

周 瑶, 周恩强, 姚梦楠, 等. 基于转录组测序数据的蚕豆 InDel 标记开发和应用[J]. 江苏农业科学, 2025, 53(8): 35–41.
doi:10.15889/j.issn.1002-1302.2025.08.005

基于转录组测序数据的蚕豆 InDel 标记开发和应用

周 瑶, 周恩强, 姚梦楠, 赵 娜, 李宗迪, 朱宇翔, 王永强, 李 波,
王学军, 缪亚梅, 汪凯华, 顾春燕, 魏利斌
(江苏沿江地区农业科学研究所经济作物研究室, 江苏南通 226000)

摘要:我国是全球最大的蚕豆生产国,然而由于蚕豆基因组巨大(约为 13.0 G),蚕豆基础研究进展缓慢。目前,受限于有限的蚕豆专用分子标记,我国蚕豆种质资源的评价和利用研究开展的不够深入。为了丰富蚕豆分子标记类型,充分有效评价和利用蚕豆种质资源,拟开展 7 份蚕豆材料的转录组测序,并对测序数据间存在的插入/缺失(InDel)位点进行特征分析。结果表明,成功挖掘出 2 666 个 InDel 位点,其中插入、缺失位点分别有 841、1 825 个,且单碱基的插入/缺失位点类型占比最大。此外,根据 InDel 位点信息,从插入/缺失碱基长度大于 5 bp 的序列中随机挑选 25 条序列进行 InDel 标记开发。利用开发的标记对国内 141 份蚕豆资源进行扩增,共筛选到有效 InDel 标记 20 对,多态性信息含量均值为 0.32,平均每对标记中检测到 2.85 个等位基因,香农指数均值为 0.62。聚类分析结果表明,江苏、浙江、湖北等省份的资源呈交叉分布,可能因为近年来各省份的蚕豆资源相互借鉴、利用,从而打破了地域限制,增加了彼此间的联系。基于转录组测序数据首次开发蚕豆 InDel 标记并成功将其应用于蚕豆种质资源的遗传多样性分析,不仅丰富了目前蚕豆的分子标记类型,而且对于今后蚕豆的基础研究和育种研究具有一定意义。

关键词:蚕豆;转录组;InDel;分子标记;遗传多样性

中图分类号:S643.603.2 **文献标志码:**A **文章编号:**1002-1302(2025)08-0035-06

蚕豆(*Vicia faba* L.)别称胡豆、罗汉豆、佛豆等,属豆科蝶形花亚科野豌豆属^[1],是一种粮、饲、菜、肥兼用的重要食用豆类,其营养价值、经济价值较高,在我国农业种植结构调整和农民脱贫致富中具有重要作用^[2-3]。我国是全球最大的蚕豆生产国,每年的蚕豆种植面积约为 86.67 万 hm²,年产量约为 180 万 t^[4]。中国农业农村部的相关数据显示,我国蚕豆种植面积呈逐年增长趋势,产业发展前景广阔。

我国蚕豆种植区域广,生态差异大,种质资源十分丰富,这些丰富的种质资源是开展蚕豆遗传改良研究的基础。然而,目前对这些资源的收集、保护、利用、研究工作还不够全面、深入,并且蚕豆基因组庞大(约为 13.0 G),限制了蚕豆基础研究的开

展,相比于其他作物,蚕豆资源在分子方面的研究已经远远落后,其遗传背景和亲缘关系仍旧模糊,资源利用率较低。分子标记技术可以从分子水平直接测定基因型,能快速、准确鉴定种质资源^[5]。目前,国内外关于用蚕豆分子标记鉴定蚕豆资源的研究已有一些进展,分子标记类型主要有 SSR(简单重复序列)、AFLP(扩增片段长度多态性)、ISSR(简单重复序列区间)、SNP(单核苷酸多态性)等。Link 等用 RAPD(随机扩增的多态性 DNA)标记研究 3 个蚕豆自交群体,发现地中海种、欧洲小粒种有明显分离,欧洲大粒种介于两者间^[6]。Kwon 等用 TRAP(目标区域扩增多态性)标记研究 151 份国内外蚕豆资源,发现地方品种间的变异较大^[7]。孙虎等基于转录组测序数据开发了 SSR 分子标记,对 80 份蚕豆资源进行遗传多样性分析及聚类分析,发现 80 份蚕豆资源的遗传多样性比较丰富^[8]。马钰利用磁珠富集法、数据库搜索法开发蚕豆 SSR 分子标记,发现用数据库搜索法开发的分子标记的多态性较高,而用磁珠富集法开发的分子标记的多态性较低^[9]。王海飞等用 11 个 ISSR(简单重复序列区间)标记分析我国 18 个省份春蚕豆资源、秋蚕豆资源的遗传多样性与亲缘关系,结果表明,相邻省份的资

收稿日期:2024-05-23

基金项目:江苏省种业振兴“揭榜挂帅”项目[编号:JBGS(2021)056];国家食用豆产业技术体系建设专项(编号:CARS-08-Z10);江苏省“双创人才”项目(编号:YSSCRC2022469);江苏沿江地区农业科学研究所青年科技基金[编号:YJ(2022)005]。

作者简介:周 瑶(1993—),女,江苏泰兴人,硕士,助理研究员,主要从事食用豆育种研究。E-mail:1090362118@qq.com。

通信作者:魏利斌,博士,研究员,主要从事豆类品种育种研究。E-mail:libinwei2013@aliyun.com。

源亲缘关系相对较近^[10]。Zong 等用 10 个 AFLP 标记研究来自国内外的 243 份秋蚕豆资源,发现我国种质资源与世界其他地区的资源有明显区别^[11]。

虽然目前国内外关于蚕豆分子标记的开发和蚕豆资源的鉴定研究已有一些进展,但是蚕豆专用分子标记的类型和数量仍然有限,严重制约着蚕豆种质资源的评价和利用。本研究利用蚕豆转录组数据首次开发插入/缺失(InDel)分子标记,并成功用其对蚕豆种质资源进行聚类分析,不仅丰富了目前蚕豆分子标记类型,而且对蚕豆相关分子遗传学研究和育种研究具有重要意义。

1 材料与方法

1.1 试验材料

本试验材料为笔者所在课题组收集的蚕豆种质资源,共 141 份。这些蚕豆资源来自国内 13 个省份,包括江苏、浙江、云南、四川、湖北、福建、重庆、青海、内蒙古等地,试验材料均种植于江苏沿江地区农业科学研究所薛窑基地(32°07'N,120°27'E),2023 年 3 月底采集嫩叶后迅速放入液氮中进行储藏,然后放入 -80 ℃ 冰箱中保存,以备后续试验。

1.2 试验方法

1.2.1 转录组数据的获得 转录组数据来源于笔者所在课题组 2022 年对地理来源、表型性状差异显著的 7 份蚕豆材料进行的二代转录组测序结果^[12]。

1.2.2 转录组数据 InDel 位点的挖掘与引物设计 以蚕豆全长转录本为参考模板^[12],用 GATK 3 软件^[13]分析上述 7 份蚕豆转录组数据中的 InDel 位点信息。根据 InDel 位点信息,从插入/缺失碱基长度大于 5 bp 的序列中随机挑选 25 条序列,用 Primer Premier 5.0 软件^[14]设计 InDel 标记引物。引物的参数设定如下:GC 含量为 40% ~ 60%,引物长度为 18 ~ 25 bp,预期片段大小为 60 ~ 150 bp,退火温度为 55 ~ 62 ℃,上下游引物差异在 2 ℃ 以内。本研究设计的所有引物尽可能避免出现引物二级结构,设计的引物由南京金斯瑞生物科技有限公司合成。

1.2.3 DNA 的提取及质量检测 利用植物基因组 DNA 提取试剂盒[天根生化科技(北京)有限公司]提取各样本蚕豆种质资源的嫩叶组织 DNA,之后用 1% 琼脂糖凝胶电泳检测所提 DNA 的质量,将符合质量要求的 DNA 放在 -20 ℃ 冰箱中保存。

1.2.4 PCR 及电泳检测 10 μL PCR 扩增反应体系如下:DNA 模板 1 μL,上下游引物各 1 μL,3G

Taq Master Mix 5 μL,ddH₂O 2 μL。PCR 扩增反应程序如下:95 ℃ 5 min;95 ℃ 30 s,55 ~ 62 ℃(不同引物退火温度不同)30 s,72 ℃ 30 s,30 个循环;72 ℃ 3 min,10 ℃ 保温 4 min。

电泳检测:用聚丙烯酰胺凝胶进行垂直电泳检测 PCR 扩增产物,点样量为 1 μL,银染、显色后观察结果。能扩增出清晰条带的引物被视为有效引物,不能扩增出清晰带或带型不可读的引物被认为是无效引物。

1.2.5 多态引物的筛选及数据统计分析 首先用合成的 25 对引物对随机挑选的 12 份蚕豆种质资源 DNA 进行初筛,选取清晰可读的 InDel 引物开展蚕豆种质资源的扩增分析。采用人工读带方法,在相同迁移位置上,有带的记为“1”,无带的记为“0”,缺失的记为“9”。用 Excel 2010 软件构建数据矩阵,记录 InDel 位点的分布特征,用 POPGENE 1.32^[15]、NTSYS 2.1^[16]软件进行标记的遗传分析和亲缘关系树状图的绘制。

2 结果与分析

2.1 转录组测序结果

挑选地理来源、表型性状具有显著差异的 7 份蚕豆材料进行种植,提取幼苗嫩叶 RNA 进行二代转录组测序分析,共获得测序数据 43.9 Gb,质量符合下一步分析的需求。转录组测序数据统计结果见表 1。

表 1 蚕豆转录组数据统计结果

样品名	片段总数 (条)	碱基总数 (bp)	GC 含量 (%)	≥Q30 (%)
C102	21 665 656	6 471 016 376	43.33	94.23
C130	20 621 496	6 157 145 976	43.77	94.22
C131	20 753 678	6 198 255 600	43.46	94.49
C229	21 837 644	6 514 193 770	43.78	95.06
C290	21 050 034	6 258 400 874	44.32	94.52
TCX7-1	21 608 376	6 439 075 410	43.74	94.57
WY7-1	19 660 417	5 865 328 020	43.67	94.33

注:Q30 表示碱基质量值,代表某个碱基的正确率为 99.9%;≥Q30(%)表示质量值≥30 的碱基数占总碱基数的比例。

2.2 转录组中 InDel 位点的特征分析

2.2.1 InDel 变异位点数量和类型分析 用 Smith-Waterman 比对算法对上述样品与蚕豆参考全长转录本^[12]中存在的插入缺失序列进行分析,发现仅有 925 条序列(占 66 484 条蚕豆参考转录组序

列的 1.39%) 存在插入缺失现象, 共含有 2 666 个 InDel 变异位点, 平均每条序列有 2.83 个 InDel 位点。

在 2 666 个 InDel 变异位点中, 分别有 841 个 (31.55%) 插入位点、1 825 个 (68.45%) 缺失位点, 变异长度范围为 1~36 bp, 以小片段 (1~5 bp) 的变异为主 (91.94%), 而大于 5 bp 的 InDel 变异位点仅有 215 个, 占 8.06%, 见表 2、图 1。

表 2 InDel 分析结果统计

搜索条目	数量 (个)	占比 (%)
InDel 序列数目	925	
InDel 变异位点总数	2 666	
插入类型 InDel 位点数	841	31.55
缺失类型 InDel 位点数	1 825	68.45
插入/缺失 1 bp 的 InDel 位点数	1 768	66.32
插入 1 bp 的 InDel 位点数	498	59.22
缺失 1 bp 的 InDel 位点数	1 270	69.59
插入/缺失 2 bp 的 InDel 位点总数	363	13.62
插入 2 bp 的 InDel 位点数	139	16.53
缺失 2 bp 的 InDel 位点数	224	12.27
插入/缺失 3 bp 的 InDel 位点总数	195	7.31
插入 3 bp 的 InDel 位点数	73	8.68
缺失 3 bp 的 InDel 位点数	122	6.68
插入 4 bp 的 InDel 位点数	37	4.40
缺失 4 bp 的 InDel 位点数	52	2.85
插入 5 bp 的 InDel 位点数	15	1.78
缺失 5 bp 的 InDel 位点数	21	1.15
插入 6~10 bp 的 InDel 位点数	50	5.95
缺失 6~10 bp 的 InDel 位点数	80	4.38
插入 11~36 bp 的 InDel 位点数	29	3.45
缺失 11~36 bp 的 InDel 位点数	56	3.07

2.2.2 InDel 位点主要碱基类型分析 在 2 666 个插入/缺失 InDel 位点中, 共有 1 768 个插入/缺失

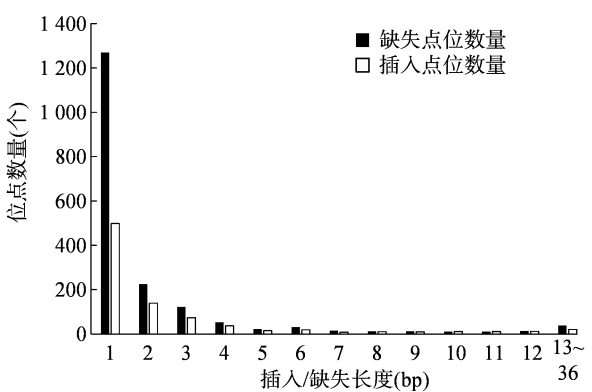


图1 InDel 插入/缺失碱基长度的位点数

1 bp 的 InDel 位点, 占比为 66.32%, 其中 A/T 的插入/缺失位点总数的比例最高, 达 65.50%; 共有 363 个插入/缺失 2 bp 的 InDel 位点, 占比为 13.62%, 其中 AA/TT 基元为优势基元, 占比为 30.58%; 共有 195 个插入/缺失 3 bp 的 InDel 位点, 占比为 7.31%, 其中优势基元为 AAG/TTC, 占比为 8.72%。

在 841 个插入 InDel 位点中, A/T 的插入位点数最多, 占比为 70.08%, 插入 2 bp 的 InDel 位点数占比为 16.53%, 其中基元为优势 AA/TT 基元, 占比为 35.25%; 在 1 825 个缺失 InDel 的位点中, A/T 的缺失位点数最多, 占比为 63.70%, 缺失 2 bp 的 InDel 位点数占比为 12.27%, 其中 AA/TT 基元为优势基元, 占比为 27.68% (图 2)。

2.3 InDel 分子标记的开发

根据转录组序列的多态位点信息, 用 Primer Premier 5.0 成功开发 25 对 InDel 分子标记 (表 3)。

2.4 InDel 分子标记的筛选和多态性分析

首先用 25 对 InDel 引物对随机挑选的 12 份蚕豆资源 DNA 进行 PCR 扩增和电泳检测, 初步筛选出 22 对具有扩增产物的 InDel 引物。然后用初筛引物对 141 份蚕豆资源 DNA 进行检测, 综合考虑分子标记扩增的有效等位基因数、杂合度、多态性信息量、

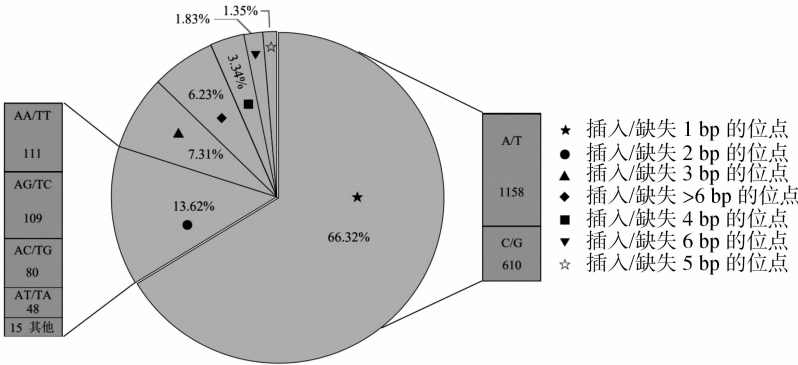


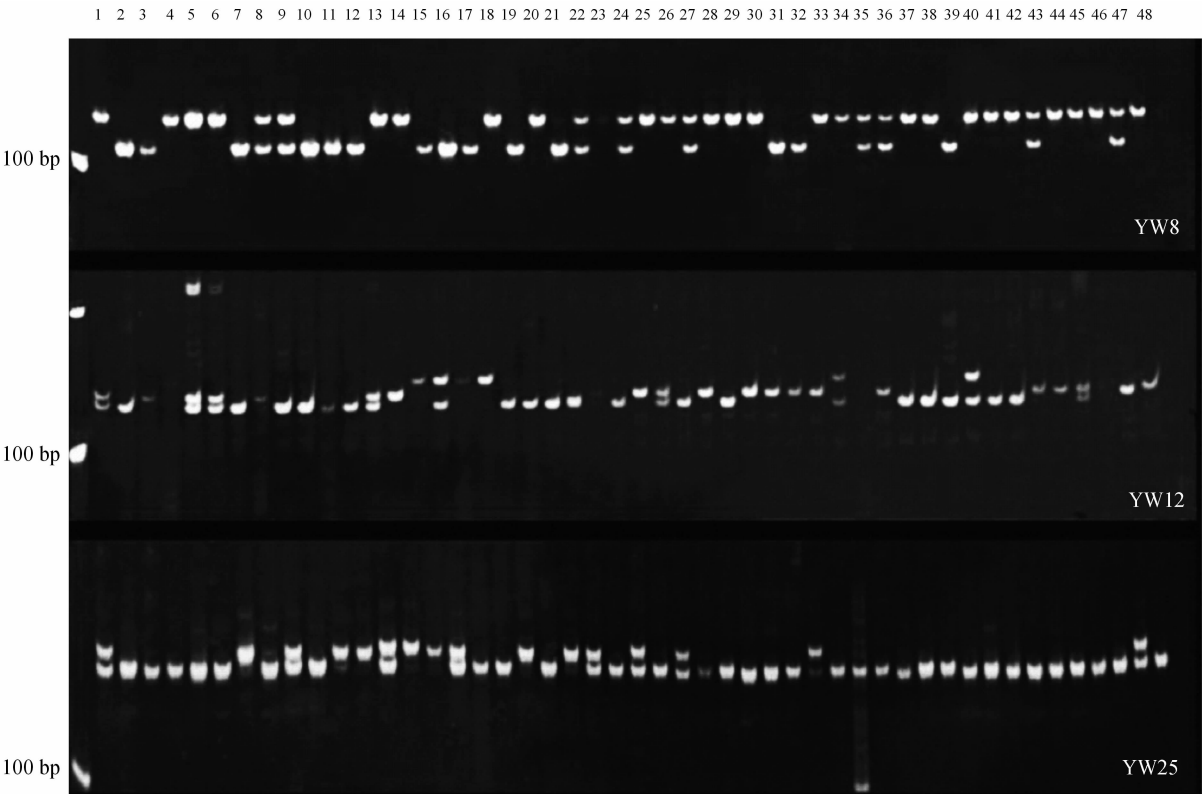
图2 InDel 插入/缺失碱基类型占比

表 3 部分 InDel 引物信息

引物 编号	前引物序列(5'→3')	退火温度 (℃)	引物大小 (bp)	后引物序列 (5'→3')	退火温度 (℃)	引物大小 (bp)	产物大小 (bp)
YW1	TCAAACGACGGAACATAACTT	55.2	21	CATTTGGTAGATAACATAGTGGAT	55.0	25	127
YW2	ATGATGTGATTTTGGCTGATTGG	61.5	23	AACACATCACCACGGTAAAATCG	61.6	23	123
YW3	TTACCTATTTCAGTTTCCTATTGTGG	57.4	25	GGTTCGAATCTGGTTTATGTCC	57.4	23	94
YW4	GTGATGCCTGGAGTTGAGATTTT	59.2	24	CCTGTGAGAAAGAATAACCAAGAAA	59.4	25	111
YW5	CTGAGTTTTTTCCTCTTTTTT	59.6	23	TGTGCTTGACTTTCCAGAATCCT	60.3	23	143
YW6	CGTGTAAGTTTTCCTCCAAGTG	58.5	24	TGATTGAAAGAAAAAACTACCAT	57.2	25	118
YW7	GGGTTCTATCAGTCATTGCTTT	59.5	23	TCAAAGGTGCTAACACGGACT	59.3	22	86
YW8	TGTCCCGTGATGTTCTGAT	58.0	20	TCAACTTCTTCATCTGATTCAACAC	58.2	25	133
YW9	GCCCAATCCACCCTCCC	60.1	17	GGTGAGTGATTTTCAAGTGTTGGG	61.0	23	101
YW10	AAATCAAGCGGAGTGAATAAGAA	58.2	23	ATTACAGAAAAAATTATGTGCTCG	57.8	25	116
YW11	TAAATTCTGGGACAAAAGGATGAAT	60.8	25	TAGGACTTGGACGCCCCAT	60.0	19	93
YW12	CATTGAAAACTGCTAGAGAAGGAT	58.8	25	CCTCTCTCTGAAATCTCCAT	59.0	22	98
YW13	GCGTGAGTTTGAATTTCTTTGTTT	58.3	24	AAACTGAAGGTGATGTTGTTGATG	58.7	24	90
YW14	ATCTTAACAACATACATCGACGAAAC	57.2	25	CACCCAGCGCCATCGT	58.4	16	132
YW15	TATCTTGTGTAGCCCTGTGACTT	58.0	24	TAACCACATCCATATCACTCAAAAT	57.8	25	76

基因多样性及扩增产物分子量、扩增带型清晰度等因素,最终选择 20 对 InDel 引物进行后续数据统计。结果(图 3)显示,20 对 InDel 标记共扩增出 57 个多态性条带,平均每对标记产生 2.85 个多态性条带。用

POPGENE 1.32 软件对上述引物在 141 份蚕豆材料中的扩增结果进行统计分析。由表 4 可以看出,20 对 InDel 标记的等位基因数在 2~6 之间,有效等位基因数在 1.10~3.15 之间,均值为 1.74;香农指数



1~48 指部分蚕豆品种编号
图3 部分 InDel 引物的扩增结果

范围为 0.19 ~ 1.41, 均值为 0.62; 多态性信息含量 (PIC) 在 0.08 ~ 0.64 之间, 均值为 0.32, 有 3 个高度多态性标记 ($PIC \geq 0.5$), 11 个中度多态性标记 ($0.25 < PIC < 0.5$) 和 6 个轻度多态性标记 ($PIC \leq 0.25$)。InDel 标记 18 的有效等位基因数、多态性信息含量最高, InDel 标记 3 的有效等位基因数、多态性信息含量最低。

表 4 20 个 InDel 标记的蚕豆多态性信息

InDel 标记编号	等位基因数 (个)	有效等位基因数 (个)	香农指数	多态性信息含量
1	4	1.52	0.58	0.29
2	2	1.10	0.20	0.09
3	2	1.10	0.19	0.08
4	2	1.77	0.63	0.34
5	2	1.90	0.67	0.36
6	3	1.77	0.70	0.36
7	2	1.84	0.65	0.35
8	2	1.71	0.61	0.33
9	2	1.76	0.62	0.34
10	2	1.24	0.35	0.18
11	2	1.48	0.51	0.27
12	6	2.39	1.00	0.50
13	2	1.99	0.69	0.37
14	3	1.25	0.37	0.18
15	4	1.56	0.70	0.33
16	2	1.84	0.65	0.35
17	3	2.94	1.09	0.59
18	6	3.15	1.41	0.64
19	4	1.18	0.36	0.15
20	2	1.33	0.41	0.22
均值	2.85	1.74	0.62	0.32

2.5 141 份蚕豆种质资源的遗传多样性分析结果

非加权组平均法 (UPGMA) 聚类分析结果 (图 4) 表明, 当遗传相似性系数为 0.56 时, 除了资源 T66 外, 其他资源都聚在 I 类; 当遗传相似性系数为 0.69 时, I 类资源又分为 1、2、3、4 亚类, 2、3、4 亚类资源包含多数西南地区及相邻省份的资源; 当遗传相似性系数为 0.74 时, 1 亚类资源又分成 a、b、c、d、e、f 6 个小亚类, 江苏、浙江、湖北等省份的资源呈交叉分布。另外, 云南的蚕豆资源基本紧密聚集在 i、ii、iii 3 个小分支。

3 讨论与结论

蚕豆是一种重要的食用豆类作物, 我国蚕豆资



图4 蚕豆遗传相似度聚类结果

源十分丰富,但是关于蚕豆的分子生物学研究基础薄弱,远远落后于其他大宗作物,蚕豆资源的遗传背景及遗传多样性仍不清晰。随着分子技术的不断发展,目前国内外关于用分子标记鉴定蚕豆资源的研究已有一些进展,但是有限的蚕豆专用分子标记数量仍然制约着蚕豆种质资源的评价和利用。目前,研究蚕豆的分子标记类型主要有 SSR、AFLP、ISSR、SNP,虽然 InDel 标记具有在基因组内分布广、密度高、变异稳定、多态性强、容易检测、可重复且成本低廉等优点^[17-20],但是该标记还未在蚕豆种质的评价中得以应用。

本研究在分析 7 份蚕豆材料转录组测序数据的基础上,成功挖掘出 2 666 个 InDel 位点,这些 InDel 位点变异长度的范围为 1 ~ 36 bp,以小片段(1 ~ 5 bp)的插入/缺失为主,占 91.94%。其中,InDel 长度为 1 bp 的位点数最多,这种分布特点与油菜、茄子、白菜、青菜等植物中 InDel 的长度分布一致^[21-24]。利用开发的 25 对 InDel 标记对国内 141 份蚕豆资源的 DNA 进行 PCR 扩增,最终筛选出 20 对多态性较好的标记,多态率为 80%,而已报道的茄子、高粱、长豇豆、槟榔、黄瓜中 InDel 标记的多态率分别为 34.7%、42.44%、51.52%、57%、57.4%^[25-29],远低于本研究中蚕豆标记 80% 的多态率,进一步证实蚕豆的遗传背景较丰富。本研究中有 5 对标记未能成功扩增出条带或多态,其原因可能与 PCR 产物内含子太大、所设计的引物位于剪切位点上、转录组拼接序列质量不高等因素有关。对于筛选得到的 20 对标记而言,平均每对检测到 2.85 个等位基因,香农指数均值为 0.62,平均多态性信息含量均值为 0.32,整体而言多态性较好。

UPGMA 聚类分析结果表明,当遗传相似性系数为 0.56 时,除了资源 T66 外,所有国内资源都聚成 I 类,推测资源 T66 的遗传背景特殊,也可能含有国外蚕豆资源的遗传背景。江苏、浙江、湖北等省份的资源呈交叉分布,这可能由于近年来各省份蚕豆资源相互借鉴、利用,从而打破了地域限制,增加了彼此间的联系。另外,云南的蚕豆资源基本紧密聚集在 3 个小分支,推测云南资源有较为独特的遗传背景。

本研究首次开发了 InDel 标记,并利用其对蚕豆种质资源进行遗传多样性分析,不仅丰富了目前蚕豆分子标记类型,而且对进一步促进分子辅助育种在蚕豆种质资源中的研究具有积极的推动作用。

参考文献:

- [1]程须珍,田静,王丽侠,等. 中国食用豆类品种志(第二辑)[M]. 北京:科学出版社,2023:294.
- [2]周恩强,周瑶,姚梦楠,等. 南方冬闲田绿肥鲜食两用蚕豆产业发展策略及建议[J]. 上海农业学报,2023,39(3):122-127.
- [3]王海飞,宗绪晓. 蚕豆种质资源、抗病育种和 QTL 定位及抗逆性研究进展[J]. 植物遗传资源学报,2011,12(2):259-270.
- [4]刘玉蛟,张红岩,郭兴莲,等. 基于“一优两高”战略的蚕豆产业认知与产业发展[J]. 青海科技,2020,27(6):18-21.
- [5]尹凤祥,张维琴. 抗病性的分子标记在作物育种上的应用[J]. 国外农学(杂粮作物),1999,19(5):15-19.
- [6]Link W, Dixkens C, Singh M, et al. Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers[J]. Theoretical and Applied Genetics,1995,90(1):27-32.
- [7]Kwon S J, Hu J, Coyne C J. Genetic diversity and relationship among faba bean (*Vicia faba* L.) germplasm entries as revealed by TRAP markers[J]. Plant Genetic Resources,2010,8(3):204-213.
- [8]孙虎. 基于蚕豆转录组数据开发 SSR 标记及蚕豆抗虫象遗传多样性分析[D]. 荆州:长江大学,2024:32-36.
- [9]马钰. 蚕豆 SSR 标记的开发及遗传连锁图谱的构建[D]. 北京:中国农业科学院,2012:30-32.
- [10]王海飞,关建平,马钰,等. 中国蚕豆种质资源 ISSR 标记遗传多样性分析[J]. 作物学报,2011,37(4):595-602.
- [11]Zong X X, Liu X J, Guan J P, et al. Molecular variation among Chinese and global winter faba bean germplasm[J]. Theoretical and Applied Genetics,2009,118(5):971-978.
- [12]Zhao N, Zhou E Q, Miao Y M, et al. High-quality faba bean reference transcripts generated using PacBio and Illumina RNA-seq data[J]. Scientific Data,2024,11(1):359.
- [13]vander Auwera G A, Carneiro M O, Hartl C, et al. From FastQ data to high confidence variant calls: the genome analysis Toolkit best practices pipeline[J]. Current Protocols in Bioinformatics,2013,43(1110):11.
- [14]Lalitha S. Primer Premier 5[R]. Biotech Software & Internet Report,2000:270-272.
- [15]Yeh F C, Yang R C, Boyle T B J, et al. POPGENE, the user-friendly shareware for population genetic analysis[EB/OL]. [2024-05-20]. <https://www.semanticscholar.org/paper/PopGene%20-the-user-friendly-shareware-for-population-Yeh-Yang/67958a632e377b561219e15748787bf6987222f1>.
- [16]Rohlf F J. NTSYS-pc 2.0. Numerical taxonomy and multivariate analysis system[M]. New York:Exeter Publishing Ltd., 2000.
- [17]潘存红,王子斌,马玉银,等. InDel 和 SNP 标记在水稻图位克隆中的应用[J]. 中国水稻科学,2007,21(5):447-453.
- [18]张红梅,翟文,金海军,等. 利用 InDel 标记分析 23 份黄瓜种质的遗传多样性及核心种质资源筛选[J]. 上海农业学报,2019,35(4):28-33.
- [19]杨洁,赫佳,王丹碧,等. InDel 标记的研究和应用进展[J]. 生物多样性,2016,24(2):237-243.
- [20]Feltus F, Wan J, Schulze S, et al. An SNP resource for rice genetics

孙晓丽,武艳芳,王永丽,等. 金银花 MYB 基因家族的功能鉴定与组织表达差异分析[J]. 江苏农业科学,2025,53(8):41-48.
doi:10.15889/j.issn.1002-1302.2025.08.006

金银花 MYB 基因家族的功能鉴定与组织表达差异分析

孙晓丽,武艳芳,王永丽,高璐,陈寻峰,蒋建雄

(江苏大学生物质能源研究院,江苏镇江 212013)

摘要:为了探究金银花 MYB 家族转录因子的调控功能,挖掘对绿原酸生物合成具有调控作用的 MYB 转录因子,从金银花的转录组数据中鉴定出 37 个 MYB 转录因子,并对其所编码蛋白的理化性质、进化关系、启动子序列、保守基序区域和不同组织中的相对表达量进行系统分析。保守基序(motif)区域分析结果表明,37 个 MYB 转录因子含有保守的 R2R3-MYB 结构域,属于 R2R3-MYB 亚族。上游启动子元件分析结果显示,这些 MYB 转录因子参与植物激素、外界非生物胁迫、生长发育的应答等。组织差异性表达分析结果显示,LjMYB6-1、LjMYB12-2、LjMYB4 为代表的 12 个转录因子的基因在白花中相对表达量较高,以 LjMYB6、LjMYB11、LjMYB12 为代表的 12 个转录因子的基因在青蕾中相对表达量较高,而金银花中的白花和青蕾是绿原酸含量最高的药用部位。系统进化树结果显示,LjMYB11、LjMYB4、LjMYB6、LjMYB6-1、LjMYB12、LjMYB12-2 共 6 个 MYB 转录因子位于参与植物苯丙素代谢途径调控的进化分支中,表明它们具有潜在的调控苯丙素代谢途径的作用。综合分析,认为这 6 个转录因子可能参与对金银花绿原酸生物合成的调控。

关键词:金银花;MYB;转录因子;功能分析;组织表达;motif 分析

中图分类号:S567.7+90.1 **文献标志码:**A **文章编号:**1002-1302(2025)08-0041-08

金银花(*Lonicera japonica* Thunb.)作为我国的传统中药,药用部位为植物未开放的花蕾,金银花早在《本草纲目》中就有记载,使用历史悠久,药用价值极高。现代医学表明,金银花具有清热解毒、凉血止痢、抗菌消炎、保肝利胆、提高免疫力等功效^[1-2],对于呼吸道感染和肺炎等疾病也有很好的治疗作用^[3]。金银花含有绿原酸、木犀草素、多酚、黄酮类、三萜皂苷类、挥发油等物质^[4-6]。其中,绿

原酸类物质是金银花发挥抗菌消炎作用的主要药效成分^[7],也是评价金银花等级的重要指标^[8]。

绿原酸(chlorogenic acid),也称为咖啡鞣酸,是由咖啡酸(caffeic acid)和奎宁酸(quinic acid)生成的缩酚酸^[9]。在金银花中,次级代谢产物绿原酸主要通过细胞的苯丙素代谢途径进行合成^[10]。绿原酸合成通路从苯丙氨酸开始,首先生成对-香豆酰 CoA(*p*-Coumaroyl CoA),再分别经历 3 种主要途径,羟基化肉桂酰辅酶 A/奎宁酸肉桂酰羟基化转移酶(HQT)途径、对-香豆酰奎宁酸(*p*-coumaroyl quinic acid)途径、羟基化肉桂酰 D-葡萄糖/奎宁酸羟基化肉桂酰转移酶(HCGQT)途径,最终生成绿

收稿日期:2024-03-05

作者简介:孙晓丽(1996—),女,山西大同人,硕士研究生,主要从事金银花绿原酸代谢调控研究。E-mail:chinamoyu@163.com。

通信作者:蒋建雄,博士,研究员,主要从事植物育种与生物质资源化利用研究。E-mail:jxjiang2002@163.com。

and breeding based on subspecies *indica* and *japonica* genome alignments[J]. *Genome Research*,2004,14(9):1812-1819.

[21]岳晓鹏. 基于甘蓝型油菜基因组重测序开发 InDel 标记[D].

武汉:华中农业大学,2014:30-31.

[22]尤园园,王帅,方莹莹,等. 茄子全基因组 InDel 变异特征及分子标记开发和应用[J]. *园艺学报*,2024,51(3):520-532.

[23]Park S, Yu H J, Mun J H, et al. Genome-wide discovery of dna polymorphism in *Brassica rapa* [J]. *Molecular Genetics & Genomics*,2010,283(2):135-145.

[24]许玉富,黄依琳,李荣华,等. 基于 RAD-seq 的菜心 InDel 标记开发及应用[J]. *广东农业科学*,2024,51(3):91-102.

[25]吉康娜,邹俊杰,林丹妮,等. 基于茄子基因组重测序的 InDel 标记开发及应用[J]. *植物遗传资源学报*,2019,20(5):1278-1288.

[26]蒲伟军,谭冰兰,贺丹晨,等. 利用重测序技术开发高粱 InDel 分子标记[J]. *生物技术进展*,2023,13(5):730-741.

[27]杨易,黎庭耀,李国景,等. 基于重测序的长豇豆基因组 InDel 标记开发及应用[J]. *园艺学报*,2022,49(4):778-790.

[28]齐兰,杨耀东,黄丽云,等. 基于重测序的槟榔 InDel 标记开发及应用[J]. *分子植物育种*,2024,22(5):1501-1507.

[29]卢霞,刘梦华,邓志军,等. 基于 InDel 标记的黄瓜种质资源遗传多样性分析[J]. *江苏农业科学*,2021,49(1):49-54.