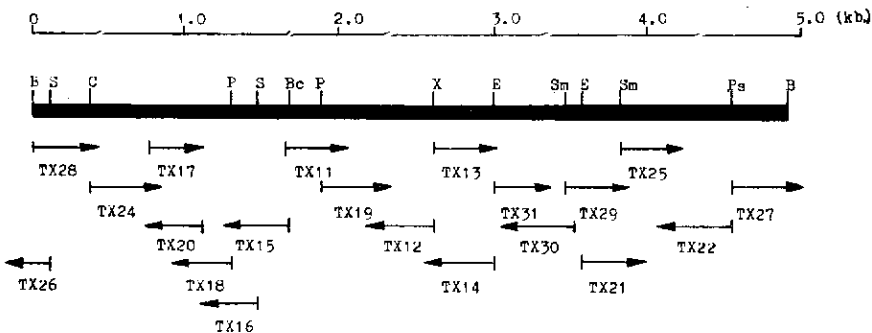


图 1 质粒pXZ10145酶切图谱和测序策略

Fig. 1 Restriction map and sequencing strategy of pXZ10145
The short lines with arrow indication the fragments for sequencing and sequencing direction.
B, BamHI, Bc, BclI, C, ClaI, E, EcoRI, P, PvuII, Ps, PstI, S, SalI, Sm, SmaI, X, XbaI

一的 BamHI 切点为序列编制的起始点。

(二) GC比例

质粒 pXZ10145 GC 比例为59.3%。

由于尚没有该质粒原宿主谷氨酸棒杆菌 1014的 GC 比例数据, 我们查阅了其它谷氨酸棒杆菌的GC比例值, 这些菌株的GC比例都低于 57.5%^[9], 此外, 我们分别测出几个可能的阅读框架的GC比例, 它们

都在 59%—63.6%。而质粒 pAM 330 的 GC 值为53.0%^[2]。

(三) 限制酶位点

多种限制酶的质粒 pXZ10145的切点分布及数目列于表1中, 其中包括 ApaI, BamHI 等18种单切位点, 在其上存在切点的常用限制酶还有 EcoRI, PstI, SmaI, SalI。有26种限制酶在该质粒上

```

GGATCCCAGG CAGGACGCC GAGCATGTCA AGAAGAAGAT CATCAACCTG CGCTGGCGGC TTCGACTGGG GCCAGCCCAG 80
ATCGTCCGCG GACTTGCTCT CTCGACGTGG ACTGTTACAG CGGTCCCTGT CCGTTGCCGC GTGAACCCGC TCTCCGATAT 160
CGATCGTGTG ACTGGCGAGC CATTGGCGGC ATATGAGCAT CCTCATCCGG GATCGTTGAT TCATGTGAT GTCACGTTGT 240
TCGGCAACAT CCCCGACGGC GGTGGACATC GTTACGTAGG TCGGCAGCAA GCGGCACGGA ACAAGCTCGC GACTCCGGGA 320
TTACCACGAG GAAAAGATCA CAAGCCGGCG ACCGGGACGG CGTTCCGTTCA CACAGTCATC GATGACCACT CCCCGCTCGC 400
ATACGCAGAA ATCTGGTCCG ATGAGCAGGC GAGCACAGCG GTGGGAGTTC TCGAACCGCG CGTGGCCTGG TTCGCCGAAC 480
GAGGCGTGAC COTCGAGCGA GTCCATCCG ACAACGGGTC GGCATACAGA TCCCACGCAT GGAGGGACTT CTGCGCTCGG 560
CTCGGCATCC GACACAAGCG GACACGCCCT ACCGGCCGCA GACGAACGGG AAGATCGAGC GATTCCACCG CACGCTCGGG 640
GACGGCTGGG CCTATGCCAG GTTTTACGGT TCAGAGGCCG AACGACGCTG GCGCTGCCCG GCTGGCTCCA CTTCTACAAC 720
CACCCACGAC ACCACTCTGC GATTGGCGGC GTACCCTTCG ACCGACTCAA CAACGTCCCT GGACATCACA TCTAGCCCAA 800
GGCGGATGAG CACCAGGGCG CCGATGACGG CGAGCAGAAG TAGTGGGAGG GGATTGCGTT GAAAAAGTTT CACGGGGTTC 880
TCCTTGTGTT GCGTTTGATT GTCTCGGCTA CACCGTGGCG TCAGAACCGC GTTGTGCAAC CAGATGGCAA ATATCAGTTC 960
CAAGAAATTT CGAGCCGTGA AAAAAAGTGT TCAGTGAGGT GTGCACGTTT TCGGAGCCAA GGCAGACCGC CAGCGGTGTC 1040
CGCGCCAGCC CTAAGAACCG CCGTAGGAGC CACCAGGAGG CTTGAGCCAA CCTCAGGGAG GGGGTAGAGC CCTGAGACGG 1120
TTAGAGGGCT ACCTGAGBGG GCGCGCCAAA GGACGGTTC GCTCGCACCG GAAGGGTCTT TCAGGGGAAC CCCCTGAAAC 1200
CCCGGAAACA TCTGACTTGG TTACGATGGA CTTTGAACAC GCGGAGGAAT GAACCCCGAC CCGCCACACT CATATGAGAA 1280
AGTCAGCTGC ATGAGACCAT CCACCCGCCC TGAGGGACCG TTTGAGCAGC TCTGGCTACC GCTATGGCCA CTAGCAAGCG 1360
ACAACCTCCG TGACGGCATE TACCGCACCT CACCGGATAA AGCCCTAGAC AAGCGCTATA TCGAAACCAA CCCGGATGCC 1440
ATCTCGAATC TCCTGGTCTG CGACATCGAC AAGCAAGACG CAATCTTGGC TTCGATGTGG GACAGAGACG GCTGGCGGCC 1520
TAACGCCGCG GTTGA AAAACC CCTTAAAACG GCACGCACAC GCGCTGTGGG CGCTCGCGGA GCCATTTACC CGCACCGAGT 1600
ACGCCAGGCG CAAGCCACTG CCTACCGCGC CGCTGTCAAC GAAGGACTAC GCGCTCTGT CGACGGCGAC AAAGGCTACT 1680
CGGGGCTGAT CACCAAAAAC CCCGAGCACA CAGCCTGGGA CAGCCACTGG GTAACCGATA AGCTCTACAC ACTCGATGAA 1760
CTGCGCTTTT GCGTGAAGA AACCGGCTTC ATGCCGCCAG AGTCCCTGGAA GAAAAACCGC AGAAAAATCGC CCATCGGGCT 1840
AGGCCGAAAC TGCGCGCTCT TCGAGAGCGC ACGCAGCTGG GCTTATEGCG AAATACGCCA CCACCTTCGGC GACCCCGAAG 1920
GGCTAGGACG CTCTATCCAA GCCACCGCCC AGGCGCTCAA CCAGGAGCTG TTTTCTGAGC CGCTACCGGT GGCAGAAGTT 2000
GACCAAAATG CCAGGTCAAT CCACCGATGG ATCATACCA AATCAGCAT GTGGACAGAC GGCCAGCGC TCTACGAAGC 2080
CACATTCACC ACAATCCAAG CCACACCGCG AAAACGAAGC GCGGAGCACC GCTGGGGCAC CACCAACCGC GAACGAATCG 2160
AAGATTTTAT CAATGACTAA GCGCACTCGT ATTCGCCGAA ACGGTAAAAA TATCCCGGAA GTAGCCGAAG GAACTGGCCT 2240
ATCAACTGGC ACAATCGAAG GCTGGACATC TGCCCTTCGC GAAGATTATC TAGCGCAGGC CAATGAAAAA CCGTCTCCGC 2320
GTCCAGGAAC TGCGTGC AAAAGTCTGTCC ATGCCGCTA TCGCGGGCGA AATCGGTTGC TCAGTCCGGC TGGTTTCATCG 2400
GTACGTCAA GAACTTGAAG AGAAGAAAAC GCGTAAATC CAGCGGTTA GTCCCTTCG GCGTGTTCAT AATAGGTCA 2480
GAACCAAGTC AGATATTCC GGGGTTTACG GCGTCTTTG CTGGCGTGGC AGCTTGGCGC GTTGTGTGCG AGTGAGCGTT 2560
TGTGGGCTG GTTGGCTACG ACAGCGACTA GTTTTCTGGC CCGCTTGACC ATCTCCGGCG AGCTGTCTAG AAGCTCCGGG 2640

```

```

CGTGATTGCA GTAGTCGGGT TGTTCGGTCC ATTTCCGCTGA CAACGGTGGC CAAAATTTCC TCTGGGCTTG CCACCAGTAA 2720
CGAGTTTTGG GTGCTTTCTT GAGTCTGCGC CTGGGCTTGT TTGAGTTCCG CCTGGTGGCC GTTGCCGGTT CGTTCACGAG 2800
CTGCTGAGGC TCTGCACGGC CCGCAGCAGT AGCGGGCTTT GCGCCTCGCG GGTGGCTTCC TGGCGGAATC TCGTTGTGCG 2880
ACCACGCGCA TTTCTGAGAC GTTTCGGCAC CAGTCATACC CCAATTATAA CGTACGTACG TATAAAATTT GCTCTCGATT 2960
TCACCGGAAA ACCGATACCT CGTGGCTCGG TTTCTGACGC GAATTCGGCA CAACTAGTTG GGTTTATTTT CCGTCTACCT 3040
GCGCGCGCAG ACTTTAGTGC ATGAGGTGTG TGGATCTTTT GGGCGAAGCC GAAAACCGAA AGTCTGATTA TCAGCAATAG 3120
GTGTGTACCA ATCGGGCTGA TGTGGTTCGG GCTAGTGGTG TTCATCGTGC GGTTCAGAAT TAGGCGCGAA AAACGGCACG 3200
AGTCCAATCA GGGTGGCAGT TGTGTCAAAG ACGGAATCTA GCTGTGATGT CCAGGGACGT TGTTCTGGCC CCGCGCCGTA 3280
ATGACACCGT GCGGATTCAA GAAATATCTG TCCAGACCGC TTGCCATTGA AGGCATGTTA GAGTTTGGGG CATGTGAGT 3360
CCCGAATCAG ACAGGTCAAG GGTTCGCGGT CACACGGCCA GCCCCTTGAA CCACTAGTTA CGAGGCCAC GATCTGTGTG 3440
GGCGTATGTC TGGCGTACCC GGGCGCTCG CCGTGGTGAC AAGAAGAACC ATTTCTTGAT CTCACACCTC GGAGTACTCG 3520
ATGCCTTTTG CCCTTACAT GCTTCCCTG GCGTCTCTCC TCATGGGAC TTGAGAATTG ATGCTGGCGG GATTGCTCCC 3600
CGCGATCGCG ACCGAACTTG ACGTCTCGGT CCGCACTGGC GGCTGCTGA CCTCCGCAAT CGCAGTGGT ATGGTGGTCC 3680
GCGCGCCAGT GATGGCGGCA TTCGCTCGCC GTTGGCCACC GCGGCTCACA TTGATCGTTT GCCTCTCGT GTTCGGGGGA 3760
AGCCACGTCA TCGGAGCGAT GACACCAAGT TTCTCTCTCC TGCTCATCAC CCGGGTCTC AGCGCTCTCG CAAACGCAGG 3840
ATTCCTCGCC GTAGCACTGA GCACGGCCAC TACCCTCGTG CCAGCGAACC AGAAGGGGGC TGCACCTGTC ATCCTGCTCT 3920
CCGGCACGAC GATCGCAACC GTCGTGGCG TCCCGCCCGG GCACTGCTC GGCACAGCCG TGGGCTGGCG AACGACGTT 4000
TGGGGGATCG CCATCTCTG TATTCCCGCG GCCGTTGGAG TCATTCGTGG CGTCAAGAAC AATGTTGGTC GGAGCGAGAC 4080
TAGCGCGACC TCACCAAGGC TCCGTGTCTA GCTCAAGCCG TTGGCAACGC CCGGCTCATC CTGGCCATGG CACTCGGAGC 4160
GCTGATCAAC GGAGGGACCT TTGCGGCATT CAGCTTCTTG GCACCCATCG TGACCGAGAC CCGGCTCTTG GCCGAAGCGT 4240
GGGTGTCCGT CCGGCTGGTG ATGTTCCGCA TCGGATCGTT CCTTGGCGTC ACGATCGCAG GACGACTATC AGATCAACGA 4320
CCTGGCCCTG TGCTCGCAGT CCGCGCAACG CTATTGCTGA CAGGCTGGAT CGTGTGGCA TGGTTCGCAT CTCATCCCCT 4400
TGCGCTTATC GTCTCTGTC TCGTTCAGGG ATTCCTGTCC TTCGGCGTCC GCAGTACTCT GATCAGCCGT GTGCTGTATG 4480
CAGCATCGGG TGCCGGAACC ATGGGCGGTT CGTACGCAAC CGCAGCATG AATATCGGAG GTGCAGCGGG GCCCGTGTCT 4560
GGTGGCTCGC GGCTCGCGAC CCGGCTGGGG CTGCTCGCGC CGGTTTGGGT CGTTCGGGTG CTGACAGGGA TCGCTCTCGT 4640
CATCATGCTT CTCACCAGAC GCGCGCTTAC GAAGACCGCG CCGGAGGCCA ATTGATGGCC GATCCGAAGC CTCTTCTCAC 4720
TCTCTGTGCC CGTCTCCGCT TAGCTCGGCT GATTGTGAAA GACGGCTATC CCGCCACGAT CGCGCGAAGG ATGTTCAATG 4800
TCTCCCGCAT CACTGCCCGG AAATGGGCAG GCGCTACCG GGAAGAGGCT GAGTTTGGGA TGCAGGATCG CTCCAGCAAG 4880
CCGCACC

```

图 2 质粒pXZ10145全核苷酸序列

Fig. 2 Total DNA sequence of plasmid pXZ10145

没有切点, 这包括 HindIII, BglII, EcoRV。

(四) 阅读框架

图 3 显示出我们经初步分析确定的 8 个可能的阅读框架 (Open reading frames) 的位置和转录方向, 这些阅读框架的启动子和核糖体结合位点 SD 序列等有待于进一步分析定出。需要指出的是, 质粒 pXZ10145 缺失突变体 pNAT65 所对应的区域内 (797—3239) 只有 3 个可能的阅

读框架, 这个缺失突变形成的质粒仍保持自我复制和氯霉素抗性^[8], 因此这两个主要功能对应的蛋白表达产物可能由上述三个框架所编码。

(五) 重复序列

在完成质粒 pXZ10145 全部核苷酸序列测定的同时, 还测定了质粒 pNAT65 的部分区域的核苷酸序列, 特别是其上对应 pXZ10145 发生断裂的部分。实验结果表明: 两个质粒在对应区域内具有完全的

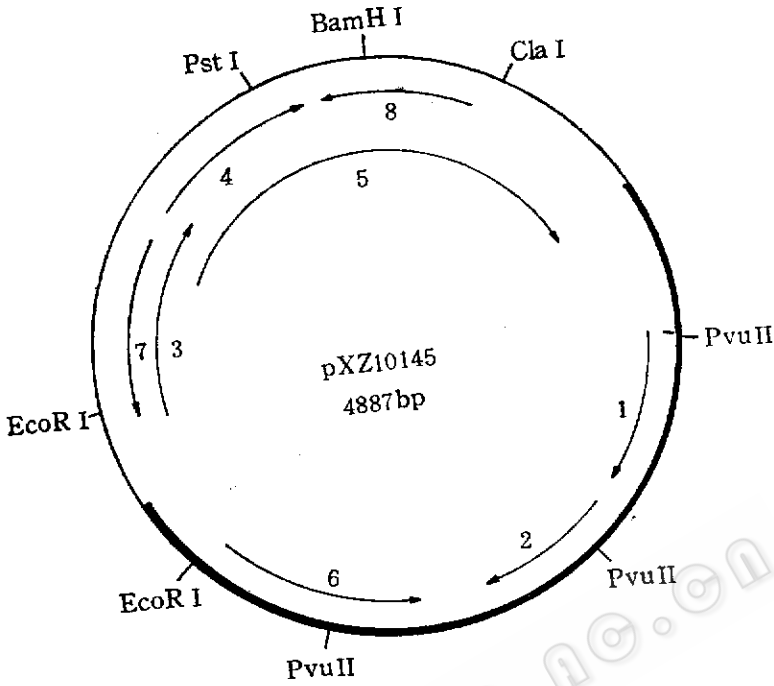


图3 质粒pXZ10145上可能的编码区分布示意图

Fig.3 Location of coding region on plasmid pXZ10145
The region with board line corresponds to that of pNAT65

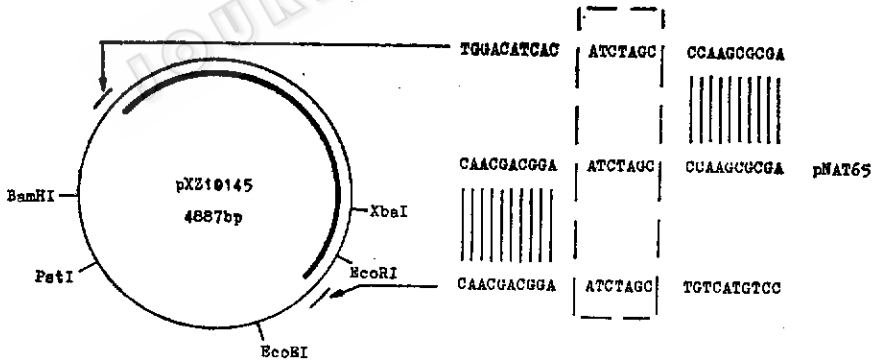


图4 质粒pXZ10145与pNAT65在相应区域中核苷酸序列的比较

Fig.4 Nucleotid sequence comparison between pXZ10145 and pNAT65 in the corresponding region

核苷酸排列顺序。图4给出了质粒 pXZ 10145上左右两个断裂点和 pNAT65 自连点上的核苷酸序列，可见它们都拥有“ATCTAGC”7个碱基的相同序列，由此我们推论，pNAT65产生的原因之一可

能正是与这个碱基同向重复序列有关的分子内重组，即 pXZ 10145 DNA通过这两个点上的同源序列重组，使3243至789或3236至798间的2441bp一段丢失产生新的质粒pNAT65。

表 1 质粒pXZ10145限制酶切点位置及数目

Table 1 Location and number of restriction enzyme cutting sites on plasmid pXZ10145

ApaI	1	4553						RsaI	9	752	1600	2933	2937	3126	3458
BamHI	1	1								3515	4455	4513			
BstEII	1	1730						SfaNI	9	206	574	1385	1444	3529	4277
BstXI	1	1304								4398	4492	4868			
DraIII	1	1563						AcyI	10	15	105	818	1139	3423	3080
MluI	1	4465								3621	4281	4375	4873		
NaeI	1	1141						AlwI	10	9	10	218	537	2037	3080
NcoI	1	4145								3918	4281	4375	4873		
NdeI	1	1272						HgiAI	10	435	812	1005	1708	2128	3819
NotI	1	1627								3863	3905	4113	4334		
PfIMI	1	4694						AsuI	11	69	122	648	2603	3287	3640
PpuMI	1	4175								4175	4345	4549	4550	4697	
PstI	1	4545						HaeII	11	698	821	1142	1416	1573	1956
RsrII	1	4345								2122	3467	3825	3980	4162	
SacI	1	4113						PleI	11	507	998	1100	1808	2388	2749
TthIII	1	94								3208	3365	3826	4046	4130	
XbaI	1	2626						Sau96I	11	69	122	648	2061	3267	3640
XmnI	1	3490								4175	4344	4549	4550	4697	
AatII	2	108	3624					FokI	12	211	217	261	432	579	1311
ApaLI	2	1002	3841							1447	4025	4150	4408	4714	4871
BbeI	2	321	1142					HphI	12	1648	1702	2047	2098	2973	3487
BspMI	2	55	3046							3819	4103	4201	4269	4664	4860
ClaI	2	160	379					EaeI	13	593	1143	1345	1627	2597	2687
DraII	2	4174	4549							3468	3715	3864	4029	4142	4229
EcoRI	2	3002	3548							4771					
NarI	2	818	1139					MaeII	13	105	274	773	1005	2403	2900
ScaI	2	3515	4455							2930	2936	2938	3257	3621	3765
SmaI	2	3460	3808							3994					
XmaI	2	3458	3806					DdeI	14	990	1092	1112	1120	1133	1309
AlwNI	3	1726	1967	2328						1976	2179	2382	2806	2896	3820
AvaII	3	122	4175	4345				MaeI	14	792	1052	1351	1405	1839	1923
BclI	3	1687	4163	4460						2290	2590	2629	3016	3154	3240
BssHI	3	1852	3681	4661						3414	4082				
CelI	3	2378	3819	4112				AlvI	15	305	1286	1329	1742	1976	1987
EagI	3	593	1627	4029						2384	2532	2622	2633	2800	3241
MstII	3	1092	1133	1310						4111	4540	4743			
PvuII	3	1286	1376	2622				BbvI	16	96	296	705	1042	1057	1298
Sall	3	107	1459	1659						1338	1885	2631	2812	2835	3058
SnaBI	3	275	2935	2936						4492	4552	4555	4612		
SpeI	3	2587	3013	3413				MaeIII	16	167	230	270	485	1219	1369
XhoII	3	1	528	3072						1634	1730	2475	2716	3388	3416
XmaIII	3	593	1626	4029						3474	4051	4207	4209		
BalI	4	1347	2689	3715	4144			MboI	16	44	47	622	1190	1788	1820
HaeII	4	1111	4113	4553	4574					1870	2293	2430	2435	3495	3567
Eco47III	4	1414	3823	3978	4160					4683	4724	4770	4854		
Ksp6321	4	1863	2423	4717	4948			NlaIII	16	27	255	641	1293	1793	2051
StyI	4	1017	4094	4145	4281										
AccI	6	108	1460	1660	2071	3034									
NruI	5	309	1888	2279	3608	4576									

续表 1

SacII	5	1043	3722	4029	4223	4679		2352	3063	3337	3355	3541	3565
								3583	4149	4647	4799		
AvaI	6	18	635	1701	3558	3810	4567	以下限制内切酶在质粒上没有切点:					
Cfr101	6	591	1141	1782	2384	2784	4599	(There are no restriction sites for the following					
HincII	6	109	1461	1861	2000	2606	3226	enzymes in pXZ10145)					
BanI	7	817	1036	1138	2136	2909	3212	AflII	AccII	AsnI	Asp718	AsnII	
		4200						AvaIII	AvrII	BglII	BspHI	BstBI	
BglII	7	185	529	702	1567	1826	1835	DraI	EcoRV	FspI	FspII	HindIII	
		3721						HpaI	KpnI	MstI	NhlI	NsiI	
PvuI	7	164	3606	3932	4008	4295	4631	SfiI	SphI	SspI	StuI	XcaI	
		4780						XhoI					

参 考 文 献

- [1] Martin, J., Society for General Microbiology Symposium, Vol. 44, Cambridge University Press, p.25 1989.
- [2] Yamaguchi, Ryuji et al., *Agric. Biol. Chem.*, 50:2771, 1986.
- [3] 雷肇祖等, 工业微生物, 1:1, 1981.
- [4] 郑兆鑫等, 生物工程学报, 3:183, 1987.
- [5] 余红等, 生物工程学报, 5:51, 1989.
- [6] 那淑敏等, 生物工程学报, 7:312, 1991.
- [7] Sambrook, J. et al., *Molecular Cloning, A Laboratory Manual*. Second edition, Cold Spring Harbor Laboratory Press, 1989.
- [8] Sanger, F. et al., *Proc. Natl. Acad. Sci. USA*, 74:546, 1977.
- [9] Shigeo Abe. et al., *J. Gen. Appl. Microbiol.*, 13:279, 1967.

Determination and Analysis of Nucleotide Sequence of Plasmid pXZ 10145

Shen Tianxiang Jia Panxing Na Shumin Men Dapeng

(Institute of Microbiology, Academia Sinica, Beijing 100080)

With ABI 370A autosequencer, the total nucleotide sequence of plasmid pXZ10145 from *Corynebacteria glutamicum* 1014-6T has been determined using the dideoxy chain termination method. The plasmid contains 4887 base pairs (bps). Computer aided analysis of the sequence showed the location and number of restriction enzyme cutting sites and revealed eight open reading frames (ORFs) on the plasmid. The two sites on the plasmid pXZ10145, at which deletion occurred to result in plasmid pNAT 65 were confirmed. At these two sites a seven base pairs sequence "ATCTAGC" was found.

Key words *Corynebacteria glutamicum*; plasmid; DNA sequencing; repeat sequence