DOI: 10.3724/SP.J.1006.2020.94110

高抗黑胫病烤烟 BAC 文库的构建及分析

董庆园^{1,4,**} 马德清^{1,4,**} 杨学^{2,**} 刘勇¹ 黄昌军¹ 袁诚¹
方敦煌¹ 于海芹¹ 童治军¹ 沈俊儒³ 许银莲⁵ 罗美中² 李永平¹
曾建敏^{1,*}

¹ 云南省烟草农业科学研究院, 云南昆明 650021; ² 华中农业大学生命科学技术学院, 湖北武汉 430070; ³ 云南省烟草公司, 云南昆明 650011; ⁴ 云南农业大学烟草学院, 云南昆明 650201; ⁵ 云南省烟草公司临沧市公司, 云南临沧 677099

摘要: 烟草是重要的模式植物。本研究利用 pIndigoBAC536-S 载体及 *Hind* III 限制性内切酶酶解烟草基因组 DNA 的方法, 构建了烟草新品系 14-60 的细菌人工染色体(BAC)文库。该文库共包含 414,720 个克隆, 保存在 1080 块 384 板中。随机挑选的 120 个烟草 BAC 克隆检测结果表明, 外源插入片段大小为 97.0~145.5 kb, 平均约为 123 kb, 空载率极低(0), 覆盖烟草基因组 11 倍。用烟草 *hemA* 基因、*eIF4E-1* 基因、*NtFT* 基因的特异引物进一步验证, 该文库质量高、可用性强, 为烟草黑胫病抗性基因的克隆以及其他重要农艺性状和品质性状等功能基因克隆研究提供了基础资源。

关键词: 烟草; 基因组 DNA; 细菌人工染色体(BAC); 基因筛选

Construction and characterization of a BAC library for flue-cured tobacco line with high resistance to blank shank

DONG Qing-Yuan^{1,4,**}, MA De-Qing^{1,4,**}, YANG Xue^{2,**}, LIU Yong¹, HUANG Chang-Jun¹, YUAN Cheng¹, FANG Dun-Huang¹, YU Hai-Qin¹, TONG Zhi-Jun¹, SHEN Jun-Ru³, XU Yin-Lian⁵, LUO Mei-Zhong², LI Yong-Ping¹, and ZENG Jian-Min^{1,*}

¹ Yunnan Academy of Tobacco Agricultural Sciences, Kunming 650021, Yunnan, China; ² College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, Hubei, China; ³ Yunnan Province Tobacco Company, Kunming 650021, Yunnan, China; ⁴ College of Tobacco Science, Yunnan Agricultural University, Kunming 650201, Yunnan, China; ⁵ Lincang City Company of Yunnan Tobacco Company, Lincang 677099, Yunnan, China

Abstract: Tobacco (*N. tabacum*) is an important model crop in molecular biology research. In this study, a bacterial artificial chromosome (BAC) library of a flue-cured tobacco line 14–60 with high blank shank resistance and good quality was constructed. High molecular weight DNA was isolated using intact nuclei from tobacco, partially cleaved with *Hind* III and cloned into the pIndigoBAC536-S vector. The BAC library consisted of 414,720 clones arrayed in one thousand and eighty 384-microtiter plates, with an average insert size of 123 kb ranging from 97.0–145.5 kb. No empty insert clone was found. Based on an estimated genome size of 4500 Mb for common tobacco, the BAC library was estimated to cover 11 times of genome equivalents. The utility of the library was further confirmed by screening the library with the primers of tobacco *hemA*, *NtFT*, and *eIF4E-1* genes. The high capacity library will serve as a giant genomic resource for map-based cloning of quantitative trait loci or genes associated with important agronomic and smoking quality traits or resistance to blank shank, physical mapping and comparative genome analysis.

Keywords: tobacco; genomic DNA; bacterial artificial chromosome (BAC); gene screening

本研究由中国烟草总公司云南省公司科技项目(2017YN04, 2019530000241004)资助。

This study was supported by the Grants of Yunnan Province Tobacco Company (2017YN04, 2019530000241004).

* 通信作者(Corresponding author): 曾建敏, E-mail: zengjm2017@163.com

** 同等贡献(Contributed equally to this work)

第一作者联系方式: 董庆园, E-mail: BlueSkies183@163.com, 马德清, E-mail: 1847494695@qq.com, 杨学, E-mail: 919231424@qq.com

Received (收稿日期): 2019-07-29; Accepted (接受日期): 2020-01-15; Published online (网络出版日期): 2020-01-24.

URL: <http://kns.cnki.net/kcms/detail/11.1809.S.20200123.2202.014.html>

普通烟草(*Nicotiana tabacum*)是林烟草(*Nicotiana sylvestris*)和绒毛状烟草(*Nicotiana tomentosiformis*)杂交后自然加倍而成的异源四倍体植物^[1-4]。研究表明,烟草在生物制药^[5-6]和生物燃料^[7]方面的应用前景较大。烟草由于容易组织培养和遗传转化^[8-9],已被广泛应用于分子生物学研究,是重要的模式植物。普通烟草基因组重复序列多、杂合度高,已公布基因组序列。菲莫国际构建了林烟草(2.636 Gb)和绒毛状烟草(2.682 Gb)的参考基因组草图,分别含有覆盖基因组 82.9%和 70.6%的遗传信息,同时发现基因组内的重复序列比例高达 72%~75%^[10],在烤烟品种 K326 和白肋烟品种 TN90 基因组中的重复序列也分别高达 73.1%和 78.9%^[11],且仅有 19%的序列锚定到参考染色体组上。随着测序成本的降低,烟草参考基因组的质量有了进一步的提高^[12]。然而,要想完全解释烟草基因组信息,还需要进一步提高测序质量和组装、注释水平;国内在这方面的研究有了一定的进步,其中烤烟品种红花大金元的基因组组装质量较好,但对烟草有利遗传性状资源的开发和利用速度有待进一步加强。

细菌人工染色体(BAC)文库是功能基因分离和基因组结构研究的重要工具,在物理图谱构建、功能基因图位克隆、基因组全序列测定、比较基因组学分析等研究中发挥了重要作用,特别是对未测序或基因组结构复杂的物种^[13-16]。BAC 文库已被广泛地应用于细菌^[17]、动物^[18]以及棉花^[19]、水稻^[20-21]、小麦^[22]、玉米^[23]、大豆^[24]、番茄^[25]、大麦^[26]、高粱^[27]、珍珠谷子^[28]等物种的研究。构建烟草 BAC 文库的报道较少,仅有烟草野生种绒毛状烟草^[29],栽培烟草品种 Hicks Broadleaf^[30]和 SR1^[31]。由于早期受到技术和成本限制,很少开展高质量、大容量的 BAC 文库构建工作,因此对这些 BAC 文库的具体参数、检测标准以及质量评价均没有做详细介绍。

烤烟是中国、津巴布韦、巴西、美国、加拿大、印度等国家主要经济作物之一。烤烟品种 K326 是全球主栽品种之一,其烟叶品质深受广大卷烟工业企业认可,同时也是现在烟草品种选育的重要亲本之一,云烟 87、云烟 85、云烟 97 等多个大面积推广种植品种均与其有亲缘关系^[32-34]。黑胫病是土壤传播的真菌病害,是烟草生产最具毁灭性病害之一,严重影响全世界烟叶生产安全^[35-36]。育种家利用自然变异或诱导变异培育抗病新品种是解决黑胫病导致生产损失的最有效方法之一。目前,烟草黑胫病

抗源主要包括质量性状抗性和数量性状抗性 2 种类型。质量性状抗性基因有 *Php* 和 *Phl*, 分别来自于野生烟 *N. plumbaginifolia*^[37]和 *N. longiflora*^[38], 这 2 个抗源仅抗黑胫病 0 号小种、不抗 1 号小种,随着大面积种植含有该抗源的烟草品种,土壤中黑胫病 1 号小种有逐渐变为优势种群的趋势^[39]。来自于黄花烟的 *Wz* 抗性基因,高抗黑胫病 0 号和 1 号生理小种,属于类似质量性状,但抗性未达到免疫水平,其对烟草产量和质量的影响小,具有兼顾抗性和烟叶品质的特点^[40-41]。

本研究获得了含有 *Wz* 抗性基因的 K326 近等基因系。通过构建该材料 BAC 文库,不仅降低了对目的性状基因克隆的成本,也弥补了对于已知烟草基因组缺少来自未测序烟草品种的遗传信息的不足,为挖掘 *Wz* 抗性基因功能以及解析 K326 烟叶品质形成机制奠定重要基础。

1 材料与方法

1.1 试验材料

以含有高抗烟草黑胫病抗性基因 *Wz* 的育种材料为供体亲本,参照文献^[42]报道的标记,通过连续回交 7 代后再连续自交 7 代,将该抗性基因导入我国主栽烤烟品种 K326,选取其中农艺性状表现优良的 14-60 作为该文库构建的材料。文库构建载体是 pIndigoBAC536-S^[43],由华中农业大学基因组资源实验室构建。

1.2 试验方法

1.2.1 烟草基因组 DNA Plug 的制备 播种试验材料 14-60,四叶期收集嫩叶 50 g,冷冻保存于 -80 °C 备用。关于高分子量烟草基因组 DNA 的制备方法和过程参照 Luo 等^[44-45]。

1.2.2 烟草基因组 DNA 部分酶切条件的优化

参考文献的方法^[46]优化烟草基因组 DNA 部分酶切条件。在相同时间内,使用不同量的 *Hind* III 内切酶对相同量烟草基因组 DNA 进行酶切。实验设置内切酶浓度梯度为 8 个(0、0.3、0.5、1.0、1.5、2.0、4.0、10.0,单位为 U μ L⁻¹)。酶切反应体系包含 H₂O (25 μ L)、10×buffer (10 μ L)、40 mmol L⁻¹ 亚精胺(10 μ L)、5 μ L 不同浓度的内切酶和切碎的 1/2 Plug (50 μ L)。将反应体系置恒温 37 °C 的水浴锅中,反应 15 min 后置冰上,每个体系中加入 10 μ L 0.5 mol L⁻¹ EDTA (pH 8.0),静置 15 min。通过脉冲场凝胶电泳检测酶解效果,所用缓冲液为 0.5×TBE,参数设为

14 、1 s-50 s、120°、6 V cm⁻¹、18 h。

1.2.3 大片段 DNA 筛选 参照刘家栋等^[46-47]的方法进行 2 次大片段 DNA 筛选。采用优化后的酶解体系,进行 16 个相同酶解反应,在相同的脉冲场凝胶电泳条件下,进行第 1 次 DNA 片段分离。电泳结束后,在点样孔两侧向内各 0.5 cm 的地方,将胶块切割成 3 份,用 EB (溴化乙锭)溶液对两边的胶块染色,漂洗后,放入凝胶成像仪进行成像分析,用透明直尺作为参照。根据凝胶成像的结果,从未染色胶块中间,切下 100~250 kb 范围内宽约 2 cm 的胶条,将切下的胶条平分成 b 胶条和 a 胶条(b>a)。将 a 胶条放在制胶模具的左边, b 胶条放在右边,倒入 1% 琼脂糖填满两块胶的周围,之后将整块胶置凝胶电泳仪内,继续进行第 2 次大片段 DNA 筛选。设凝胶电泳参数为 120°、6 V cm⁻¹、4 s-4 s、温度 14 ,时间 18 h。第 2 次筛选结束后,切下 a、b 胶条两侧的胶块,染色、漂洗后,进行凝胶成像。在 a、b 胶条上切下未染色大片段 DNA 的胶块,并分别横切为 a1、a2 和 b1、b2,其中 a1<a2, b1<b2。在 4 条件下,将回收的 4 根胶条分别浸泡在 1×TAE 缓冲液中过夜。电洗脱回收胶条中的大片段 DNA,并检测洗脱产物的浓度。

1.2.4 BAC 文库的构建 参照前人方法^[44-47]构建烟草 BAC 文库。根据已知的 λDNA 浓度,判断每个回收区段 DNA 的浓度,只有回收区段 DNA 的浓度达到 1 ng μL⁻¹ 以上才进行连接反应。连接反应体系含有 84 μL 的大片段 DNA、4 μL 载体、2 μL T4 DNA 连接酶,以及 10 μL T4 DNA 连接酶 buffer,反应在 16 恒温水浴中进行 16 h。反应结束后,将连接体系置 65 干浴锅 15 min,连接产物在脱盐胶(1% 琼脂糖和 2% 葡萄糖)内脱盐 1 h。将 2 μL 脱盐后的产物与 18 μL 大肠杆菌 DH10B 感受态细胞充分混合,之后进行电转化,将转化产物置 500 μL SOC 培养基中 37 复苏 55 min。吸取复苏产物 100 μL,均匀涂布于 LB 固体培养基上, LB 固体培养基含有氯霉素(12.5 mg L⁻¹)、IPTG (100 mg L⁻¹)、X-gal (80 mg L⁻¹),将培养皿倒置于恒温 37 培养箱内,培养 18 h。培养结束后,随机从每个培养皿上挑取一定数量的白斑,检测插入外源 DNA 片段的大小,若达到目标,则挑拣白斑至含有冰冻培养基的 384 孔板中、37 培养,并复制 1 份,将培养好的细胞保存于-80 冰箱中。

1.2.5 BAC 文库的质量检测与基因筛选 参照前人方法^[46],随机从文库中挑取 120 个克隆,在含

有氯霉素(12.5 mg L⁻¹)的 LB 液体培养基中培养过夜,用碱裂解法提取质粒,用 I-Sce I 内切酶进行酶切、3.5 h,通过脉冲场凝胶电泳,检测外源 DNA 片段的大小、载体空载率等信息,脉冲场凝胶电泳缓冲液为 0.5×TBE,电泳条件为 5~15 s、120°、6 V cm⁻¹、14 、16 h。文库覆盖度依据烟草基因组大小、BAC 文库总克隆数目、平均插入外源 DNA 片段大小和载体空载率计算。

从 BAC 文库中随机挑取 40 个克隆,提取高纯度和浓度的质粒 DNA,进行 BAC 单末端测序(部分克隆进行双末端测序)。查询 NCBI 数据库中的烟草细胞器基因组数据库序列,与末端序列(BES)比对,估测该文库的细胞器污染率。

在 NCBI 数据库中,查询烟草 *hema* 基因(5'-GG GATGCACTAGGATGGTTCG-3', 5'-AGCATAAGCCC GCAGTTTCT-3')、*eIF4E-1* 基因(5'-TGCATACGGTG GTCCAAACA-3', 5'-GCCTGAGGTTTGTCTGAGAGT-3') 和 *NtFT* 基因(5'-CGACATGTTTGTTCGG TGG-3', 5'-TGAGAGGTTCTCCGTCCAGT-3')的序列、设计相关引物,对文库 1080 个一级混合池进行筛选并验证。

2 结果与分析

2.1 烟草基因组 DNA Plug 的制备与部分酶切

制备烟草细胞核 DNA 时,在 NIB 溶液中加入 0.15% β-巯基乙醇与 4% PVP-40,防止烟草基因组 DNA 氧化降解。加入 1% Triton 裂解烟草细胞器膜,避免细胞器 DNA 的污染。用蛋白酶 K 溶解液处理消化烟草细胞核膜,释放出烟草基因组 DNA,共得到烟草高分子量 DNA 包埋块 36 块。

从图 1 可看出,未加 *Hind* III 酶的泳道, DNA 拖带亮度低,证明提取过程中损伤低、提取的 DNA 质量好,满足 BAC 文库构建要求。*Hind* III 酶浓度梯度的酶切效果差异显著,其中 1.0 U μL⁻¹ 的 *Hind* III 酶浓度的效果最好(图 1)。

2.2 烟草基因组大片段 DNA 筛选

根据预先部分酶切的结果,选择了 1.0 U μL⁻¹ 的 *Hind* III 酶浓度和 14 min 酶切时间进行大量酶切,随后进行酶切产物的第 1 次和第 2 次筛选。图 2 为烟草基因组大片段 DNA 第 1 次筛选结果。参照荧光尺刻度,切下未经 EB 染色 5.4~6.4 cm 范围内的、6.4~7.4 cm 范围内的 b 片段胶块,得到 100~250 kb 的烟草基因组 DNA 片段。

图 1 烟草 plug DNA 部分酶切

Fig. 1 Partial digestions of tobacco DNA plugs

图中第 1 泳道为 λ ladder PFG Marker, 2~9 泳道分别为半块 plug 的酶切条带。酶浓度依次为 0、0.3、0.5、1.0、1.5、2.0、4.0、10.0 $\text{U } \mu\text{L}^{-1}$ 。酶切时间为 15 min。

The first lane is the λ ladder PFG marker, lanes 2~9 are enzyme digested bands of half plug. The order of enzyme concentration is 0, 0.3, 0.5, 1.0, 1.5, 2.0, 4.0, and 10.0 $\text{U } \mu\text{L}^{-1}$. The enzyme digestion time is 15 min.

图 2 烟草基因组 DNA 的第 1 次筛选

Fig. 2 First size selection of tobacco genomic DNA

1%琼脂糖凝胶电泳条件为 0.5×TBE, 1 s-50 s, 120°, 6 V cm^{-1} , 14 h, 18 h。Marker 为 λ ladder PFG。

DNA was separated on 1% CHEF agarose gel at 1 s-50 s linear ramp, 6 V cm^{-1} , 14 h in 0.5×TBE buffer for 18 h. The marker is λ ladder PFG marker.

图 3 显示, 第 1 次筛选得到 b 片段胶块中几乎不含小片段 DNA; 而 a 片段中则含有较多的小片段, 经过第 2 次脉冲场凝胶电泳分离, 胶块中 DNA 小片段的浓度明细降低。依据透明直尺参照的位置, 将 2 个未染色片段 DNA 分别切成 2 份, 每份胶块 0.5 cm 左右, 其中 a 片段从 7.5~8.5 cm 范围切下, 得到 a1、a2 ($a1 < a2$), b 片段从 7.8~8.8 cm 范围切下, 得到 b1、b2 ($b1 < b2$)。

图 3 烟草基因组 DNA 的第 2 次筛选

Fig. 3 Second size selection of tobacco genomic DNA

1%琼脂糖凝胶电泳条件为 0.5×TBE, 14 h, 4 s-4 s, 120°, 6 V cm^{-1} 。

DNA fragments were separated on 1% CHEF agarose gel at 4 s-4 s linear ramp, 6 V cm^{-1} , 14 h in 0.5×TBE buffer for 18 h. The marker used is λ ladder PFG marker.

2.3 BAC 文库构建及检测

从图 4 可以看出, a1 的浓度最低, 在 1 $\text{ng } \mu\text{L}^{-1}$ 左右, a2、b1、b2 的浓度都在 2~3 $\text{ng } \mu\text{L}^{-1}$ 之间, 满足与载体连接要求。

根据预转化后的结果进行连接产物大量转化, 最终获得 414,720 个克隆(表 1), 保存在 1080 块 384 孔板中。对随机挑选的 120 个克隆进行了分析, 其中部分 BAC 克隆平均插入片段检测结果见图 5。如 λ ladder PFG Marker 所示, 烟草 BAC 文库随机插入片段的大小在 97.0~145.5 kb 范围, 平均约为 123 kb, 空载率为 0。对 40 个 BAC 克隆进行了末端序列测定, 测序结果显示没有烟草细胞器和细菌 DNA 序列污染。按烟草基因组 4500 Mb 计算, 该文库覆盖度约为 11 倍基因组。

图 4 洗脱产物 DNA 的浓度检测

Fig. 4 Concentration of recovered DNA fragments

λ DNA 标准液 1 μL , 浓度分别为 1、2、3、4 $\text{ng } \mu\text{L}^{-1}$; a1、a2、b1、b2 电泳洗脱产物各 1 μL 。

The loading volumes for the recovered DNA fractions a1, a2, b1, b2, and the standard λ DNA solutions of 1, 2, 3, 4 $\text{ng } \mu\text{L}^{-1}$ are 1 μL each.

表 1 烟草 BAC 文库统计分析
Table 1 Statistics of common tobacco BAC library

名称 Name	指标 Index
克隆数 Total clones	414720
384 孔板数 Total 384-well plates	1080
载体 BAC vector	pIndigoBAC536-S
感受态细胞 Competent cell	DH10B
平均插入片段大小 Average insert size (kb)	123
烟草基因组覆盖度 Tobacco genome equivalent	11×
空载率 Insert-empty BACs (%)	0

图 5 烟草 BAC 文库随机克隆的插入片段检测(部分)
Fig. 5 Insert detection of random tobacco BAC library clones (partially)
为部分随机挑选的 40 个克隆的插入片段检测图片。Marker 为 λ ladder PFG。
It is the insert detection image of 40 randomly selected clones. Marker is λ ladder PFG.

2.4 烟草 BAC 文库基因筛选

为验证文库的可用性，利用 PCR 对 3 个基因进

行了筛选。从图 6 可以看出，*hemA* 基因引物的扩增条带约 460 bp，在烟草 BAC 文库一级混合池共筛选到 26 个阳性克隆。测序结果表明，这 26 个克隆的序列与 NCBI 中 *hemA* 基因序列结果一致。另外，*eIF4E-1* 基因共筛选到 11 个阳性克隆(图 7)，*NtFT* 基因共筛选到 13 个阳性克隆(图 8)，表明所构建的烟草 BAC 文库覆盖度达到理论覆盖度 11 倍以上，质量可靠。

3 讨论

大片段基因组 DNA 是制备 BAC 文库前提，然而大片段 DNA 容易降解。因此，本 BAC 文库构建过程中提取 HMW DNA 时，所有操作尽量保持在卫生干净、低温环境下，避免因为空气或者灰尘中的微生物等外界因素污染文库的构建质量。基因组 DNA 浓度低，获得大片段 DNA 的机会也低，但浓度

图 6 一级混合池 2 号板 132~166 编号的 PCR 筛选结果(*hemA*)
Fig. 6 PCR screening of Primary Mixing Pools 132~166 in No.2 plate (*hemA*)
Marker 为 Trans2K DNA marker, 从上到下依次是 2000、1000、500、250、100 bp。
The marker is Trans2K DNA marker, and the bands from top to bottom represent 2000, 1000, 750, 500, 250, and 100 bp.

图 7 一级混合池 2 号板 161~190 编号的 PCR 筛选结果(*eIF4E-1*)
Fig. 7 PCR Screening of Primary Mixing Pools 161~190 in No. 2 plate (*eIF4E-1*)
Marker 为 Trans2K DNA marker, 从上到下依次是 2000、1000、750、500、250、100 bp。
The marker is Trans2K DNA marker, and the bands from top to bottom represent 2000, 1000, 750, 500, 250, and 100 bp.

图 8 一级混合池 9 号板 898-927 编号 PCR 筛选结果(*NiFT*)

Fig. 8 PCR Screening of Primary Mixing Pools 898-927 in No. 9 plate (*NiFT*)

Marker 为 Trans2K DNA marker, 从上到下依次是 2000、1000、750、500、250、100 bp。

The Marker is Trans2K DNA marker, and the bands from top to bottom represent 2000, 1000, 750, 500, 250, and 100 bp.

过高的基因组 DNA 在脉冲场中很难完全分离, 无法通过 PFGE 分开, 大、小片段的 DNA 容易堆积在一起^[44-45]。此外, DNA 浓度太高还会增加包埋块的粘稠度, 使限制性内切酶与基因组 DNA 无法均匀混合, 酶切的效果差。本文库构建选用的 DNA 浓度合适(图 1), 为成功构建烟草 BAC 文库奠定重要基础。

外源 DNA 片段大, 与载体连接的产物转化效率低; 外源 DNA 片段小, 与载体相连产物转化却很容易。因此, 筛选合适的 DNA 片段也是构建 BAC 文库的重要步骤。本研究通过 2 次筛选, 获得了随机插入片段大小在 97.0~145.5 kb 范围(图 2 和图 3), 平均插入片段为 123 kb (图 5)。通过优化限制性内切酶 *Hind* III 的用量及反应时间, 实现对烟草基因组 DNA 的随机酶切, 控制了高分子量大小片段均一插入, 效果和文献报道接近^[48-49]。本研究通过随机检测 120 个烟草 BAC 克隆, 发现本文库的空载率为 0 (图 5)。有研究表明, 如果内切酶的用量过大, 将产生非特异性的酶切, 产生假阳性比例高, 文库的空载率也高^[50], 严重影响文库的质量。

用于构建烟草 BAC 文库的克隆载体由华中农业大学基因组资源实验室构建, 通过酶切, 除去复合载体 pHZAUBAC1 高拷贝复制子后, 获得克隆载体 pIndigoBAC536-S。利用 pIndigoBAC536-S 克隆载体上 *I-Sce* I 释放完整插入片段的特点, 当电泳泳道出现 2 条大小相近条带时, 可判断为双克隆, 极大改善了 BAC 文库的构建分析和评价等工作。pHZAUBAC1 复合载体已经被成功应用于众多 BAC 文库的构建^[47,51]。

BAC 文库基因筛选方法, 包括 Southern 杂交法、高密度影印膜(HDR)菌落杂交筛选法、PCR 筛选法^[51]。前两方法均需要将 BAC 克隆转移在尼龙膜上, 再使用探针筛选克隆, 过程复杂、工作量大, 而 PCR 筛选方法与前两者相比具有操作简单, 成本低的优点, 应用范围广。早花基因(*NiFT*)能够有效缩短

烟草开花时间, 该技术成功应用于烟草育种^[11,52]。研究表明 *va* 基因型烟草植株的抗性是由于对 PVY 感病的基因片段的缺失造成的^[53], 即 *eIF4E-1* 基因, 属于真核生物翻译起始因子(eucaryotic initiation factor, 简称为 *eIF*)的一种类型^[54]。5-氨基乙酰丙酸(ALA)的合成是植物体内四吡咯生物合成的限速步骤, 谷氨酰 t-RNA 还原酶基因(*hemA*)是控制 ALA 合成最关键酶^[55]。本研究采用烟草 *hemA* 基因、*eIF4E-1* 基因、*NiFT* 基因的特异引物对构建的文库进行筛选验证, 结果表明这 3 个基因的文库覆盖度高达 11~26 倍, 可见, 所构建文库的质量比预期还好。

种植不同类型抗性品种是解决该病害的重要途径。然而关于 Wz 抗性位点的遗传信息非常有限, 全面解释其基因功能和结构是利用该抗性的重要基础。本研究构建的 BAC 文库一共有 414,720 个克隆, 保存在 1080 个 384 孔板中, 平均插入片段约 123 kb, 覆盖烟草基因组 11 倍, 文库的插入片段较大, 均一性较好, 叶绿体和线粒体 DNA 污染率低、空载率低。虽然烟草的参考基因组序列已经发表, 但是其参考基因组序列的质量仍然有待提升。由于 Wz 抗性基因来自黄花烟, 属于另一个烟草种, 与现有测序的烟草品种存在种间差异, 已知的参考基因组中并未包含本研究中导入的 Wz 抗性基因的相关信息, 因此完全依赖参考基因组无法进行 Wz 抗性基因的克隆。此外, 该 BAC 文库是以 K326 为背景构建的, 因此可以作为现有烟草基因组在组装和注释等方面存在的问题的一个重要参考, 为解析 K326 其他优良性状提供基础资源。综上所述, 本研究构建的烟草 BAC 文库能够很好地覆盖烟草的基因组信息, 可以作为提升和完善烟草基因组序列的重要资源, 包括部分基因组信息无法锚定到具体染色体位置, 同时为解决现有烟草基因组中缺少来自黄花烟抗性基因 Wz 的遗传信息提供最直接的补充, 也为挖掘 Wz 抗性基因功能, 解析 K326 烟叶品质形成机理以及

烟草数量性状相关功能基因的克隆提供基础材料。

4 结论

本研究构建的高抗黑胫病烤烟 BAC 文库含有 414,720 个克隆, 保存在 1080 个 384 孔板中, 外源插入片段大小在 97.0 kb 至 145.5 kb 间、平均约为 123 kb, 覆盖普通烟草基因组 11 倍。文库的插入片段较大, 均一性较好, 叶绿体和线粒体 DNA 污染率低、空载率低。为 *Wz* 抗性基因定位克隆和功能挖掘, 解析 K326 烟叶品质形成机制提供基础材料。

References

- [1] Liao L J, Liu L F, Lai K L. Somatic hybridization in the genus *Nicotiana*: *N. sylvestris* and *N. tomentosiformis*. *Memoirs College Agric Natl Taiwan Univ*, 1990, 30: 73–82.
- [2] Goodspeed, Harper T. The genus *Nicotiana*. *Soil Sci*, 1955, 80: 250.
- [3] Kenton A, Parokonny A S, Gleba Y Y, Bennett M D. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. *Mol Gen Genet*, 1993, 240: 159–169.
- [4] Gazdová B, Siroky J, Fajkus J, Brzobohaty B, Kenton A, Parokonny A, Heslop-Harrison J S, Palme K, Bezděk M. Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the *Nicotiana tomentosiformis* genomic component. *Chromosome Res*, 1995, 3: 245–254.
- [5] Ma J K, Drossard J, Lewis D, Altmann F, Boyle J, Christou P, Cole T, Dale P, van Dolleweerd C J, Isitt V, Katinger D, Lobedan M, Mertens H, Paul M J, Rademacher T, Sack M, Hundleby P A C, Stiegler G, Stoger E, Twyman R M, Vcelar B, Fischer R. Regulatory approval and a first-in-human phase I clinical trial of a monoclonal antibody produced in transgenic tobacco plants. *Plant Biotechnol J*, 2015, 13: 1106–1120.
- [6] Tusé D, Tu T, McDonald K A. Manufacturing economics of plant-made biologics: case studies in therapeutic and industrial enzymes. *BioMed Res Int*, 2014, 2014: 1–16.
- [7] Vanhercke T, El Tahchy A, Liu Q, Zhou X R, Shrestha P, Divi U K, Ral J P, Mansour M P, Nichols P D, James C N, Horn P J, Chapman K D, Beaudoin F, Ruiz-López N, Larkin P J, de Feyter R C, Singh S P, Petrie J R. Metabolic engineering of biomass for high energy density: oilseed-like triacylglycerol yields from plant leaves. *Plant Biotechnol J*, 2014, 12: 231–239.
- [8] Vasil V, Hildebrandt A C. Differentiation of tobacco plants from single, isolated cells in microcultures. *Science*, 1965, 150: 889–892.
- [9] Hoekema A, Hirsch P R, Hooykaas P J J, Schilperoot R A. A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, 1983, 303: 179–180.
- [10] Siervo N, Battey J N, Ouadi S, Bovet L, Goepfert S, Bakaher N, Peitsch M C, Ivanov N V. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. *Genome Biol*, 2013, 14: R60.
- [11] Siervo N, Battey J N, Ouadi S, Bakaher N, Bovet L, Willig A, Goepfert S, Peitsch M C, Ivanov N V. The tobacco genome sequence and its comparison with those of tomato and potato. *Nat Commun*, 2014, 5: 1–9.
- [12] Edwards K D, Fernandez-Pozo N, Drake-Stowe K, Humphry M, Evans A D, Bombarely A, Allen F, Hurst R, White B, Kernodle S P, Bromley J R, Sanchez-Tamburrino J P, Lewis R S, Mueller L A. A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics*, 2017, 18: 448.
- [13] Chen M, Presting G, Barbazuk W B, Goicoechea J L, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S, Hingbottom S, Phimpilalai J, Phimpilalai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Main D, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Budiman M A, Tomkins J P, Luo M, Bancroft I, Salse J, Regad F, Mohapatra T, Singh N K, Tyagi A K, Soderlund C, Dean R A, Wing R A. An integrated physical and genetic map of the rice genome. *Plant Cell*, 2002, 14: 537–545.
- [14] Mueller L A, Tanksley S D, Giovannoni J J, van Eck J, Stack S, Choi D, Kim B D, Chen M, Cheng Z, Li C, Ling H, Xue Y, Seymour G, Bishop G, Bryan G, Sharma R, Khurana J, Tyagi A, Chattopadhyay D, Singh N K, Stiekema W, Lindhout P, Jesse T, Lankhorst R K, Bouzayen M, Shibata D, Tabata S, Granell A, Botella M A, Giuliano G, Frusciante L, Causse M, Zamir D. The tomato sequencing project, the first cornerstone of the international solanaceae project (SOL). *Comp Funct Genomics*, 2005, 6: 153–158.
- [15] Luo M, Gu Y Q, You F M, Deal K R, Ma Y, Hu Y, Huo N, Wang Y, Wang J, Chen S, Jorgensen C M, Zhang Y, McGuire P E, Pasternak S, Stein J C, Ware D, Kramer M, McCombie W R, Kianian S F, Martis M M, Mayer K F, Sehgal S K, Li W, Gill B S, Bevan M W, Simková H, Dolezel J, Weining S, Lazo G R, Anderson O D, Dvorak J. A 4-gigabase physical map unlocks the structure and evolution of the complex genome of *Aegilops tauschii*, the wheat D-genome progenitor. *Proc Natl Acad Sci USA*, 2013, 110: 7940–7945.
- [16] Edwards D, Batley J, Snowden R J. Accessing complex crop genomes with next-generation sequencing. *Theor Appl Genet*, 2013, 126: 1–11.
- [17] Wang X, Liu Q, Wang H, Luo C X, Wang G, Luo M. A BAC based physical map and genome survey of the rice false smut fungus *Villosiclava virens*. *BMC Genomics*, 2013, 14: 883.
- [18] Spiranzlova P, Dhome-Pollet S, Fellah J S, Da Silva C, Tlapakova T, Labadie K, Weissenbach J, Poulain J, Jaffredo T, Wincker P, Krylov V, Pollet N. Construction and characterization of a BAC library for functional genomics in *Xenopus tropicalis*. *Dev Biol*, 2017, 426: 255–260.
- [19] Liu Y, Zhang B, Wen X, Zhang S, Wei Y, Lu Q, Liu Z, Wang K, Liu F, Peng R. Construction and characterization of a bacterial artificial chromosome library for *Gossypium mustelinum*. *PLoS One*, 2018, 13: e0196847.
- [20] Zhang J, Chen L L, Xing F, Kudrna D A, Yao W, Copetti D, Mu T, Li W, Song J M, Xie W, Lee S, Talag J, Shao L, An Y, Zhang C L, Ouyang Y, Sun S, Jiao W B, Lv F, Du B, Luo M, Maldonado C E, Goicoechea J L, Xiong L, Wu C, Xing Y, Zhou D X, Yu S, Zhao Y, Wang G, Yu Y, Luo Y, Zhou Z W, Hurtado B

- E, Danowitz A, Wing R A, Zhang Q. Extensive sequence divergence between the reference genomes of two elite indica rice varieties Zhenshan 97 and Minghui 63. *Proc Natl Acad Sci USA*, 2016, 113: E5163–E5171.
- [21] Ammiraju J S, Luo M, Goicoechea J L, Wang W, Kudrna D, Mueller C, Talag J, Kim H, Sisneros N B, Blackmon B, Fang E, Tomkins J B, Brar D, MacKinnon D, McCouch S, Kurata N, Lambert G, Galbraith D W, Arumuganathan K, Rao K, Walling J G, Gill N, Yu Y, SanMiguel P, Soderlund C, Jackson S, Wing R A. The *Oryza* bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus *Oryza*. *Genome Res*, 2006, 16: 140–147.
- [22] Janda J, Safár J, Kubaláková M, Bartos J, Kovárová P, Suchánková P, Pateyron S, Ciháliková J, Sourdille P, Simková H, Faivre-Rampant P, Hribová E, Bernard M, Lukaszewski A, Dolezel J, Chalhou B. Advanced resources for plant genomics: a BAC library specific for the short arm of wheat chromosome 1B. *Plant J*, 2006, 47: 977–986.
- [23] Wang C, Shi X, Liu L, Li H, Ammiraju J S, Kudrna D A, Xiong W, Wang H, Dai Z, Zheng Y, Lai J, Jin W, Messing J, Bennetzen J L, Wing R A, Luo M. Genomic resources for gene discovery, functional genome annotation, and evolutionary studies of maize and its close relatives. *Genetics*, 2013, 195: 723–737.
- [24] Wu C, Nimmakayala P, Santos F, Springman R, Scheuring C, Meksem K, Lightfoot D, Zhang H. Construction and characterization of a soybean bacterial artificial chromosome library and use of multiple complementary libraries for genome physical mapping. *Theor Appl Genet*, 2004, 109: 1041–1050.
- [25] Frary A, Hamilton C M. Efficiency and stability of high molecular weight DNA transformation: an analysis in tomato. *Transgenic Res*, 2001, 10: 121–132.
- [26] Schulte D, Ariyadasa R, Shi B, Fleury D, Saski C, Atkins M, Pieter D, Wu C, Andreas G, Peter L, Nils S. BAC library resources for map-based cloning and physical map construction in barley (*Hordeum vulgare* L.). *BMC Genomics*, 2011, 12: 247.
- [27] Woo S S, Jiang J, Gill B S, Paterson A H, Wing R A. Construction and characterization of bacterial artificial chromosome library of *Sorghum bicolor*. *Nucleic Acids Res*, 1994, 22: 4922–4931.
- [28] Allouis S, Qi X, Lindup S, Gale M, Devos K. Construction of a BAC library of pearl millet, *Pennisetum glaucum*. *Theor Appl Genet*, 2001, 102: 1200–1205.
- [29] 高晓明, 陈艳玲, 刘贵山, 李凤霞, 王卫峰, 任媛媛, 孙玉合. 绒毛状烟草 BAC 文库的构建. *中国烟草科学*, 2012, 33(3): 68–71.
- Gao X M, Chen Y L, Liu G S, Li F X, Wang W F, Ren Y Y, Sun Y H. BAC library construction of villi tobacco (*Nicotiana tomentosiformis*). *Chin Tob Sci*, 2012, 33(3): 68–71 (in Chinese with English abstract).
- [30] Sierro N, van Oeveren J, van Eijk M J, Martin F, Stormo K E, Peitsch M C, Ivanov N V. Whole genome profiling physical map and ancestral annotation of tobacco Hicks Broadleaf. *Plant J*, 2013, 75: 880–889.
- [31] Nagaki K, Shibata F, Suzuki G, Kanatani A, Ozaki S, Hironaka A, Kashiwara K, Murata M. Coexistence of NtCENH3 and two retrotransposons in tobacco centromeres. *Chrom Res*, 2011, 19: 591–605.
- [32] 李永平, 王颖宽, 马文广, 谭彩兰. 烤烟新品种云烟 87 的选育及特征特性. *中国烟草科学*, 2001, 22(4): 38–42.
- Li Y P, Wang Y K, Ma W G, Tan C L. Breeding and selecting of a new flue-cured tobacco variety Yunyan 87 and its characteristics. *Chin Tob Sci*, 2001, 22(4): 38–42 (in Chinese with English abstract).
- [33] 谭彩兰, 李永平, 王颖宽, 马文广, 雷永和. 烤烟新品种云烟 85 的选育及其特征特性. *中国烟草科学*, 1997, 18(1): 7–10.
- Tan C L, Li Y P, Wang Y K, Ma W G, Lei Y H. Breeding and selecting a new flue-cured tobacco of Yunyan 85 and its characteristics. *Chin Tob Sci*, 1997, 18(1): 7–10 (in Chinese with English abstract).
- [34] 李永平, 肖炳光, 焦芳婵, 张谊寒, 于海芹, 卢秀萍. 烤烟新品种云烟 97 的选育及其特征特性. *中国烟草科学*, 2012, 33(4): 28–31.
- Li Y P, Xiao B G, Jiao F C, Zhang Y H, Yu H Q, Lu X P. Breeding and characteristics of a new flue-cured tobacco variety Yunyan 97. *Chin Tob Sci*, 2012, 33(4): 28–31 (in Chinese with English abstract).
- [35] Lucas G B. Diseases of tobacco. *Quarterly Rev Biol*, 1965, 4548: 1421.
- [36] 朱贤朝, 王彦亭, 王智发. 中国烟草病害. 北京: 中国农业出版社, 2002. pp 21–22.
- Zhu X C, Wang Y T, Wang Z F. Tobacco Diseases of China. Beijing: China Agriculture Press, 2002. pp 21–22 (in Chinese).
- [37] Chaplin J F. Transfer of black shank resistance from *Nicotiana plumbaginifolia* to flue-cured *N. tabacum*. *Tob Sci*, 1962, 6: 184–189.
- [38] Valteau W D, Stokes G W, Johnson E M. Nine years' experience with the *Nicotiana longiflora* factor for resistance to *Phytophthora parasitica* var. *nicotianae* in the control of black shank. *Tob Sci*, 1960, 4: 92–94.
- [39] Sullivan M J, Melton T A, Shew H D. Managing the race structure of *Phytophthora parasitica* var. *nicotianae* with cultivar rotation. *Plant Dis*, 2005, 89: 1285–1294.
- [40] Drake K E, Lewis R S. An introgressed *Nicotiana rustica* genomic region confers resistance to *Phytophthora nicotianae* in cultivated tobacco. *Crop Sci*, 2013, 53: 1366–1374.
- [41] Zeng J M, Nifong J, Liu Y, Huang C J, Fang D H, Lewis R S, Li Y P. Evaluating diverse systems of tobacco genetic resistance to *Phytophthora nicotianae* in Yunnan, China. *Plant Pathol*, 2019, 68, 1616–1623.
- [42] Drake K E, Moore J M, Bertrand P, Fortnum B, Peterson P, Lewis R S. Black shank resistance and agronomic performance of flue-cured tobacco lines and hybrids carrying the introgressed region Wz. *Crop Sci*, 2015, 55: 1–8.
- [43] Shi X, Zeng H, Xue Y, Luo M. A pair of new BAC and BIBAC vectors that facilitate BAC/BIBAC library construction and intact large genomic DNA insert exchange. *Plant Methods*, 2011, 7: 33.
- [44] Luo M, Wang Y H, Frisch D, Joobeur T, Wing R A, Dean R A. Melon bacterial artificial chromosome (BAC) library construction using improved methods and identification of clones linked to the locus conferring resistance to melon *Fusarium*

- Wilt (Fom-2). *Genome*, 2001, 44: 154–162.
- [45] Luo M, Wing R A. An improved method for plant BAC library construction. *Methods Mol Biol*, 2003, 236: 3.
- [46] 刘家栋, 王革娇, 罗美中. 阿维链霉菌 BAC 文库的构建及分析. *华中农业大学学报*, 2016, 35(5): 45–50.
- Liu J D, Wang G J, Luo M Z. Construction and analysis of a BAC library of *Streptomyces avermitilis* genome. *J Huazhong Agric Univ*, 2016, 35(5): 45–50 (in Chinese with English abstract).
- [47] Thorsen J, Zhu B, Frengen E, Osoegawa K, de Jong P J, Koop B F, Davidson W S, Høyheim B. A highly redundant BAC library of Atlantic salmon (*Salmo salar*): an important tool for salmon projects. *BMC Genomics*, 2005, 6: 50.
- [48] Kim U J, Birren B W, Slepak T, Mancino V, Boysen C, Kang H L, Simon M I, Shizuya H. Construction and characterization of a human bacterial artificial chromosome library. *Genomics*, 1996, 34: 213–218.
- [49] Osoegawa K, Woon P Y, Zhao B, Frengen E, Tateno M, Catanese J J, de Jong P J. An improved approach for construction of bacterial artificial chromosome libraries. *Genomics*, 1998, 52: 1–8.
- [50] 刘庆丽, 王晓明, 王革娇, 罗朝喜, 谭新球, 罗美中. 稻曲病菌 UV-2 菌株细菌人工染色体文库构建及分析. *微生物学通报*, 2013, 40: 1715–1722.
- Liu Q L, Wang X M, Wang G J, Luo C X, Tan X Q, Luo M Z. Construction of a bacterial artificial chromosome library of *Villosiclava virens* UV-2 genome. *Microbiol China*, 2013, 40: 1715–1722 (in Chinese with English abstract).
- [51] Asakawa S, Abe I, Kudoh Y, Kishi N, Wang Y, Kubota R, Kudoh J, Kawasaki K, Minoshima S, Shimizu N. Human BAC library: construction and rapid screening. *Gene*, 1997, 191: 69–79.
- [52] Lewis R S, Kernodle S P. A method for accelerated trait conversion in plant breeding. *Theor Appl Genet*, 2009, 118: 1499–1508.
- [53] Noguchi S, Tajima T, Yamamoto Y, Ohno T, Kubo T. Deletion of a large genomic segment in tobacco varieties that are resistant to potato virus Y (PVY). *Mol Gen Genet*, 1999, 262: 822–829.
- [54] Liu Y, Zeng J, Yuan C, Guo Y, Yu H, Li Y, Huang C. Cas9-PF, an early flowering and visual selection marker system, enhances the frequency of editing event occurrence and expedites the isolation of genome-edited and transgene-free plants. *Plant Biotechnol J*, 2019, 17: 1191–1193.
- [55] Richter A S, Banse C, Grimm B. The GluTR-binding protein is the heme-binding factor for feedback control of glutamyl-tRNA reductase. *eLife*, 2019, 8: e46300.