

基于 cp DNA 的当归野生与栽培种质鉴定与遗传变异分析

张明惠¹, 朱田田^{1,2,3*}, 晋玲^{1,2,3}, 王富胜⁴, 康舒淇¹, 徐丽¹, 王圆圆¹

(1. 甘肃中医药大学, 兰州 730000; 2. 西北中藏药省部共建协同创新中心, 兰州 730000;
3. 甘肃省珍稀中药资源评价与保护利用工程研究中心, 兰州 730000;
4. 定西市农业科学研究院, 甘肃定西 743000)

[摘要] 目的: 基于叶绿体基因(cp DNA)对甘肃省 11 个当归栽培品种(系)及 7 个野生当归居群进行了遗传变异分析, 为当归的种质鉴定和新品种的选育提供参考。方法: 利用 3 对 cp DNA 引物对当归样品进行聚合酶链式反应(PCR)扩增并测序, 利用 MegaX 软件对序列特征进行统计, 并计算当归居群间平均遗传距离, 利用 NTSYS 2.10e 软件构建遗传距离非加权组平均法(UPGMA)聚类图。利用 DanSP v6 软件计算当归序列多态性及 Tajima 中性检验; 利用 PERMUT 软件计算当归居群结构; 利用 Arlequin v3.5 软件进行分子变异分析; 最后利用 PopART 1.7 软件构建 TCS 单倍型网络图。结果: 3 对 cp DNA 引物扩增、测序、比对、合并后的序列长度为 1 759 bp, 野生当归检测到 1 个变异位点, 栽培当归检测到 480 个变异位点, 其中单一突变位点 97 个, 简约信息位点 383 个, 插入-缺失位点 152 个。在野生组和栽培组当归 cp DNA 的 *matK*、*psbA-trnH* 和 *rbcL* 3 个区域中, 多态性最高的区域均为 *matK*。全部当归联合序列的 Tajima 中性检验均为不显著负值, 但在栽培当归中 *psbA-trnH* 和 *rbcL* 基因中呈显著负值, 说明当归整体上遵循中性进化, 而 *psbA-trnH* 和 *rbcL* 基因经历过选择。野生居群间的遗传分化程度($F_{st}=0$)低于栽培当归居群间的分化程度($F_{st}=0.114\ 19, P<0.05$), 且二者之间遗传分化程度较高($F_{st}=0.942\ 55, P<0.01$)。栽培当归居群中变异主要来源于居群内部(89%)。遗传距离 UPGMA 聚类树显示野生当归和栽培当归各自聚为一支, 亲缘关系较远, 栽培当归中居群 3 距离其他栽培当归居群较远。TCS 单倍型网络图由 15 个单倍型和 4 个未知单倍型构成, 分为 3 部分, 各部分间变异次数较多, 共享单倍型仅分布于野生或栽培组之内, 组间不存在共享单倍型。结论: 当归在物种水平上遗传多样性偏低, 野生当归居群多样性低于栽培当归, 野生与栽培当归组间的遗传分化程度高, 而野生当归居群内和栽培当归居群内的分化程度均较低, 栽培当归中遗传变异主要存在于居群内。

[关键词] 当归; 叶绿体基因(cpDNA); 野生; 栽培; 遗传变异

[中图分类号] R284.2; R289; R22; R2-031; R33 **[文献标识码]** A **[文章编号]** 1005-9903(2022)15-0129-10

[doi] 10.13422/j.cnki.syfjx.20220915

[网络出版地址] <https://kns.cnki.net/kcms/detail/11.3495.r.20220531.1559.001.html>

[网络出版日期] 2022-06-02 8:40

Identification and Genetic Variation Analysis of Wild and Cultivated Germplasm of *Angelica sinensis* Based on cp DNA

ZHANG Minghui¹, ZHU Tiantian^{1,2,3*}, JIN Ling^{1,2,3}, WANG Fusheng⁴,
KANG Shuqi¹, XU Li¹, WANG Yuanyuan¹

(1. Gansu University of Chinese Medicine, Lanzhou 730000, China;
2. Northwest Collaborative Innovation Center for Traditional Chinese Medicine (TCM)
Co-constructed by Gansu Province & MOE of PRC, Lanzhou 730000, China;
3. Engineering Research Center for Evaluation, Protection, and Utilization of
Rare TCM Resources, Lanzhou 730000, China;

[收稿日期] 2022-03-21

[基金项目] 甘肃省科技厅创新基地和人才计划目(20JR5RA182); 道地药材生态种植及质量保障项目(国中医药科技[2020]153号); 甘肃省教育厅“双一流”科研重点项目(GSSYLXM-05); 甘肃中医药大学科学研究与创新基金项目(2021KCZD-4); 中国工程院战略研究与咨询项目(GS2021ZDA06-2)

[第一作者] 张明惠, 在读硕士, 从事中药资源保护、评价与可持续利用研究, E-mail: 2244670610@qq.com

[通信作者] * 朱田田, 副教授, 硕士生导师, 从事中药资源评价与分子生药学研究, Tel: 0931-5161169, E-mail: ztt0935@163.com

4. Dingxi Academy of Agricultural Sciences, Dingxi 743000, China)

[Abstract] **Objective:** To conduct genetic variation analysis of 11 cultivars and 7 wild populations of *Angelica sinensis* in Gansu province based on the chloroplast gene (cp DNA), and provide references for germplasm identification and breeding of new cultivars of *A. sinensis*. **Method:** Three pairs of cp DNA primers were used for polymerase chain reaction (PCR) amplification and sequencing of *A. sinensis* samples. MegaX was used to perform statistics on sequence characteristics and calculate mean genetic distances among *A. sinensis* populations. Unweighted pair-group method with arithmetic means (UPGMA) clustering tree based on genetic distance was constructed by NTSYS 2.10e. DanSP v6 was used to calculate sequence polymorphism and Tajima's D of *A. sinensis*. PERMUT was used to calculate the population structure of *A. sinensis*. Arlequin v3.5 was used to perform molecular variation analysis, and PopART1.7 was used to construct TCS haplotype network. **Result:** Three pairs of cp DNA primers were amplified, sequenced, compared, and combined to give a sequence length of 1 759 bp. One variable site was detected in the wild *A. sinensis* and 480 variable sites were detected in the cultivated *A. sinensis*, including 97 singleton variable sites, 383 parsimony informative sites, and 152 insertion-deletion sites. In the three regions of *matK*, *psbA-trnH*, and *rbcL* of cp DNA in the wild and cultivated *A. sinensis*, *matK* was the region with the highest polymorphism. Tajima's D of all the combined sequences of *A. sinensis* were not significantly negative, but *psbA-trnH* and *rbcL* genes of the cultivated *A. sinensis* were significantly negative, indicating that the *A. sinensis* followed neutral evolution on a whole, while *psbA-trnH* and *rbcL* genes had undergone selection. The degree of genetic differentiation ($F_{st}=0$) among wild populations was lower than that among cultivated populations ($F_{st}=0.114\ 19$, $P<0.05$). The degree of genetic differentiation between wild and cultivated *A. sinensis* was relatively high ($F_{st}=0.942\ 55$, $P<0.01$). Genetic variation in the cultivated *A. sinensis* was mainly found within the populations (89%). UPGMA cluster tree based on genetic distance showed that the wild *A. sinensis* and the cultivated *A. sinensis* were clustered into one branch, respectively, with a distant genetic relationship, and the population 3 in the cultivated *A. sinensis* was far from other cultivated populations. The TCS haplotype network consisted of 15 haplotypes and 4 unknown haplotypes, which was divided into 3 parts, with a large number of variations among each part. Shared haplotypes were only found in the wild or cultivated groups, and there were no shared haplotypes between groups. **Conclusion:** The genetic diversity of *A. sinensis* was low at species level, and the population diversity of the wild was lower than that of the cultivated. The degree of genetic differentiation between the wild and the cultivated *A. sinensis* was high, but that in the wild and the cultivated populations were low. Genetic variation in the cultivated *A. sinensis* was mainly found within populations.

[Keywords] *Angelica sinensis*; chloroplast DNA(cp DNA); wild; cultivated; genetic variation

当归来源于伞形科植物当归的干燥根,具有补血活血、调经止痛、润肠通便的功效^[1]。据《中国植物志》记载,当归主产甘肃东南部,以岷县产量多,质量好,其次为云南、四川、陕西、湖北等省,均为栽培^[2],野生资源极少^[3-5]。调查发现当归是典型的忌连作药用植物,各道地产区均已出现连作障碍问题,麻口病和根腐病发生严重,而野生当归尽管常年生长在同一地域内,却未发现野生当归病虫害危害植株,表明当归野生种具有较强的抗病虫害能力^[3]。

目前关于野生和栽培当归的研究主要集中在外观形态^[3-7]、种群特性^[8-9],少数涉及成分分析^[10],

对分子水平的研究还较少。冯伟萌等^[11]利用高通量测序手段获得当归全长转录本信息,通过比较野生当归和栽培当归在转录本层次的差异表达,为当归优良种质资源的筛选培育、抗性研究和次生代谢途径解析提供基础信息。本课题组前期应用内部简单重复序列(ISSR)分子标记比较了野生当归与栽培当归的遗传多样性,探讨遗传多样性对当归资源的保护意义^[12]。叶绿体基因(cp DNA)组序列较短,核酸置换率适中,多为单拷贝序列,结构比较保守,在分子水平上具有明显差异,在研究种及种以上较高阶元的系统发育关系上优势明显^[13]。在植

物系统发育、遗传多样性和亲缘演化关系等研究方面发挥着重要作用^[14]。本研究利用3对cp DNA引物扩增来自甘肃省的11个栽培当归品种(系)与7个野生当归居群,分析3个cp DNA序列片段的遗传变异,解析当归居群在cp DNA水平上的遗传多样性,探讨居群间的系统演化关系,为当归的种质鉴定和新品种的选育提供参考。

1 材料

1.1 植物 野生当归于2013年8月采自甘肃岷县和迭部县,共有7个居群12份样本;栽培当归于

2019年7月采自甘肃省定西市渭源县当归育苗基地,共有11个品种(系)164份样本,取样时挑选长势良好的植株,部分材料因受自然灾害缺失,故样本数量较少,所有样品经甘肃中医药大学中药资源学教研室晋玲教授鉴定为伞形科当归属当归 *Angelica sinensis* 的叶片。同时采集了西藏凹乳芹 (*Vicatia tibetica*)、欧当归 (*Levisticum officinale*)、东当归 (*A. acutiloba*)、牡丹叶当归 (*A. paeoniifolia*)、白芷 (*A. dahurica*) 的样品作为外类群,以上样品均经晋玲教授鉴定。具体信息见表1。

表1 当归样品信息

Table 1 Sample collection information

居群编号	名称	采集地	经度	纬度	海拔/m	数量/个	类型
1	岷归1号	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	18	栽培
2	岷归2号	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	16	栽培
3	岷归3号	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	3	栽培
4	岷归4号	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	18	栽培
5	岷归5号	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	17	栽培
6	岷归6号	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	25	栽培
7	D42017-01	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	26	栽培
8	半紫半绿茎	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	25	栽培
9	10G 辐照	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	4	栽培
10	20G 辐照	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	7	栽培
11	30G 辐照	甘肃省定西市渭源县	103°58'32"	35°05'09"	2 322	5	栽培
12	当归	甘肃省定西市岷县	104°49'34"	35°20'03"	2 179	1	野生
13	当归	甘肃省定西市岷县	104°47'54"	34°20'05"	2 335	2	野生
14	当归	甘肃省定西市岷县	103°56'52"	34°16'35"	2 936	2	野生
15	当归	甘肃省定西市岷县	104°06'03"	34°15'42"	2 614	2	野生
16	当归	甘肃省定西市岷县	103°56'52"	34°16'34"	2 964	2	野生
17	当归	甘肃省定西市岷县	104°49'34"	34°20'30"	2 180	2	野生
18	当归	甘肃省甘南藏族自治州迭部县	103°36'13"	33°53'02"	2 602	1	野生
19	西藏凹乳芹	西藏自治区林芝市米林县	94°21'16"	29°37'29"	2 989	2	栽培
20	欧当归	甘肃省定西市陇西县	104°24'52"	35°05'19"	1 910	2	栽培
21	东当归	吉林农业大学药圃	125°24'13"	43°48'38"	222	2	野生
22	牡丹叶当归	西藏自治区林芝市巴宜区	94°29'57"	27°36'24"	3 146	2	野生
23	白芷	甘肃省定西市陇西县	104°24'52"	35°05'19"	1 910	2	栽培

1.2 仪器与试剂 SureCycler 8800型梯度聚合酶链式反应(PCR)扩增仪(美国Agilent Technologies); Gel Doc XR⁺型凝胶成像仪(美国Bio-Rad); Eppendorf 5418R型台式高速冷冻离心机(德国Eppendorf); 微DYY-T型转移电泳仪(北京六一仪器厂); HH-S₂₄型电热恒温水浴锅(上海一恒科学仪器有限公司); WFH-201-B型紫外透射反射仪

(上海精密仪器仪表有限公司); BCD-451WDEMU1型冰箱(青岛海尔股份有限公司)。

NEP003-2型植物基因组提取试剂盒(北京鼎国昌盛科生物技术有限责任公司,批号B00B00103), PCR试剂β-巯基乙醇、Agarose琼脂糖、牛血清白蛋白(BSA)、引物、Taq DNA Polymerase、dNTPs、GoldView核酸染料[(上海)生工生物工程股份有限

公司,批号分别为20191103、0000590598、20191203、W9910e、R6409、321523BB],DNA Marker(大连宝生物股份有限公司,批号20190601);乙醇、三氯甲烷、异戊醇、异丙醇(国产分析纯,天津市大茂化学试剂厂,批号分别为20190901、20190826、20180701、20170701),水为屈臣氏纯净水。

2 方法

2.1 基因组DNA提取与纯度和浓度检测 采用试剂盒法提取当归DNA,参照课题组前期研究中的试验方法^[15]。采用1.0%琼脂糖凝胶电泳,电压110 V下20 min检测DNA浓度,根据凝胶成像系统拍照后的DNA条带的亮度、整齐度、是否有拖尾等判断DNA纯度,检测合格后的DNA进行PCR扩增。

表2 3对cp DNA引物基本信息

Table 2 3 pairs of cp DNA primers basic information

cp DNA 间区	上游(5'-3')	下游(5'-3')	退火温度/°C	长度/bp
<i>matK3F-matK1R</i>	CGTACAGTACTTTTGTGTTTACGAG	ACCCAGTCCATCTGGAAATCTTGGTTC	50	756-837
<i>psbA-trnH</i>	GTTATGCATGAACGTAATGCTC	CGCGCATGGTGGATTCACAATCC	58	211-221
<i>rbcL1F-rbcL724R</i>	ATGTCACCACAAACAGAAAC	TCGCATGTACCTGCAGTAGC	56	672-701

2.3 数据统计与分析 利用Chromas 2.6.5软件和ContigExpress软件将测序所得的序列文件进行校对峰图并去除低质量的碱基,并将正反向序列拼接,人工校正无法判读的碱基^[21]。利用MegaX软件^[22]中ClustalW程序分别将3对引物扩增出的序列进行比对,对当归序列的基本特点进行统计分析。利用PhyloSuite软件^[23]将3个cp DNA序列片段进行拼接;利用Kimura2-Parameter模型^[24]计算当归居群间平均遗传距离,利用NTSYS 2.10e软件^[25]构建非加权组配对算术平均法(UPGMA)聚类树。利用DanSP v6软件^[26]对当归所有个体的序列片段及拼接片段进行多态性计算,并进行Tajima中性检验^[27]。利用PERMUT(2.0版)软件^[28]计算当归居群

表3 当归cp DNA片段的序列特征

Table 3 Sequence characteristics of DNA fragments in *Angelica sinensis*

cp DNA	扩增片段长度/bp	A/%	T/%	C/%	G/%	(A+T)/%
<i>matK3F-matK1R</i>	837	29.5	35.8	17.9	16.8	65.3
<i>psbA-trnH</i>	221	36.9	31.3	12.2	19.6	68.2
<i>rbcL1F-rbcL724R</i>	701	28.2	28.1	22.6	21.1	56.3
combine	1 759	29.9	32.1	19.1	18.9	57.6

3.2 当归cp DNA多态性 野生当归居群中,核苷酸多样性最高的是*matK3F-matK1R*序列($Pi=0.00022$),检测到1个变异位点,1个单一突变位点,

2.2 PCR扩增和测序 本实验所用3对cp DNA引物^[16-20]见表2,利用单因素法对扩增体系中DNA模板浓度、引物浓度、*Taq*酶浓度进行条件优化,最终确定当归PCR最佳反应体系(25 μ L):DNA模板1.0 μ L、上下游引物各0.5 μ L、*Taq*DNA酶0.25 μ L、10 \times Ex*Taq* buffer 2.5 μ L、dNTPs 2.0 μ L、BSA 0.5 μ L、ddH₂O 17.75 μ L。扩增程序95 $^{\circ}$ C预变性5 min,95 $^{\circ}$ C变性30 s,根据不同引物的退火温度复性1 min,72 $^{\circ}$ C延伸1 min 40 s,35次循环,72 $^{\circ}$ C延伸7 min,4 $^{\circ}$ C保存结束反应。各引物退火温度见表2。扩增产物用0.8%的琼脂糖凝胶电泳进行检测,将条带明亮、单一的PCR产物送至生工生物工程(上海)股份有限公司进行双向测序。

内平均遗传多样性(HS)、总遗传多样性(HT)及居群遗传分化系数Gst和Nst;利用Arlequin v3.5^[29]软件进行分子变异分析(AMOVA),用20 000次排列来检测不同分组变异的差异显著性,从组间、组内居群间及居群内评价当归遗传分化。用PopART 1.7^[30]软件选择TCS^[31]模型构建cp DNA单倍型网络图。

3 结果与分析

3.1 当归cp DNA的序列特征 18个居群176个样本被成功扩增和测序,将全部个体测序获得的3个cp DNA片段进行比对,得到这3个片段及合并序列的序列特征,见表3。在4条序列中A+T占有所有组成的比率分别为65.3%、68.2%、56.3%、57.6%。

0个简约信息位点,0个插入-缺失位点;*psbA-trnH*和*rbcL1F-rbcL724R*序列均未检测到变异,见表4。在cp DNA的3个区域中,单倍型数目分别为2、1、

1个,以 *matK3F-matK1R* 区域单倍型多样性(Hd=0.167)最高, *psbA-trnH* 和 *rbcL1F-rbcL724R* 序列单倍型多样性均为0。野生当归居群3个cp DNA基

因片段拼接后共形成2个单倍型,其单倍型多样性、单倍型多样性方差和标准差分别为0.167、0.018 04、0.134,见表5。

表4 当归3个cp DNA区域多态性信息

Table 4 Polymorphism information of three cp DNA regions in *Angelica sinensis*

类型	cp DNA 间区	变异位点	单一突变位点	简约信息位点	插入/缺失位点	核苷酸多样性	平均核苷酸差异
野生	<i>matK3F-matK1R</i>	1	1	0	0	0.000 22	0.167
	<i>psbA-trnH</i>	0	0	0	0	0	0
	<i>rbcL1F-rbcL724R</i>	0	0	0	0	0	0
	combine	1	1	0	0	0.000 10	0.167
栽培	<i>matK3F-matK1R</i>	383	5	378	118	0.060 80	43.712
	<i>psbA-trnH</i>	88	88	0	15	0.005 27	1.085
	<i>rbcL1F-rbcL724R</i>	7	7	0	15	0.000 13	0.085
	combine	480	97	383	152	0.028 62	45.422

表5 当归单倍型多样性信息

Table 5 Haplotype diversity information of *Angelica sinensis*

类型	cp DNA 间区	单倍型数目	单倍型多样性	单倍型多样性方差	单倍型多样性标准差
野生	<i>matK3F-matK1R</i>	2	0.167	0.018 04	0.134
	<i>psbA-trnH</i>	1	0	0	0
	<i>rbcL1F-rbcL724R</i>	1	0	0	0
	combine	2	0.167	0.018 04	0.134
栽培	<i>matK3F-matK1R</i>	10	0.347	0.002 26	0.048
	<i>psbA-trnH</i>	3	0.024	0.000 28	0.017
	<i>rbcL1F-rbcL724R</i>	2	0.012	0.000 14	0.012
	combine	13	0.367	0.002 31	0.048

栽培当归居群中,核苷酸多样性最高的是 *matK3F-matK1R* 序列(Pi=0.060 80),检测到383个变异位点,5个单一突变位点,378个简约信息位点,118个插入-缺失位点;核苷酸多样性最低的是 *rbcL1F-rbcL724R* 序列(Pi=0.000 13),检测到7个变异位点,7个单一突变位点,0个简约信息位点,15个插入-缺失位点,见表4。3个cp DNA基因片段中,单倍型数目分别为10、3、2个,以 *matK3F-matK1R* 区域单倍型多样性(Hd=0.347)最高, *RbcL1F-RbcL724R* 单倍型多样性最低(Hd=0.012)。栽培当归3个cp DNA基因片段拼接后共形成13个单倍型,其单倍型多样性、单倍型多样性方差和标准差分别为0.367、0.002 31、0.048,见表5。

对当归3个cp DNA联合序列的所有个体进行中性检验,栽培当归3个序列及联合序列的Tajima's D值均为负值, *matK3F-matK1R* 序列与联合序列均差异无统计学意义, *psbA-trnH* 序列为极显著负值(P<0.001), *rbcL1F-rbcL724R* 序列为显著负值(P<

0.05)。野生当归的 *matK3F-matK1R* 及联合序列的Tajima's D值均为不显著负值; *psbA-trnH* 和 *rbcL1F-rbcL724R* 序列无多态性,无法计算Tajima's D值。见表6。

表6 当归cp DNA中性检验

Table 6 Tajima's test results of cp DNA in *Angelica sinensis*

类型	cp DNA 间区	Tajima's D 值	P
野生	<i>matK3F-matK1R</i>	-1.140 53	>0.10
	<i>psbA-trnH</i>	-	-
	<i>rbcL1F-rbcL724R</i>	-	-
	combine	-1.140 53	>0.10
栽培	<i>matK3F-matK1R</i>	-1.176 59	>0.10
	<i>psbA-trnH</i>	-2.914 51	<0.001
	<i>rbcL1F-rbcL724R</i>	-2.030 40	<0.05
	combine	-1.531 87	>0.10

3.3 居群遗传结构 去除样本量为1的野生当归居群,基于cp DNA数据计算得出野生当归居群内

平均遗传多样性(HS)为0.200、总遗传多样性(HT)为0.200,两个遗传分化系数Gst和Nst均为0。基于cpDNA数据计算得出栽培当归居群内HS为0.441、HT为0.510,两个遗传分化系数Gst和Nst分别为0.136和0.278, $Nst > Gst (P < 0.01)$ 说明当归中亲缘关系较近的单倍型发生在同样居群中的概率极高。

3.4 当归居群间的遗传分化 当归居群间分化系数见表7,除去两个样本量为1的野生当归居群,在剩余的5个野生居群中,各居群之间的分化系数均为0。在栽培当归居群中,居群3与居群9分化不明显,其他栽培居群分化均较大,分化系数(FST)最低为0.3037(居群4),最高为0.4987(居群10、居群11);分化程度适中的有居群4与居群5、居群7、居群10、居群11, FST依次为0.0574、0.0590、0.0582、0.0582;居群5与居群7、居群10、居群11, FST依次为0.0861、0.0625、0.0625;居群7与居群10、居群11, FST均为0.1390;其他居群分化均不明显。野生当归居群与栽培当归居群的FST范围为0.4491~0.7881,说明其分化程度较高。

当归18个居群的176个样本分为野生和栽培两组,进行不同水平的分子变异分析结果显示94.25%的变异来源于组间,0.46%的变异来源于组内居群间,5.28%的变异来源于居群间,说明当归物种的变异主要来自于野生与栽培之间,见表8。总的遗传分化系数FST为0.94718,说明当归居群间存在较高的遗传分化。栽培当归居群中11%的变异来源于居群间,89%的变异来源于居群内,说明栽培当归的变异主要来自于居群内,栽培当归的FST为0.11419,分化程度远低于总的遗传分化系数。

3.5 当归居群间的遗传关系 基于cpDNA联合序列片段的遗传距离矩阵见表7,由遗传距离矩阵可以看出,野生当归居群间的遗传距离极低,在0~0.0003。当归11个栽培居群中居群3距离其他栽培居群最远,与其他居群的遗传距离为0.1658~0.1902,平均为0.1780;其次为居群9,与其他居群的遗传距离为0.0708~0.1658,平均为0.1183;剩余9个栽培居群两两之间的遗传距离均低于0.050。当归野生居群距离栽培居群中的居群3和居群9稍近,遗传距离分别为0.4491和0.6595,距离其他9个栽培居群均>0.700,野生当归距离栽培当归居群7最远达到了0.7881,当归野生居群与栽培居群之间的遗传距离平均为0.7288,远大于野生居群内部及栽培居群内部。

由于野生当归居群之间遗传距离极低,在构建遗传距离UPGMA聚类树时将居群12至居群18合并为一个居群,见图1。聚类树在0.75处分为2支,外类群与野生当归聚为一支,栽培当归聚为一支;在栽培当归中,居群3与其他栽培当归居群亲缘关系较远,在0.19处分开;居群5、居群10、居群11和居群7亲缘关系极近,聚为一支。

3.6 单倍型网络图 图2为176个当归样本的cpDNA单倍型中介邻接网络图。本图由15个单倍型和4个未知单倍型构成,两两单倍型之间连接线的长度上的数字代表突变的次数。第一部分的单倍型存在于野生当归的样本中;第二部分是由单倍型2(Hap2)经375次突变成Hap5,再由Hap5经过数次变异形成了Hap9~Hap11;第三部分是由第二部分Hap5经过多次变异形成Hap6, Hap6经2次变异形成Hap1, Hap1经1~2次变异形成Hap4、Hap6、Hap13~Hap15, Hap14经7次变异形成Hap7, Hap8是由Hap1先经1次突变成1个未知单倍型,再次88次突变成。Hap1、Hap13与1个未知单倍型形成一个回环结构, Hap1、Hap4、Hap12与Hap14形成一个回环结构。

单倍型分布情况见表9,在cpDNA联合片段中,当归野生居群与栽培居群的单倍型分布差异明显, Hap1只分布在栽培当归居群中,且在11个栽培居群中均有分布,其数量占栽培当归样本量的76.8%; Hap2只分布在野生当归居群中,且在7个野生居群中均有分布;特有单倍型共有9个,分布在居群2(2个)、居群3(1个)、居群4(2个)、居群6(1个)、居群7(1个)、居群8(1个)、居群13(1个)。包含单倍型数量最多的是居群7,共检测到5个单倍型,其单倍型多样性为0.62462。

4 讨论

4.1 当归遗传多样性 野生当归的核苷酸多样性和单倍型多样性均低于栽培当归,说明野生当归居群的遗传多样性低于栽培居群,与课题组前期应用ISSR分子标记对野生和栽培当归遗传多样性研究结果一致^[15]。按照GRANT等^[32]提出的标准,单倍型多样性以0.5为临界值,核苷酸多样性以0.005为临界值,二者的值越大,群体的多样性程度越高,当归物种的多样性较低,遗传多样性亟待保护。在野生和栽培当归居群中,均以matK基因片段的多态性最高,野生当归样本中检测到1个变异位点,栽培当归样本中检测到383个变异位点。这是因为matK基因是在叶绿体基因组中进化速率最快的基因^[13],

表7 基于cp DNA联合序列片段的遗传距离与分化系数矩阵

Table 7 Genetic distance and fixation index matrix based on cp DNA combined sequence fragment

居群编号	1	2	3	4	5	6	7	8
1	-	-0.055	0.398	-0.038	-0.001	-0.043	0.001	-0.048
2	0.036	-	0.377	-0.042	0.000	-0.044	0.003	-0.044
3	0.188	0.192	-	0.304	0.498	0.358	0.497	0.427
4	0.045	0.050	0.183	-	0.057	-0.044	0.059	-0.017
5	0.017	0.022	0.193	0.032	-	0.043	0.086	0.000
6	0.037	0.043	0.186	0.051	0.024	-	0.039	-0.025
7	0.017	0.022	0.194	0.033	0.001	0.024	-	0.006
8	0.027	0.032	0.190	0.042	0.012	0.034	0.013	-
9	0.081	0.086	0.169	0.089	0.072	0.085	0.073	0.078
10	0.017	0.022	0.193	0.033	0.001	0.024	0.002	0.012
11	0.017	0.022	0.194	0.033	0.001	0.024	0.001	0.012
13	0.756	0.766	0.449	0.728	0.783	0.746	0.787	0.765
14	0.757	0.767	0.449	0.729	0.784	0.747	0.788	0.766
15	0.757	0.767	0.449	0.729	0.784	0.747	0.788	0.766
16	0.757	0.767	0.449	0.729	0.784	0.747	0.788	0.766
17	0.757	0.767	0.449	0.729	0.784	0.747	0.788	0.766
居群编号	9	10	11	13	14	15	16	17
1	-0.100	0.001	0.001	0.970	0.971	0.971	0.971	0.971
2	-0.105	0.000	0.000	0.959	0.960	0.960	0.960	0.960
3	-0.001	0.499	0.499	0.745	0.746	0.746	0.746	0.746
4	-0.160	0.058	0.058	0.943	0.944	0.944	0.944	0.944
5	-0.001	0.063	0.063	0.999	1.000	1.000	1.000	1.000
6	-0.125	0.043	0.043	0.958	0.959	0.959	0.959	0.959
7	0.001	0.139	0.139	0.999	0.999	0.999	0.999	0.999
8	-0.074	0.000	0.000	0.979	0.979	0.979	0.979	0.979
9	-	0.000	0.000	0.855	0.855	0.855	0.855	0.855
10	0.072	-	0.000	0.999	1.000	1.000	1.000	1.000
11	0.073	0.001	-	0.999	1.000	1.000	1.000	1.000
13	0.659	0.785	0.785	-	0.000	0.000	0.000	0.000
14	0.660	0.786	0.786	0.000	-	0.000	0.000	0.000
15	0.660	0.786	0.786	0.000	0.000	-	0.000	0.000
16	0.660	0.786	0.786	0.000	0.000	0.000	-	0.000
17	0.660	0.786	0.786	0.000	0.000	0.000	0.000	-

注:遗传距离(对角线下),遗传分化系数(对角线上)

在科级、属级水平均有广泛的应用,而且也被用来探讨某些类群种间、种下的系统进化研究^[33-36]。

居群1~居群6是甘肃省已育成的6个当归新品种,其中岷归1号(居群1)是目前种植面积最大的品种^[37],而岷归2号(居群2)、岷归3号(居群3)、岷归4号(居群4)、岷归6号(居群6)当归具有比岷归1号更高的遗传多样性,如能广泛推广种植,有望能够

改善甘肃省栽培当归品质退化的问题^[38];岷归5号(居群5)当归可能是因为样本量的问题,遗传多样性较岷归1号更低。在尚未育成的5个当归新品系(居群7~居群11)中,D42017-01(居群7)和10G辐照(居群9)当归有着较高的遗传多样性,推测其环境适应能力更强。长期生长在自然环境中的野生当归具抗病虫害的优良基因^[3],且与栽培当归的遗传

表8 当归 cp DNA 联合片段的分子方差分析

Table 8 Analysis of molecular variance for cp DNA combined sequence fragment of *Angelica sinensis*

类型	变异来源	自由度 <i>df</i>	平方和	变异组分	变异百分比/%	固定指数	<i>P</i>
野生和栽培	组间	1	7 903.950	352.309 88	94.25	0.942 55	<0.01
	组内居群间	16	583.639	1.732 67	0.46	0.080 69	<0.05
	居群内	158	3 119.155	19.741 49	5.28	0.947 18	<0.01
	总计	175	11 606.744	373.784 04	100		
栽培	居群间	10	583.223	2.627 75	11	0.114 19	<0.05
	居群内	153	3 118.655	20.383 37	89		
	总计	163	3 701.878	23.011 12	100		

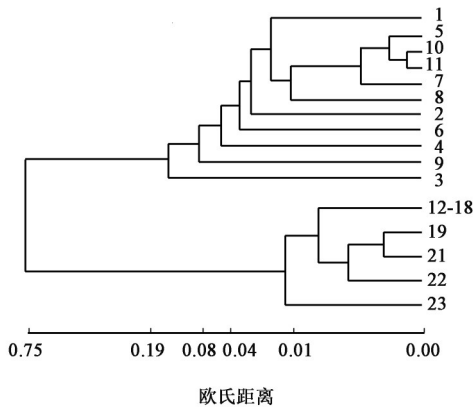


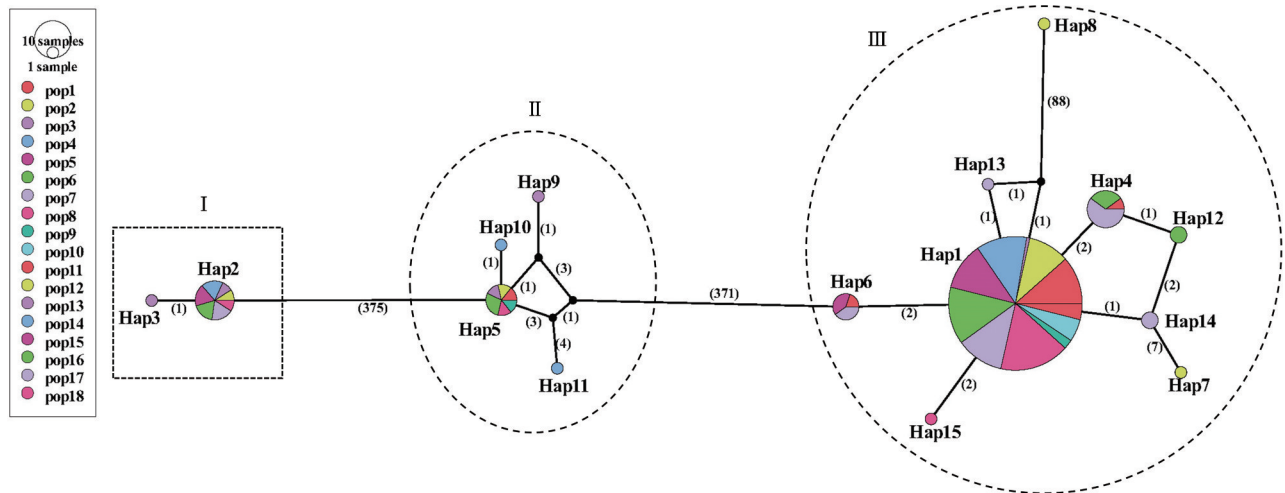
图1 基于 cp DNA 联合序列片段的遗传距离 UPGMA 聚类树
Fig. 1 UPGMA clustering tree of genetic distance based on cp DNA combined sequence fragments

距离较远,可为栽培居群基因库的改良及当归新品种的选育提供优良的种质资源。

4.2 当归遗传变异 当归的分子遗传变异主要存在于野生和栽培两组之间,组间遗传分化系数 F_{ST} 为 0.947 255 ($P < 0.01$), WRIGHT^[39] 认为当 $F_{ST} > 0.25$ 时,种群间极度分化,当归组间 F_{ST} 远 > 0.25 ,说明野

生当归与栽培当归之间产生了显著的遗传分化。采集到的野生当归居群间分化系数均为 0,低于栽培居群的分化系数;在栽培当归居群中变异主要来源于居群内。由此推测,采集的野生当归居群为原始野生居群,不是栽培居群逸生而来,而当归栽培居群也并非来自采集的野生居群,与课题组前期研究结果一致^[12]。基于所选的 3 个 cp DNA 联合序列,野生当归与栽培当归分化程度较高,亲缘关系较远,二者之间的遗传距离远大于野生居群内部或栽培居群内部,可以作为鉴别当归野生与栽培种质资源的候选 cp DNA 条形码。

4.3 当归单倍型分布 176 份当归的 3 个 cp DNA 区域联合片段的单倍型邻接网络图中,当归单倍型呈现一种混杂的分布格局,Hap1 分布在全部栽培当归居群中,Hap2 分布在全部野生当归居群中,而 Hap7~Hap13、Hap15 作为特有单倍型分布于某些居群中;共享单倍型仅分布于各组之内,组间不存在共享单倍型且单倍型之间突变次数较多,说明组间



注:pop.居群

图2 基于 cp DNA 联合片段单倍型的网络

Fig. 2 Network evolution plots of haplotypes based on cp DNA combined sequence fragments

表9 当归单倍型频率及在18居群中的分布

Table 9 Frequency of haplotypes and distribution in 18 populations of *Angelica sinensis*

居群编号	单倍型(个数)	单倍型数目	单倍型多样性	核苷酸多样性
1	1(15)、4(1)、5(1)、6(1)	4	0.314	0.026 86
2	1(13)、5(1)、7*(1)、8*(1)	4	0.350	0.037 57
3	1(1)、5(1)、9*(1)	3	1	0.159 97
4	1(16)、10*(1)、11*(1)	3	0.216	0.050 22
5	1(15)、6(2)	2	0.221	0.000 28
6	1(18)、4(3)、5(2)、12*(2)	4	0.473	0.037 22
7	1(15)、4(6)、6(2)、13*(1)、14(2)	5	0.625	0.000 80
8	1(22)、5(1)、14(1)、15*(1)	4	0.230	0.019 30
9	1(3)、5(1)	2	0.500	0.119 66
10	1(7)	1	0	0
11	1(5)	1	0	0
12	2(1)	1	-	-
13	2(1)、3*(1)	2	1	0.000 64
14	2(2)	1	0	0
15	2(2)	1	0	0
16	2(2)	1	0	0
17	2(2)	1	0	0
18	2(1)	1	-	-

注:*表示特有单倍型

的交流不明显,即野生与栽培当归之间分化显著。跟据 Tajima's D 检验的原理,当 Tajima's D 显著 >0 时,可用于推断瓶颈效应和平衡选择;当 Tajima's D 显著 <0 时,可用于推断群体规模放大和定向选择^[27]。本研究中野生和栽培当归联合序列的 Tajima's D 检验均为不显著负值,推断当归在进化过程中遵循中性进化模型,但栽培当归中 *psbA-trnH* 和 *rbcL1F-rbcL724R* 基因为显著负值,推断栽培当归中该两区域可能经历过定向选择或群体扩张。在栽培当归产生的单倍型中,以 Hap 1 为中心呈辐射状演化关系,与 Tajima's D 检验结果一致。若单倍型在种群或居群中覆盖范围最广、出现频率最高,并处于单倍型网络的中心位置,那么这种单倍型很可能是网络图的所有单倍型中最为古老的类型^[40],由此推断 Hap1 为栽培当归中最古老的单倍型。

[利益冲突] 本文不存在任何利益冲突。

[参考文献]

[1] 国家药典委员会. 中华人民共和国药典:一部[M]. 北京:中国医药科技出版社,2020:139.
[2] 《国植物志》中编辑委员会. 中国植物志:第五十五卷

第三分册[M]. 北京:科学出版社,1992:50-75.

[3] 赵锐明,陈垣,郭凤霞,等. 甘肃岷县野生当归资源分布特点及其与栽培当归生长特性的比较研究[J]. 草业学报,2014,23(2):29-37.
[4] 张尚智,贺莉萍,韩黎明,等. 野生当归及其当归植物资源的研究[J]. 中国现代中药,2012,14(4):33-36.
[5] 张宏意,罗连,余意,等. 当归种质资源调查研究[J]. 中药材,2009,32(3):335-337.
[6] 杜旭. 当归基原植物的研究(2):北海道野生当归的成分分析[J]. 国际中医中药杂志,2006(6):359-360.
[7] 鲍隆友,刘智能. 西藏野生当归资源及人工栽培技术研究[J]. 中国林副特产,2003(4):1-2.
[8] 王鑫,袁庆军,郭增祥,等. 野生当归种子的生物学特性研究[J]. 中国中药杂志,2020,45(10):2368-2373.
[9] 王鑫,袁庆军,孙楷,等. 甘肃地区野生当归的种群特性及其致危关联研究[J]. 中国中药杂志,2019,44(14):2987-2995.
[10] 李涛,何璇. GC-MS测定野生当归挥发油中的化学成分[J]. 华西药理学杂志,2015,30(2):249-250.
[11] 冯伟萌,刘培,严辉,等. 基于高通量测序的野生和栽培当归转录组分析[J]. 中国中药杂志,2020,45(8):1879-1886.
[12] 朱田田,晋玲,黄得栋,等. 野生与栽培当归遗传多样性比较[J]. 中草药,2018,49(1):211-218.

- [13] 毕毓芳,温星,潘雁红,等. 叶绿体DNA条形码在林木中的应用及研究进展[J]. 分子植物育种, 2020, 18(16):5444-5452.
- [14] 付涛,王志龙,钱萍仙,等. 高等植物DNA条形码最新研究进展及其应用[J]. 核农学报, 2016, 30(5): 887-896.
- [15] 朱田田,晋玲,张裴斯,等. 基于ISSR的甘肃不同产区栽培当归遗传多样性研究[J]. 中草药, 2015, 46(23):3549-3557.
- [16] 胡建霖,刘志芳,慈秀芹,等. DNA条形码在热带龙脑香科树种鉴定中的应用[J]. 植物学报, 2019, 54(3): 350-359.
- [17] 董斌,田夏红,谢月亮,等. 剑麻种质资源DNA条形码遗传多样性分析[J]. 分子植物育种, 2021, 19(16):5546-5554.
- [18] 张亚光,陈丹青,朱国琴. ITS2和*psbA-trnH*序列对人参属常见药用植物的DNA分子鉴定研究[J]. 沈阳药科大学学报, 2021, 38(5):523-529.
- [19] 吕瑞华,冯昭,马添翼,等. 陕西关中野生商陆资源的ITS2和*sbA-trnH*条形码序列研究[J]. 药学报, 2020, 55(8):1951-1956.
- [20] 熊淼,袁灿,彭芳,等. 川芎种质鉴定标记开发和系统发育研究[J]. 中草药, 2020, 51(1):169-181.
- [21] Technelysium Pty Ltd. Chromas version 2. 6. 5 [EB/OL]. [2022-03-10]. <https://technelysium.com.au/wp/chromas/>.
- [22] KUMAR S, STECHER G, LI M, et al. MEGA X: Molecular evolutionary genetics analysis across computing platforms[J]. Mol Biol Evol, 2018, 35(6): 1547-1549.
- [23] ZHANG D, GAO F, JAKOVLIĆ I, et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies [J]. Mol Ecol Resour, 2020, 20(1):348-355.
- [24] KIMURA M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences[J]. J Mol Evol, 1980, 16(2):111-120.
- [25] ROHLF F J. NTSYS-pc: Numerical taxonomy and multivariate analysis system, version 2.1 [M]. New York: Applied Biostatistics, 2000.
- [26] JULIO R, ALBERT F M, SÁNCHEZ-DELBARRIO J C, et al. DnaSP v6: DNA sequence polymorphism analysis of large dates [J]. Mol Bio Evol, 2017, 34: 3299-3302.
- [27] TAJIMA F. Evolutionary relationship of DNA sequences in finite populations[J]. Genetics, 1983, 105(2):437-460.
- [28] PONS O, PETIT R J. Measuring and testing genetic differentiation with ordered versus unordered alleles [J]. Genetics, 1996, 144(3):1237-1245.
- [29] EXCOFFIER L, LISCHER H E. Arlequin suite ver 3. 5: A new series of programs to perform population genetics analyses under Linux and Windows [J]. Mol Ecol Resour, 2010, 10(3):564-567.
- [30] KANE J S, PAQUIN M J. POPART: Partial optical implementation of adaptive resonance theory 2 [J]. IEEE Trans Neural Netw, 1993, 4(4):695-702.
- [31] CLEMENT M J, SNELL Q, WALKER P, et al. TCS: Estimating gene genealogies [C]//16th International Parallel and Distributed Processing Symposium; 2002. LauderdaleFort, FL, USA: CD-ROM/Abstracts Proceedings. IEEE Computer Society, 2002, 184-192.
- [32] GRANT W A S, BOWEN B W. Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation [J]. J Hered, 1998, 89(5): 415-426.
- [33] 许丹芸,张辉菊,李可心,等. 22种樟科植物DNA条形码分子鉴定[J]. 中国实验方剂学杂志, 2021, 27(16):159-166.
- [34] 林晓霞,鹿炎,李会丽,等. 基于*matK*和*rbcL*序列的石斛属植物亲缘关系研究[J]. 生物学杂志, 2021, 38(1):46-50.
- [35] 何爱芳,傅鹏,陈恒宇,等. 基于DNA条形码的金槐系统发育和变异位点分析[J]. 中国实验方剂学杂志, 2022, 28(12):183-191.
- [36] 满金辉,石玥,张靖晗,等. 基于叶绿体基因*matK*及UPLC对市售大黄的种质资源鉴定和质量分析[J]. 药学报, 2022, 57(2):514-524.
- [37] 黄璐琦,晋玲. 当归生产加工适宜技术(中国中药资源大典-中药材系列中药材生产加工适宜技术丛书) [M]. 北京:中国医药科技出版社, 2018:6-10.
- [38] 李应东,晋玲,藺海明,等. 当归研究[M]. 北京:科学出版社, 2021:36-47.
- [39] WRIGHT S W. The interpretation of population structure by F-statistics with special regard to systems of mating [J]. Evolution, 1965, 19(3):395-420.
- [40] POSADA D, CRANDALL K A. Intraspecific gene genealogies: Trees grafting into networks [J]. Trends Ecol Evol, 2001, 16(1):37-45.

[责任编辑 顾雪竹]