

鱼腥草转录组 SSR 位点信息分析及其多态性研究

黎晓英, 刘胜贵, 王 丹, 黄豪兰, 龙 强, 蒋玉红, 郭文博, 魏 麟*

怀化学院生命科学系 民族药用植物资源研究与利用湖南省重点实验室, 湘西药用植物与民族植物学湖南省高校重点实验室, 湖南 怀化 418008

摘 要: 目的 分析鱼腥草 *Houttuynia cordata* 转录组中的 SSR 位点信息, 并设计简单重复序列 (SSR) 引物, 以期鱼腥草分子标记辅助育种提供有力工具。方法 利用 MISA 工具筛选了鱼腥草转录组测序获得的 63 954 条 unigenes, 对其 SSR 位点信息进行了分析; 在此基础上利用 Primer 3 设计 SSR 引物, 并随机选择 50 对 SSR 引物对 16 株不同来源的鱼腥草进行多态性扩增分析。结果 在鱼腥草的转录组中, 共搜索到 4 800 个 SSRs, 分布于 4 413 条 unigenes, SSR 位点出现频率为 7.51%。SSRs 位点中三核苷酸重复是主要类型, 占总 SSRs 的 41.54%; 其次是单核苷酸重复基序 (占 27.35%)。SSRs 所包含的重复基元中, 单核苷酸重复基元 A/T 是优势重复基元类型, 占总 SSRs 的 27.0%。利用 Primer3 共设计出 3 068 对 SSR 引物。随机选择 50 对引物进行 PCR 扩增, 其中 43 对 (86.0%) 扩增出清晰、可重复的条带, 且表现出多态性。利用 UPGMA 作图, 将 16 个样品分为 2 类。结论 鱼腥草转录组中 SSR 位点出现频率高, 类型丰富; 大量的 SSR 为其遗传多样性分析和遗传图谱构建奠定基础。

关键词: 鱼腥草; 转录组; SSR 信息; 多态性; 引物

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2016)10-1762-06

DOI: 10.7501/j.issn.0253-2670.2016.10.022

Analysis on SSR loci information in transcriptome of *Houttuynia cordata* and its polymorphism

LI Xiao-ying, LIU Sheng-gui, WANG Dan, HUANG Hao-lan, LONG Qiang, JIANG Yu-hong, GUO Wen-bo, WEI Lin

Key Laboratory of Research and Utilization of Ethnomedicinal Plant Resources of Hunan Province; Key Laboratory of Xiangxi Medicinal Plant and Ethnobotany of Hunan Higher Education, Department of Life Sciences, Huaihua University, Huaihua 418008, China

Abstract: Objective Simple sequence repeats (SSR) loci information in the transcriptome of *Houttuynia cordata* was analyzed in this study and the more powerful tools were provided for molecular marker-assisted breeding in this plant. **Methods** SSR loci were searched in all 63 954 unigenes by using MISA. SSR loci information was analyzed and SSR primers were designed by Primer 3. Furthermore, 50 pairs of primers were randomly selected for the polymorphic analysis on 16 *H. cordata* plants collected from different habitats. **Results** A total of 4 800 SSRs were found in the transcriptome of *H. cordata*, which distributed in 4 413 unigenes with the distribution frequency of 7.51%. Tri-nucleotide repeat was the main type, accounted for as much as 41.54% of all SSRs, followed by mono-nucleotide repeat motif (27.35%). The mononucleotide repeat motifs of A/T were the predominant repeat type (27.0%). A total of 3068 pairs of SSR primers were designed by using Primer 3. For validating the availability of those SSR primers, 50 pairs of primers were randomly selected for PCR amplification. Among them, 43 pairs of primers (86.0%) produced clear and reproductive bands, which showed polymorphism, and 16 *H. cordata* plants collected from different places were divided into two groups by UPGMA. **Conclusion** There are numerous SSRs in *H. cordata* transcriptome with high frequency and various types, this will provide the basis for study on genetic diversity and genetic map for *H. cordata*.

Key words: *Houttuynia cordata* Thunb.; transcriptome; simple sequence repeats information; polymorphism; primer

收稿日期: 2015-06-22

基金项目: 国家自然科学基金资助项目 (30870230); 湖南省科技计划项目 (2013FJ6090, 2014FJ4207); 湖南省高校创新平台开放基金项目 (15K100); 湖南省生物类专业大学生创新训练中心及大学生创新项目资助

作者简介: 黎晓英 (1976—), 女, 土家族, 硕士, 讲师, 研究方向为植物资源评价与分子生物学。Tel: (0745)2851037 E-mail: hhlxy@126.com

*通信作者 魏 麟, 博士, 副教授, 主要从事分子生物学与分子遗传学方面研究。Tel: (0745)2851037 E-mail: hhweilin@163.com

鱼腥草 *Houttuynia cordata* Thunb. 为三白草科 (Saururaceae) 蕺菜属 *Houttuynia* Thunb. 植物^[1], 因其全株有特殊的鱼腥气味而得名, 鱼腥草别名侧耳根、猪鼻孔、狗贴耳等。鱼腥草作为一种具有清热解毒, 利尿消肿等功效的中药材, 使用历史悠久^[2]。其注射液常用来治疗呼吸道、泌尿系统炎症疾病, 外科手术中也常用鱼腥草提取液进行消炎、抑菌。对消化系统疾病、皮肤病等也有很好的疗效^[3]。

DNA 分子标记能在分子水平上检测出 DNA 遗传多态性, 具有较大的应用价值^[4]。迄今在鱼腥草植物中所用到的分子标记仍局限于随机扩增多态 DNA (RAPD)^[5-6]和扩增简单重复序列间区 (ISSR)^[7]等随机引物标记。由于其稳定性和重现性一直备受质疑, 因此目前已较少使用。最常用的特异标记简单序列重复 (simple sequence repeat, SSR) 因其重现性好, 扩增靶序列在基因组中位置确定等优点, 备受研究者青睐。然而, SSR 标记的开发难度大, 不仅费时费力, 且费用很高, 多局限在具备全基因组数据或大量表达序列标签的物种中进行开发^[8]。目前尚未见鱼腥草植物中开发 SSR 标记的报道。

前期课题组成员对鱼腥草进行了转录组测序, 获得了大量的 unigenes 序列, 本研究将尝试利用这些序列, 首次开发鱼腥草的 SSR 标记, 期望这些标记可以用于鱼腥草的亲缘关系分析、种源鉴定及遗传图谱构建等方面的研究。

1 材料与方法

1.1 鱼腥草样品采集、转录组测序与 SSR 的筛选

采集不同产地的样品 16 份 (表 1), 经怀化学院伍贤进教授鉴定为鱼腥草 *Houttuynia cordata* Thunb.。鱼腥草地下茎、地上茎、花和叶片 RNA 混

合样品的转录组测序由华大基因公司进行, 共获得 185 276 条 Contigs, 组装出 63 954 条 unigenes。以组装出来的 Unigene 作为参考序列, 使用 SSR 软件 MicroSATellite (<http://pgrc.ipk-gatersleben.de/misa/>), 来找出所有的 SSR。参数设置为“1-12、2-6、3-5、4-5、5-4、6-4”, 1-12 代表单碱基重复至少 12 次才算 SSR, 双碱基 6 次, 三碱基 5 次, 以此类推, 重复单元最多有 6 个碱基。

1.2 SSR 引物设计

对所有 SSR 重复单元在 unigene 上前后序列的长度进行筛选。只保留前后序列均不小于 150 bp 的 SSR, 用 Primer3 软件对其设计引物。为验证 SSR 引物是否可用和检测 SSR 的可靠性, 随机选择合成 50 对引物, 由生工生物工程 (上海) 股份有限公司合成, 然后进行 PCR 扩增。

1.3 鱼腥草基因组 DNA 提取与 PCR 扩增

采用改进的 CTAB 法提取鱼腥草基因组 DNA。利用核酸/蛋白定量仪定量。定量后的 DNA 稀释至 30 μg/mL 以备 PCR 扩增用。反应体系: 10×buffer 2 μL, 1 mol/L MgCl₂ 1 μL, 10 μmol/L 上游引物和下游引物各 0.7 μL, 2.5 mmol/L dNTPs 1 μL, Taq DNA polymerase 0.5 μL, 10 mg/L 模板 DNA 2 μL, 加 ddH₂O 至总 16 μL。反应条件为 94 °C 预变性 1 min, 94 °C 变性 30 s, 最佳 *T_m* 退火 30 s, 72 °C 延伸 30 s, 循环 35 次, 72 °C 延伸 5 min。取扩增产物 2 μL 先用 1.5% 琼脂糖判断是否具有多态性条带, 确定具有多态性条带的, 将剩余产物用 8% 非变性聚丙烯酰胺凝胶检测, 进行基因分型与条带读取。

1.4 数据处理

采用人工读带的方法, 将电泳图上可重复的、清晰的条带计为“1”, 同一位置无带或不易分辨的弱带计为“0”, 建立原始数据矩阵。利用软件 Popgene 32 计算每对引物扩增位点的等位位点数、多态性信息量、遗传距离, 用 Mega4 软件进行 UPGMA 聚类分析。

2 结果与分析

2.1 鱼腥草转录组 SSR 位点数量与分布

对鱼腥草转录组测序获得总长度为 43 395 361 bp 的 63 954 条 unigenes 数目进行检测, 在 4 413 条 unigenes 数目中找到符合条件的 SSR, 发生频率 (含有 SSR 的 unigenes 数目与总 unigenes 数目之比) 为 6.9%, 其中 357 条 unigenes 含 1 个以上 SSR 位点, 4 056 条 unigenes 数目含单个 SSR 位点, 共检出 4 800

表 1 不同来源的鱼腥草材料

Table 1 *H. cordata* from different sources

编号	来源	编号	来源
1	云南昆明	9	广西陆川
2	贵州花溪	10	江苏连云港
3	四川乐山	11	湖北大悟
4	四川峨眉山	12	湖南株洲
5	重庆	13	江西瑞昌
6	陕西西安	14	湖南溆浦
7	浙江兰溪	15	湖南鹤城杨村
8	广西博白	16	湖南鹤城杨村

个 SSR 位点, 复合 SSR 位点 164 个。转录组中 SSR 位点的出现频率 (SSR 数目与总 unigene 的数目比值) 为 7.51%。从分布情况看, 鱼腥草转录组中平均 9.04 kb 有 1 个 SSR。但不同重复基序长度 SSR 分布的平均距离差别巨大, 其中三核苷酸重复最多 (1 994 个), 分布平均距离 21.76 kb/SSR, Quad-nucleotide repeat 最少 (39 个), 每条 SSR 分布平均距离 1 112.7 kb (表 2)。鱼腥草转录组 SSR 位点的序列总长达 80 754 bp, 其中单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸重复基元的 SSR 位点的碱基总长分别是 17 587、25 362、33 012、800、1 365、2 628 bp。各类型 SSR 位点的平均长度分别是 13、19、17、21、21、24 bp (表 2)。

2.2 鱼腥草转录组 SSR 的特性

鱼腥草转录组 SSR 种类丰富, 一至六核苷酸

重复类型都能看到, 但各类型出现频率不同, 主要是单核苷酸、二核苷酸和三核苷酸重复, 占总 SSR 的 95.52%, 其中三核苷酸重复最多, 占 41.54%, 其次是二核苷酸重复和单核苷酸重复类型, 分别占 26.63%和 27.35%, 四核苷酸、五核苷酸和六核苷酸重复类型的数量较少, 分别占 0.81%、1.38%、2.29%。从重复次数来看, 发现重复基元以重复 5 次出现的频率最高, 有 1 426 个, 占总 SSRs 的 29.71%, 其次是重复 6 次、≥13 次、12 次、7 次和 8 次, 分别有 980、710、608、492、223 个 SSR, 分别占总 SSR 的 20.42%、14.77%、12.59%、10.25%和 4.65%, 最少的为重复基元重复 11 次, 只有 80 个, 占 1.7% (表 3)。对鱼腥草转录组 SSR 统计分析, 发现共有 232 种重复基元, 其中单核苷酸的 A/T、G/C 重复基元 2 种。

表 2 SSR 在鱼腥草转录组中的出现频率

Table 2 Occurrence frequency of SSRs in *H. cordata* transcriptome

重复类型	数目	各类型比例/%	频率/%	平均距离/kb	总长度/bp	平均长度/bp
单核苷酸	1 313	27.35	2.05	33.05	17 587	13
二核苷酸	1 278	26.63	2.00	33.96	25 362	19
三核苷酸	1 994	41.54	3.12	21.76	33 012	17
四核苷酸	39	0.81	0.06	1 112.70	800	21
五核苷酸	66	1.38	0.10	657.55	1 365	21
六核苷酸	110	2.29	0.17	394.50	2 628	24

表 3 鱼腥草 SSR 不同基序长度和重复的数量分布

Table 3 Distribution of SSR with different motif types and repeat numbers in *H. cordata*

重复类型	重复次数								
	5	6	7	8	9	10	11	12	≥13
单核苷酸	—	—	—	—	—	—	—	604	709
二核苷酸	—	446	284	202	145	119	78	4	0
三核苷酸	1 275	481	208	21	3	0	2	0	4
四核苷酸	0	34	5	0	0	0	0	0	0
五核苷酸	57	9	0	0	0	0	0	0	0
六核苷酸	94	10	2	3	0	0	0	0	1

二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸重复基元分别有 12、58、26、52、84 种。从出现频率来看, 占优势的重复基元类型是单核苷酸重复基元, 其次是二核苷酸和三核苷酸重复基元, 其他四核苷酸、五核苷酸和六核苷酸重复基元类型分布较少, 出现频率较低 (图 1)。

对二核苷酸和三核苷酸重复基元深入分析, 按照 4 种碱基随机随机组合的方式统计后, 发现二核苷酸和三核苷酸重复基元分别为 4 种和 10 种, 即这 2 种核苷酸重复基元类型在鱼腥草 SSR 序列中全部存在。其中二核苷酸重复基元出现最多的是 AG/CT, 共有 1 108 个, 占该重复基元的 86.70%,

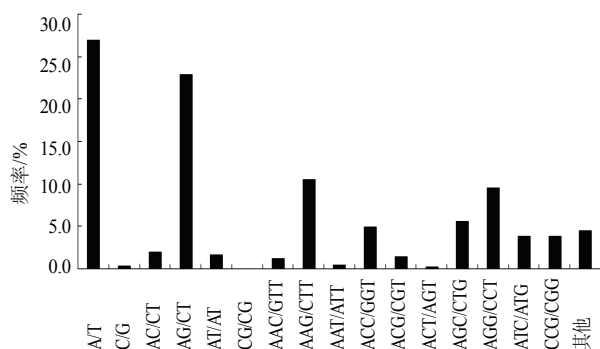


图 1 鱼腥草不同基序 SSR 类型的频率分布

Fig. 1 Frequency distribution of SSR based on different motif sequence types in *H. cordata*

其他的 2 种 AT/GC、AC/GT 则分别仅占 6.34% 和 6.96%；三核苷酸重复种类较多，最多的重复基元为 GAA/CTT，共有 336 个，GAA/CCT 也有 333 个，分别占三核苷酸重复基元的 16.85%、16.70%。其次的为 GGT/CCA（共有 168 个）、AGA/TCT（共有

153 个）、GAG/CTC（共有 153 个）、GCT/CGA（共有 89 个）、CAG/GTC（共有 62 个）。

2.3 SSR 引物筛选与验证

选择前后序列均不小于 150 bp 的 SSR 50 个，根据引物设计原则，应用软件 Primer3 设计引物，对所获得的 4 800 个 SSR 位点设计引物，共开发出 3 068 个候选 SSR 引物。随机挑选了 50 对引物，其中有 43 对引物能扩增出清晰条带，并表现出多态性差异，多态信息量（PIC）为 0.49~0.96，平均值为 0.72；等位基因数范围为 3~10，平均为 5.74，表观杂合度（ H_o ），期望杂合度（ H_e ）见表 4。利用 SSR 多态性对 16 个鱼腥草材料进行聚类分析，聚类结果显示，峨嵋蕺菜单独聚为一类；其他群体聚为一类。四川群体与其他各个省群体分别列为 2 个亚类；云南、贵州、重庆等的遗传距离较近聚为一群；湖南、湖北的遗传距离较近聚为一群；广西、江西、浙江等的遗传距离较近聚为一群（图 2）。

表 4 SSR-PCR 筛选引物

Table 4 Screening primers of SSR-PCR

引物	SSRs	正向引物 (5'-3')	反向引物 (5'-3')	等位基因数	H_o	H_e	PIC
HM_1	(TGA) ₅	GAAGCTTAAGGAGGTAGAGGCTG	CAGTAAGTTGTGTCCAAAGGTGC	3	1.00	0.59	0.49
HM_2	(AGA) ₆	GAAGGACTGCAAGAAACTCTGAA	CCCCATCTTCTGTCTCTTCTCTT	7	1.00	0.86	0.84
HM_3	(AC) ₇	AAACATGCATTGTACATACAGC	AGAGCTCGCGAAATATACTGTTG	6	0.89	0.93	0.93
HM_4	(GGA) ₅	CGACGACGATGAGACGGT	GAGCCACCCCATCTTTC	4	1.00	0.71	0.63
HM_5	(AG) ₁₀	AAGGGGAGAGAAAAGAAAGAGAGC	CCATAATAAAGCTCAATGCTGCT	6	1.00	0.84	0.79
HM_6	(CTC) ₅	CTACCAAACCCCTCTCCAAAAT	GCAGCAGCAGTAGAGGATTG	8	0.60	0.79	0.68
HM_7	(GCT) ₅	CTCGTTTCGGCAAACGCT	CGCGTATTATGAGCAGGATTACT	5	0.91	0.90	0.89
HM_8	(TCC) ₅	CAGTCCTGGTGGACATACTTGAT	CAATCGATGATGAGGAAGAAAAC	4	1.00	0.73	0.66
HM_9	(TGA) ₆	CTGCTGAGTCTGCATTAATTCCT	CAGCAGCTTCAGATACCTCAGA	7	1.00	0.79	0.75
HM_10	(AG) ₆	TATGGTGATAATAATCCGGTCG	TTCTCCCAAATGATGGATGATAA	8	0.67	0.80	0.72
HM_11	(TGG) ₅	GGTGCAGGTGTTGATGTCCT	CAAGGAGAAGTTCGAGGAGGT	10	0.65	0.82	0.73
HM_12	(TTC) ₆	GAGTCTTCCATTCTTTTGCTGA	GAGTGATGTGGAATGCTCTCTTC	4	1.00	0.76	0.70
HM_13	(TGA) ₇	GAGCATGACCTTAAAAAGGATGA	TAGACAAAAAGAATACCATCGGC	4	1.00	0.75	0.68
HM_14	(TA) ₇	CATGTCAACGTCATCTGTAGCAT	ATTTGGGTAGGGTACTGGAGCTA	5	0.80	0.69	0.65
HM_15	(CT) ₆	ACCATCTCCTCCGTTCCAC	GGAAGTAAAGATCGAGGAGGTCT	8	1.00	0.89	0.84
HM_16	(GCG) ₆	AGTGATACCTCCTTCTCCGCTA	AGCAACAGGTGACAACGTACC	7	0.63	0.82	0.65
HM_17	(TTC) ₅	GAGTGAAAGTGTTTGCTTTGGAG	CCAACAGAACCAAATAAAACCAA	6	1.00	0.85	0.80
HM_18	(ACT) ₅	AATTGCAACGCGTAATCATCTT	CTAACTTCGTCGAATTGGGTCTT	6	0.54	0.65	0.63
HM_19	(CCA) ₅	CCTCCTCCTCTATGGGGG	GTGGTGGCGGTCTATTTCTAA	4	0.58	0.96	0.96
HM_20	(TCA) ₅	TCAGTATTGAATTTGTCTCCCC	CTTGTTAAACGAGTGATATCGGG	7	0.88	0.79	0.78
HM_21	(GA) ₈	AAATAGAGTTACAGCCCCAAAA	CTTTGTACTTTTGCCTCCACTTT	8	0.64	0.83	0.76
HM_22	(TTC) ₅	TTCTGCTTCTTCCCCTTCTTATT	GCTTTTACAGAGACCTCTTGTC	4	0.50	0.67	0.62
HM_23	(TCT) ₆	GCATTACAAGAACCCAAACACAT	ACTACTGGTTCTAAGCGAGGAGC	4	1.00	0.68	0.60
HM_24	(GAA) ₅	AAATTTTAAACAGCAATTCCTCCG	TCCTTCTCTGATCTAGGGTTTCC	4	1.00	0.65	0.57
HM_25	(GCG) ₆	AAGGAGTCCCTGGATACTGCTAC	ACCACAAAACGAACGGATTC	7	0.40	0.61	0.52
HM_26	(CT) ₇	TTCTTCGCTTCTAGATTCCTTT	TCTAGCCAAAGTCTTCCTTCCTT	8	0.67	0.76	0.65
HM_27	(GA) ₈	ATCAAAGCTGCAAGCTTATATCG	TTGTTCAACCACTGGAATCACA	5	0.60	0.80	0.73
HM_28	(GGA) ₅	GGTCTACGGCAAACAGACTAGC	CGATCCTCCAACCTCAAAAAA	7	0.59	0.80	0.71

续表 4

引物	SSRs	正向引物 (5'-3')	反向引物 (5'-3')	等位基因数	Ho	He	PIC
HM_29	(AT) 8	CTAGCAGACTGGATGTTGGTTTT	GTAAGGCGCATAATACTTGTGTC	4	1.00	0.73	0.66
HM_30	(GCT) 5	AGATCCCCACCCTTTCTTGAT	ATGACTTGATGAAAAAGAGGCAA	4	1.00	0.75	0.68
HM_31	(GCCCCA) 4	GGCATTGCCTATAGAAGAAGTATCA	CATTGTCCACCTGGGTAAGAT	5	0.75	0.63	0.59
HM_32	(TGT) 7	GTACGAAGAGGAGGAATTGGG	TCAGAAAAGAAAAACCAGACCAA	8	1.00	0.87	0.82
HM_33	(TCC) 5	GTCAGATCGAGGAGCGTAAATC	CCTCCTGATGGAATGAAGATG	6	0.68	0.78	0.67
HM_34	(GAA) 7	GACGCAAGGTAACCCTCAATAAT	ATTGCTTTCACTCTGGTTGTCAT	5	0.52	0.71	0.64
HM_35	(GGCGAT) 7	GAGGGAAGAAGGCGATCAG	CTGTCTTCTCGAGGACTGCTCT	3	1.00	0.63	0.54
HM_36	(TC) 9	ATTCCCTTCTTCTCTGCCTCTC	CGTGCTCAAAGTAGCAGAATCA	3	0.86	0.91	0.90
HM_37	(AGA) 5	CCATAAATACATCACCTCGATCC	CCAGCAGAGTTCAACTTGTCTT	5	0.78	0.80	0.77
HM_38	(GA) 7	ATCTTATCCACAACAAGAAGCCA	AAGCTCATGAACATGACACACAG	3	0.91	0.90	0.89
HM_39	(TC) 6	CTCAGGAAAAAGTAACACTGGCAT	GAGACTTGTCTCCTGCATCAACT	7	1.00	0.85	0.80
HM_40	(GA) 9	AGCCTTCAGGATTACTTCTCCAC	AAGATGTCTTCAATCTTCACCCA	5	0.80	0.69	0.65
HM_41	(ACC) 6	GCAATTTGTTTGGATTGACAAAG	TGGAGTGCTTTTATGCAAATTGT	6	1.00	0.79	0.75
HM_42	(TAT) 5	GAAGAAGATGATGCTGATGGTTC	ATTCTGTGCAGCTAGTTACCACC	9	0.90	0.89	0.83
HM_43	(CAC) 7	CTGTTGAGGCTGAGGATATTGAG	AAACGCGTCAAAGAAGAACATC	8	1.00	0.86	0.81
均值				5.74	0.83	0.78	0.72

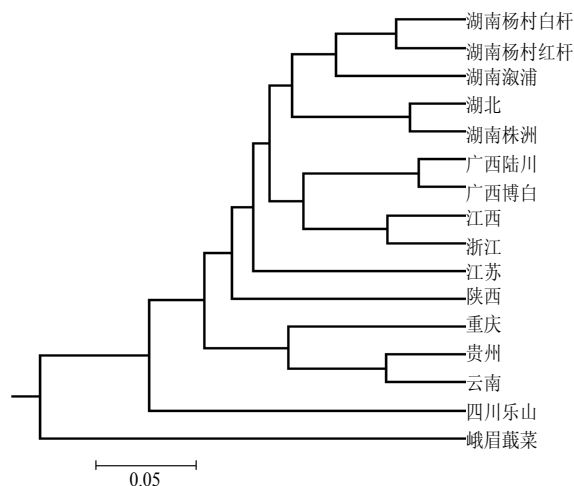


图 2 鱼腥草的 UPGMA 聚类图

Fig. 2 UPGMA dendrogram for *H. cordata*

3 讨论

基于 PCR 技术的分子标记,较常用的有随机扩增多态 DNA (RAPD)^[9]、扩增片段长度多态性 (AFLP)^[10]和扩增简单重复序列间区 (ISSR)^[11]等。这些标记因其特异性不强而广泛应用于物种的遗传多样性与亲缘关系等方面的研究,尤其对不具任何基因组信息的非模式生物的研究,凸显其优势。对不具任何基因组信息的鱼腥草,目前已利用 RAPD^[5-6]、AFLP^[12]、ISSR^[7]等分子标记进行了研究,由于这些标记的扩增产物是不确定、非特异的,因而用它们得出的结果不能有效地应用于遗传研究和标记辅助育种。SSR 分子标记是特异性较强的标记,能有效地克服不确定、非特异的缺点,而且对已开发出大量特异标记的植物,已很少使用非特异标记

对其进行遗传多样性与亲缘关系等方面的研究。为了更加有效的开发和利用药用植物鱼腥草的遗传资源,完全有必要开发其特异性较强的分子标记,而且迄今尚未见鱼腥草植物中开发 SSR 标记的报道。

从鱼腥草转录组中搜索出 4 800 个 SSRs,平均出现频率是 1/9.04 kb,和其他物种相比,鱼腥草 SSRs 出现频率高于小麦(1/15.6 kb)^[13]、拟南芥(1/13.83kb)、番茄 (1/11.1 kb)、棉花 (1/20.0kb)^[14]、杨树 (1/14.0kb)^[15]等植物,表明鱼腥草转录组中 SSR 数量相当丰富。从鱼腥草 SSR 重复基元分析,鱼腥草 SSR 与其他植物相比存在相似性和差异性。其相似性表现在鱼腥草 SSR 序列的重复基元种类繁多,单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸、六核苷酸重复基元等。而重复基元主要集中在单核苷酸、二核苷酸和三核苷酸重复。其二核苷酸重复基元以 AG/CT 为主 (占总数的 86.70%),这与烟草^[16]、茶树^[15]、马尾松^[17]、杏树^[18]等的 SSR 序列二核苷酸重复基元的情况相似。鱼腥草三核苷酸重复以 GAA/CTT 重复为主 (占 16.85%),这与拟南芥^[19]、柑橘^[20]、烟草^[21]等报道的情况基本一致,这可能是这些重复基元与其关键蛋白编码的密码子使用频率有较大的关系^[22]。鱼腥草 SSR 重复基元与其他物种的情况也存在一定的差异性,鱼腥草 SSR 序列以三核苷酸重复基元重复为主,与茶树^[15]、烟草^[16]等有所差异。

鱼腥草转录组 SSR 不但出现频率高,而且类型丰富;利用 unigenes 序列开发鱼腥草的 SSR 标记进行试验验证,所开发的 SSR 标记在鱼腥草基因组中

的检测结果多表现为清晰可读的单一主带或 2 条带,表明被检测的标记在基因组 DNA 中的位置是特异的,能有效反映其多态性,从多态性潜能的角度考虑,搜索到的这些 SSR 也具有较高的可用性。可成为鱼腥草基因组 DNA 遗传多样及其分子生物学其他相关研究的有力工具。本研究结果为进一步发掘新的鱼腥草功能基因 SSR 标记奠定了基础,这种标记的建立对于加速鱼腥草功能基因资源的开发利用、遗传资源评价及实现特定性状的辅助选择标记育种和开展鱼腥草的比较基因组学研究均具有重要意义与应用价值。

参考文献

- [1] 马林, 吴丰, 陈若芸. 三白草科植物化学及生物活性研究进展 [J]. 中国中药杂志, 2003, 28(3): 196-198.
- [2] 中国药典 [S]. 一部. 2015.
- [3] 伍贤进, 张俭, 钟晓丽, 等. 鱼腥草储藏过程中化学成分含量的变化 [J]. 食品工业科技, 2007, 28(5): 215-217.
- [4] 方宣钧, 吴为人, 唐纪良. 作物 DNA 标记辅助育种 [M]. 北京: 科学出版社, 2001.
- [5] 吴卫, 郑有良, 陈黎, 等. 鱼腥草种质资源的 RAPD 分析 [J]. 药学学报, 2002, 37(12): 986-992.
- [6] 蓝云龙, 吴令上, 裘波音, 等. 鱼腥草 RAPD 分子标记的多态性 [J]. 浙江林学院学报, 2008, 25(3): 309-313.
- [7] Wei L, Wu X J. Genetic variation and population differentiation in a Medical Herb *Houttuynia cordata* in China revealed by ISSRs [J]. *Int J Mol Sci*, 2012, 13(16): 8159-8170.
- [8] 易官美, 黎建辉, 王冬梅, 等. 南方红豆杉 SSR 分布特征分析及分子标记的开发 [J]. 园艺学报, 2013, 40(3): 571-578.
- [9] Williams J G K, Kubelik A R, Livak K J, *et al.* DNA polymorphism amplified by arbitrary primers are useful as genetic markers [J]. *Nucl Acids Res*, 1990, 18(13): 6531-6535.
- [10] Vos P, Hogers R, Bleeker M, *et al.* AFLP: A new technique for DNA fingerprinting [J]. *Nucl Acid Res*, 1995, 23: 4407-4414.
- [11] 黎晓英, 魏麟, 伍贤进, 等. 中国不同地理居群鱼腥草遗传多样性分析 [J]. 中草药, 2010, 41(2): 285-288.
- [12] Zietkiewicz E, Rafalski A, Labuda D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification [J]. *Genomics*, 2003, 19(20): 176-183.
- [13] Kantety R V, Rota M L, Matthews D E, *et al.* Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat [J]. *Plant Mol Biol*, 2002, 48(9): 501-506.
- [14] Gao L F, Tang J F, Li H W, *et al.* Analysis of microsatellites in major crops assessed by computational and experimental approaches [J]. *Mol Breed*, 2003, 12(13): 245-256.
- [15] 金基强, 崔海瑞, 陈文岳, 等. 茶树 EST-SSR 的信息分析与标记建立 [J]. 茶叶科学, 2006, 26(1): 17-23.
- [16] 赵世浩. 烟草 EST-SSR 标记的建立及其应用 [D]. 长沙: 湖南农业大学, 2007.
- [17] 刘公秉, 季孔庶. 基于松树 EST 序列的马尾松 SSR 引物开发 [J]. 分子植物育种, 2009, 7(4): 833-838.
- [18] Jung S, Abbott A, Jesudurai C, *et al.* Frequency type distribution and annotation of simple sequence repeats in Rosaceae ESTs [J]. *Funct Integr Genomics*, 2005, 19(5): 136-143.
- [19] Morgante M, Hanafey Wayne Powell. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes [J]. *Nat Genet*, 2000, 30(16): 194-200.
- [20] Chen C X, Zhou P, Choi YA, *et al.* Mining and characterizing microsatellites from citrus ESTs [J]. *Theor Appl Genet*, 2006, 112(19): 1248-1257.
- [21] 赵世浩. 烟草 EST-SSR 标记的建立及其应用 [D]. 长沙: 湖南农业大学, 2007.
- [22] 范三红, 郭蔼光, 单丽伟, 等. 拟南芥基因密码子偏爱性分析 [J]. 生物化学与生物物理进展, 2003, 23(30): 221-225.