

梁燕妮,黄银霞,魏诗琴,等. 紫芽六堡茶转录组 SSR 位点序列分析[J]. 江苏农业科学,2022,50(3):42-48.
doi:10.15889/j.issn.1002-1302.2022.03.007

紫芽六堡茶转录组 SSR 位点序列分析

梁燕妮,黄银霞,魏诗琴,陈秋雨

(梧州学院,广西梧州 543002)

摘要:以广西梧州紫芽六堡茶为研究对象,运用 MISA 软件对转录组测序获得的 165 570 条 unigene 序列(总长度为 240 177 233 bp),进行 SSR 位点搜索。使用 Excel 软件对 SSR 位点的出现频率、基元类型、基元组成、平均距离等进行处理分析。结果表明,共在 10 058 条 unigene 中找到符合条件的 SSR 位点 11 482 个,出现频率为 6.93%,平均距离为 20.92 kb,平均长度为 20.80 bp。紫芽六堡茶转录组 SSR 中,双碱基重复是主要的重复类型,占总 SSR 的 19.03%,GA/TC、GAA/TTC、AAAT/ATTT、AAAAG/CTTT 分别是双碱基、三碱基、四碱基、五碱基的优势重复单元。序列长度大于 20 bp 的 SSR 位点数占总 SSR 的 44.05%,其中长度为 24 bp 的 SSR 位点数最多,占 SSR 总数的 16.01%。说明紫芽六堡茶转录组 SSR 种类丰富、可用性高、多态性高。通过对 SSR 位点的分布特征和序列特征进行分析,可为后续六堡茶的进化、品种鉴定、分子育种等研究提供依据。

关键词:茶树;六堡茶;转录组;SSR 位点;序列分析

中图分类号:S571.101 **文献标志码:**A **文章编号:**1002-1302(2022)03-0042-07

六堡茶因其原产于广西壮族自治区梧州苍梧县六堡镇而得名,为广西特色优质的历史名茶,并以其“红、浓、陈、醇”的特性闻名于世。六堡茶的制作以六堡镇的群体种茶树为原料,六堡群体种茶树为灌木型中小叶种,树势开展,分枝密。从芽色分,以青苗茶为主,占 60%,紫芽茶占 20%,茶树红紫色芽叶是一种特异茶叶资源,紫芽六堡茶树由于富含花青素成为当前茶树新品种开发利用的热点及研

究的重点。

简单重复序列(simple sequence repeat,简称 SSR)别称微卫星 DNA,以 1~6 个碱基为核心序列,具有多态性高、重复性好、稳定性高等特点^[1],随着现代分子技术的发展,SSR 技术已广泛应用于植物分子辅助育种、种质资源遗传多样性、绘制遗传图谱等方面的研究中。毛娟等利用 30 对茶树核心 SSR 引物,对 3 个代表性的野生和栽培大理茶居群进行遗传分析^[2]。陈春林等从紫娟茶树转录组 46 041 条单基因簇(unigene)中筛选出 57 976 个 SSR 位点,并对 SSR 序列的分布特征进行了分析,筛选出 44 对高质量引物^[3]。班秋艳等对陕西及临近 8 个省份 118 份茶树资源利用表达序列标签(EST)-SSR

收稿日期:2021-10-17

基金项目:广西自然科学基金青年科学基金(编号:2020GXNSFBA159007);梧州高新区梧州学院产学研项目(编号:2020G012);广西大学生创新创业项目(编号:202111354113)。

作者简介:梁燕妮(1980—),女,广西梧州人,硕士,高级实验师,主要从事分子生物学研究。E-mail:371200021@qq.com。

207-214.

- [19] Hirano T, Kishi M, Sugimoto H, et al. Thioesterase activity and subcellular localization of acylprotein thioesterase 1/lysophospholipase 1[J]. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 2009, 1791(8):797-805.
- [20] Meldrum E, Parker P J, Carozzi A. The PtdIns - PLC superfamily and signal transduction [J]. *Biochimica et Biophysica Acta* (- Molecular Cell Research), 1991, 1092(1):49-71.
- [21] Kocourková D, Krčková Z, Pejchar P, et al. The phosphatidylcholine - hydrolysing phospholipase C NPC₄ plays a role in response of *Arabidopsis* roots to salt stress[J]. *Journal of Experimental Botany*, 2011, 62(11):3753-3763.

- [22] Pokotylo I, Kolesnikov Y, Kravets V, et al. Plant phosphoinositide - dependent phospholipases C: Variations around a canonical theme [J]. *Biochimie*, 2014, 96:144-157.
- [23] Salinas - Mondragon R E, Kajla J D, Perera I Y, et al. Role of inositol 1,4,5 - triphosphate signalling in gravitropic and phototropic gene expression [J]. *Plant, Cell & Environment*, 2010, 33(12):2041-2055.
- [24] Wang X M, Devaiah S P, Zhang W H, et al. Signaling functions of phosphatidic acid[J]. *Progress in Lipid Research*, 2006, 45(3):250-278.
- [25] 杨学利,马沁沁,袁向华,等. 磷脂酶 D 及其催化机制的研究进展[J]. *中国油脂*, 2016, 41(3):80-84.

分子标记进行了遗传多样性分析^[4]。但目前尚未见针对紫芽六堡茶 SSR 位点的分析报告。本研究基于转录组测序,首次对紫芽六堡茶转录组 SSR 位点的分布及其序列特征进行分析,以期为今后紫芽六堡茶品种选育、功能基因标记等提供基础依据。

1 材料与方法

1.1 试验材料

所用鲜茶叶于 2021 年 6 月采自广西壮族自治区梧州六堡镇四柳村群体种六堡茶。分别采摘 1 芽 1 叶的紫芽、绿芽六堡茶叶片,用离心管装好,迅速放入液氮中保存。每个样品 3 次重复。将采集的样品委托深圳华大基因股份有限公司进行转录组测序。

1.2 SSR 位点搜索

针对 Raw 数据,本试验采用 SOAPnuke 软件^[5]进行数据的过滤与质控,将包含接头污染的读长(reads)、未知碱基 N 含量大于 5% 的 reads 以及质量值低于 10 的碱基含量大于 20% 的 reads 去除,以获得高质量序列。使用 Trinity 对干净读序(clean reads)进行组装,然后使用 Tgicl 对转录本进行聚类去冗余得到 unigene。使用 MISA(<http://pgsc. ipk - gatersleben. de/misa>)对获得的 unigene 进行 SSR 位点搜索,搜索标准为单碱基重复次数 12 次以上,双碱基重复次数 6 次以上,三碱基和四碱基重复次数 5 次以上,五碱基和六碱基重复次数 4 次以上;2 个 SSR 之间的间隔碱基数 > 100 bp,SSR 位点距离两端侧翼序列 > 150 bp。使用 Primer 3 对检测到的 SSR 进行引物设计。

1.3 数据分析

所得数据使用 Excel 软件进行处理分析,包括 SSR 位点的出现频率、基元类型、基元组成、平均距离等。其中,SSR 位点出现频率是搜索到的 SSR 总数与 unigene 序列总数的比值;SSR 位点的平均距离是搜索到的 SSR 总数与总 unigene 长度之比,根据这些数据所反映出的特征可以对 SSR 位点的分布特征和序列特征进行分析。

2 结果与分析

2.1 紫芽六堡茶转录组中 SSR 位点的分布特征

通过对紫芽六堡茶叶片进行转录组测序,共获得去冗余的 unigene 序列 165 570 条,总长度为 240 177 233 bp,其中 G + C 占 41.49%。利用软件

SSRFinder 对这些序列数据进行搜索,共在 10 058 条 unigene 中找到符合条件的 SSR,总长为 238 833 bp,SSR 位点发生的频率为 6.04%,其中,8 867 条 unigene 含单个 SSR 位点,1 191 条 unigene 序列含有 1 个以上的 SSR 位点。共检出 11 482 个 SSR 位点,出现的频率为 6.93%。SSR 分布的平均距离为 20.92 kb,即平均每 20.92 kb 就出现 1 个 SSR(表 1)。

表 1 转录组各 SSR 位点的分布特征

位点特征指标	搜索结果
uinigene 序列总数(条)	165 570
序列总长度(bp)	240 177 233
SSR 位点总数(个)	11 482
包含 SSR 位点的序列数(条)	8 867
含 1 个以上 SSR 位点的序列数(条)	1 191
SSR 发生频率(%)	6.04
SSR 出现频率(%)	6.93
G + C 的含量(%)	41.49

2.2 紫芽六堡茶转录组 SSR 的基元类型

紫芽六堡茶转录组中具有较为丰富的 SSR 碱基类型(表 2)。其中重复率最高的为双碱基,占 SSR 位点总数的 55.01%,其次为三碱基、六碱基,分别占 28.64%、8.22%,其他的重复碱基类型所占比例较低(图 1)。各重复碱基类型的出现频率存在一定差异。主要集中在双碱基和三碱基重复上,占总 SSR 的 83.65%。其中,双碱基重复类型出现频率最高,为 3.81%,出现频率最低的是单碱基重复,仅为 0.11%。从分布情况看,转录组中平均每 20.92 kb 就出现 1 个 SSR,不同重复类型间平均分布距离存在差别。其中,单碱基重复的平均分布距离最长,为 1 264.09 kb;其次是四碱基重复,平均分布距离为 839.78 kb;双碱基重复的平均距离最短,为 38.03 kb。紫芽六堡茶转录组 SSR 位点的序列总长度为 238 833 bp,平均长度为 20.80 bp。

2.3 紫芽六堡茶转录组 SSR 重复基元分析

紫芽六堡茶转录组 SSR 重复基元的序列特征详见表 3。由单碱基到六碱基重复基元种类数分别为 2、6、30、48、121、460 种,共 667 种。各碱基重复基元的数量分布不均衡,其中,单碱基重复基元数量最多的是 A/T,共 186 个,占 SSR 总位点数的 1.62%,占单碱基重复基元总数的 97.89%。双碱基重复基元中,数量最多的类型是 GA/TC,共 2 185 个,占 SSR 总位点数的 19.03%,占双碱基重复基元

表 2 紫芽六堡茶转录组 SSR 位点不同基元类型的分布状况

重复碱基类型	SSR 位点数量 (个)	总长度 (bp)	平均长度 (bp)	各类型比例 (%)	频率 (%)	平均距离 (kb)
单碱基重复	190	2 934	15.44	1.65	0.11	1 264.09
双碱基重复	6 316	12 9318	20.47	55.01	3.81	38.03
三碱基重复	3 289	64 644	19.65	28.64	1.99	73.02
四碱基重复	286	6 260	21.89	2.49	0.17	839.78
五碱基重复	457	10 075	22.05	3.98	0.28	525.55
六碱基重复	944	25 602	27.12	8.22	0.57	254.43
总计	11 482	238 833	20.80	100.00	6.93	20.92

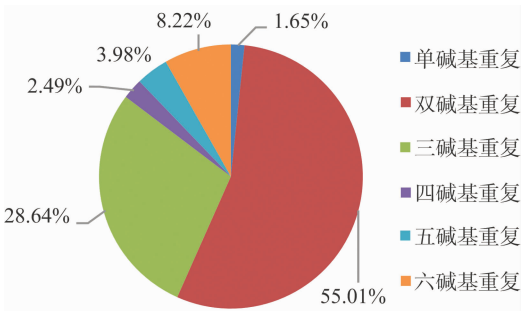


图1 不同 SSR 重复基元类型占比情况

总数的 35.45% ;其次是 AG/CT,共 2 165 个,占 SSR 总位点数的 18.86% ,占双碱基重复基元总数的 34.28% 。三碱基重复基元中,数量最多的类型是 GAA/TTC,共 345 个,占 SSR 总位点数的 3.00% ,占三碱基重复基元总数的 10.49% ;其次是 AGA/TCT,共 288 个,占 SSR 总位点数的 2.51% ,占三碱基重复基元总数的 8.76% 。四碱基重复基元中,数量最多的类型是 AAAT/ATTT,共 55 个,占 SSR 总位点数的 0.48% ,占四碱基重复基元总数的 19.23% 。

表 3 SSR 位点重复的基元序列特征

重复类型	SSR 位点数量(个)	基元种类(种)	重复基元数量最多的碱基组成	重复类型最多的碱基数量(个)	重复类型最多的占总 SSR 比例(%)
单碱基重复	190	2	A/T	186	1.62
双碱基重复	6 316	6	GA/TC、AG/CT	2 185	19.03
三碱基重复	3 289	30	GAA/TTC、AGA/TCT	345	3.00
四碱基重复	286	48	AAAT/ATTT、TAAA/TTTA	55	0.48
五碱基重复	457	121	AAAAG/CTTTT、AAAAT/ATTTT	48	0.42
六碱基重复	944	460	AAAAAT/ATTTTT、CAAAAA/TTTTTG、GATGAA/TTCATC	22	0.19
总计	11 482	667		2 841	24.74

五碱基重复基元中,数量最多的类型是 AAAAG/CTTTT,共 48 个,占 SSR 总位点数的 0.42% ,占五碱基重复基元总数的 10.50% 。六碱基重复基元中,数量最多的类型分别是 AAAAAT/ATTTTT、CAAAAA/TTTTTG、GATGAA/TTCATC,各有 22 个,各占 SSR 总位点数的 0.19% ,各占六碱基重复基元总数的 2.33% 。转录组中不同 SSR 重复基元类型的比例详见图 2。

2.4 紫芽六堡茶转录组中各基元的重复次数

SSR 的多态性受重复碱基序列长度的影响,而重复碱基序列长度由 SSR 重复次数决定^[5-6]。通过对紫芽六堡茶转录组中不同基元类型 SSR 位点重复出现次数的统计(表 4)可知,6 种基元重复次数主要集中在 4~12 次,重复出现 4~12 次的 SSR

位点数共计 9 941 个,占总 SSR 位点数的 86.58% 。其中,重复出现 6 次和 5 次的 SSR 位点数最多,分别占总 SSR 位点数的 16.63% (1 910 个)、15.89% (1 825 个);其次为重复 7 次和 8 次的,分别占总 SSR 位点数的 10.53% (1 209 个)、9.80% (1 125 个)。统计发现,在重复出现 6 次的 SSR 位点中,双碱基最多(1 062 个),其次是三碱基(673 个),五碱基最少(仅 20 个)。重复出现 5 次的 SSR 位点中,三碱基最多(1 375 个),四碱基、五碱基、六碱基的数量差别不大。SSR 位点重复次数在 12 次以上的基元占总 SSR 位点数的 13.42% 。不同基元类型不同重复次数的 SSR 位点数分布情况见图 3。从图 3 可以看出,随着重复次数的增多,SSR 位点的数量逐渐下降。

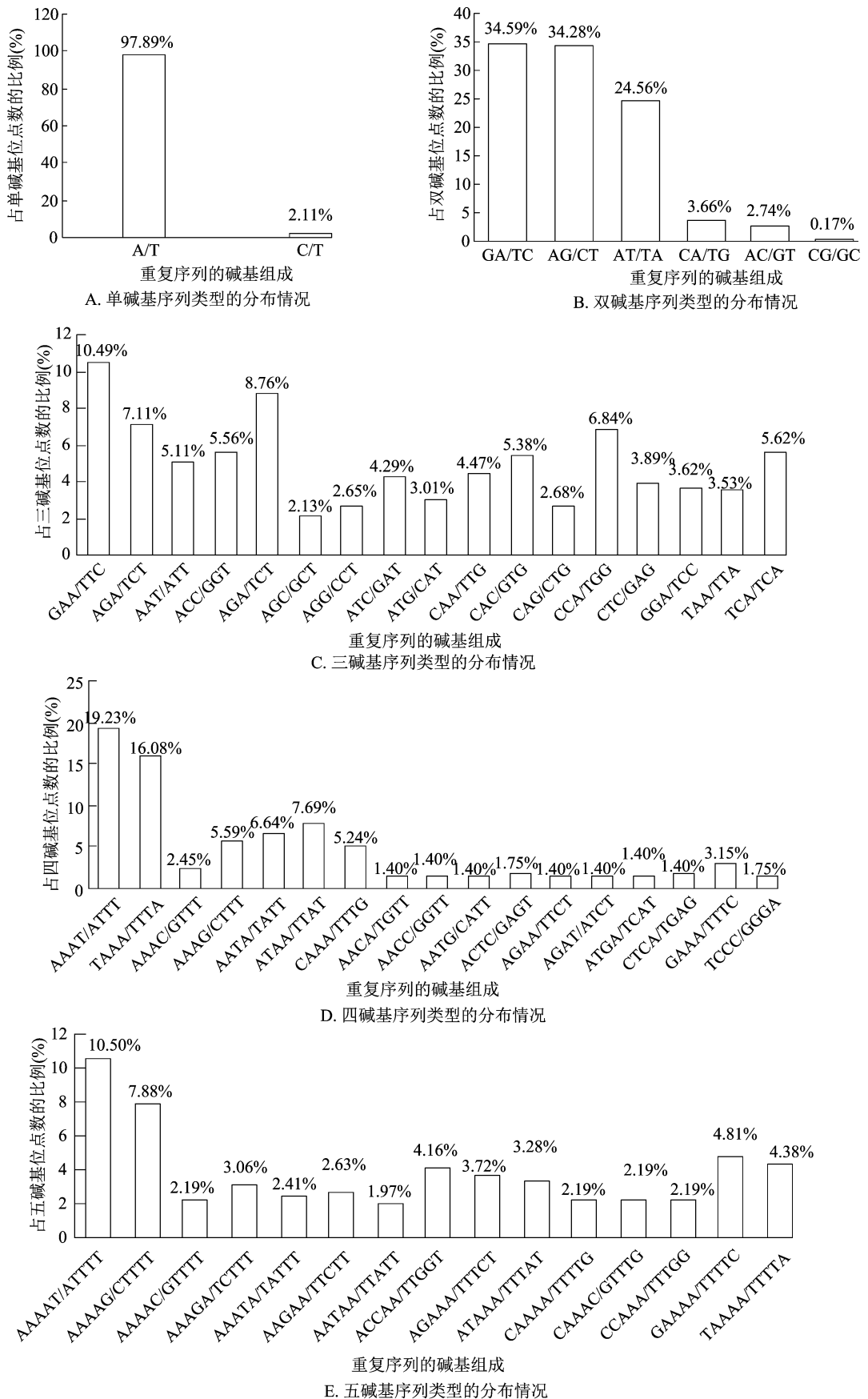


图2 紫芽六堡茶转录组中不同 SSR 序列类型的比例

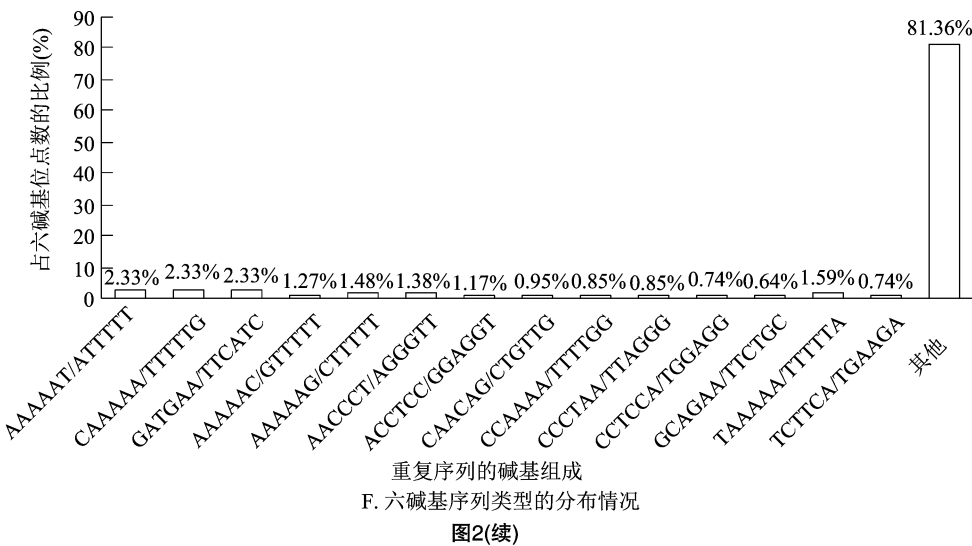
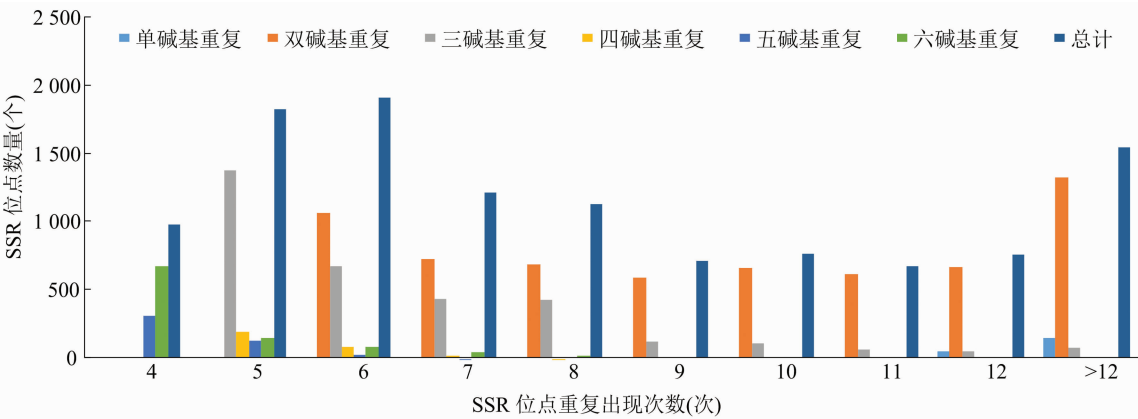


表 4 不同基元类型不同重复次数的 SSR 位点数量

重复碱基类型	不同重复出现次数的 SSR 位点的数量(个)									
	4 次	5 次	6 次	7 次	8 次	9 次	10 次	11 次	12 次	12 次以上
单碱基重复									45	29
双碱基重复			1 062	721	685	588	659	613	664	284
三碱基重复		1 375	673	428	421	115	101	59	45	28
四碱基重复		186	78	12	7	3	0	0	0	0
五碱基重复	307	122	20	7	1	0	0	0	0	0
六碱基重复	668	142	77	41	11	3	1	1	0	0
总计	975	1 825	1 910	1 209	1 125	709	761	673	754	1 541
占总 SSR 位点数的比例(%)	8.49	15.89	16.63	10.53	9.80	6.17	6.63	5.86	6.57	13.42



2.5 紫芽六堡茶转录组 SSR 重复片段的长度分布

紫芽六堡茶转录组 SSR 重复片段的长度分布情况见图 4。SSR 序列长度变化在 12 ~ 70 bp 之间, 平均长度为 20.80 bp。其中长度为 24 bp 的 SSR 位点数最多, 共有 1 838 个, 占 SSR 总数的 16.01%。长度在 12 ~ 20 bp 的 SSR 有 6 424 个, 占 SSR 总数的 55.95%; > 20 bp 的 SSR 有 5 058 个, 占 SSR 总数的 44.05%。双碱基中最常见的长度是 12 bp

(1 021 个, 9.24%), 长度变化幅度最大, 最小为 12 bp 最长为 70 bp。三碱基中最多的长度为 15 bp (1 375 个, 11.98%)。四碱基、五碱基均以 20 bp 为主, 分别有 186 (1.62%)、307 (2.67%) 个。六碱基中长度以 24 bp (668 个, 5.82%) 数量最多。

3 讨论与结论

通过对紫芽六堡茶进行转录组测序, 共获得

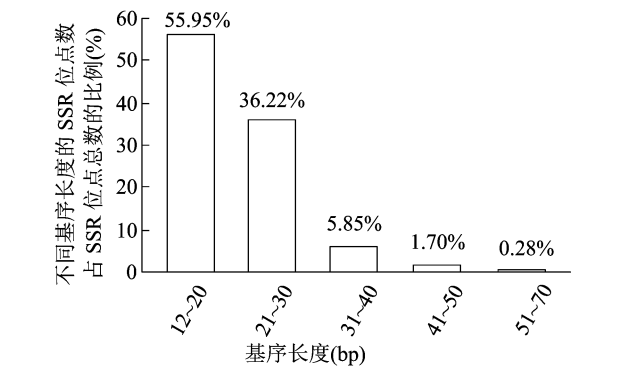


图4 转录组 SSR 重复片段的长度分布情况

165 570 条 unigene 序列,总长度为 240 177 233 bp,共搜索到 11 482 个 SSR 位点,出现频率为 6.93%,略低于杨华等报道的茶树转录组中 SSR 位点出现频率(9.64%)^[7],但高于南方红豆杉^[8](2.07%)、杜仲^[9](2.90%)、短丝木樨^[10](4.64%)、银杏^[11](5.95%)、马尾松^[12](3.45%)、云南松^[13](3.07%)、甘薯^[14](4.4%)等植物。紫芽六堡茶转录组 SSR 分布的平均距离为 20.92 kb,分布水平高于云南松^[13](29.00 kb)、杜仲^[9](26.31 kb)、藤茶^[15](31.69 kb)等植物。与其他已报道的茶树比较(表 5)可知,紫芽六堡茶的 SSR 位点出现频率和平均距离与其他茶树存在较大差异,这可能与本研

表 5 不同茶树品种 SSR 位点信息比较

品种	出现频率 (%)	平均距离 (kb)	平均长度 (bp)	主要重复类型
龙井 43	9.64	3.68	16.00	双碱基重复
舒茶早	37.19	2.85	16.00	双碱基重复
四球茶	23.79	2.07	9.53	双碱基重复
紫芽六堡茶	6.93	20.92	20.80	双碱基重复

表 6 不同茶树品种 SSR 位点重复基元序列最多的碱基组成

品种	SSR 位点重复基元序列最多的碱基组成				
	双碱基重复	三碱基重复	四碱基重复	五碱基重复	六碱基重复
龙井 43	AG/CT	GAA/TTC	AAAT/ATTT	AAAAG/CTTTT	AAAAAG
舒茶早	AG/CT	AAG/CTT	AAAT/ATTT	AAAAT/ATTTT	AACCCT/AGGGTT
云抗 10 号	TC/CT	AAT/TTA	AAAT/ATTT	AAAAT/ATTTT	AAAAAT/ATTTTT
四球茶	AG/CT	ACC/GGT	AAAC/GTTT	AAAAC/GTTTT	ACCGCC
紫鹃	AG/CT	AAG/CTT	AAAT/ATTT	AGATG/ATCTC	CGCCAT/GCGGTA
紫芽六堡茶	GA/TC	GAA/TTC	AAAT/ATTT	AAAAT/ATTTT	AAAAAT/ATTTTT

有研究认为,SSR 的多态性与 SSR 基元重复次数及序列长度密切相关。一般认为重复次数越多,长度越长,多态性越高,当重复次数 > 12 次时,多态位点含量高,紫芽六堡茶基元重复次数在 12 次以上

究中搜索 SSR 的算法不一致、数据处理方法不同、以及进化程度不同等因素有关。总体来说,紫芽六堡茶转录组中 SSR 位点的种类和数量还是较为丰富的,可为今后分子育种提供依据。

大多数研究表明,植物 SSR 重复基元以双碱基与三碱基重复为主。本研究发现紫芽六堡茶转录组 SSR 重复基元以双碱基重复最多,占 SSR 位点总数的 55.01%,其次为三碱基重复,占 SSR 位点总数的 28.64%,这与陈春林等报道的茶树 SSR 优势重复类型^[3,7,16-19]一致。银杏^[11]、猕猴桃^[20]、杏树和桃树^[21]等植物也是以双碱基重复为主。而玉米、水稻、小麦等禾谷类作物则以三碱基重复为主^[22]。经统计,紫芽六堡茶单碱基的优势重复基元为 A/T,与已报道的大多茶树单碱基优势重复基元一致;双碱基的优势重复基元为 GA/TC,而已报道的茶树中则多以 AG/CT 为主;三碱基的优势重复基元为 GAA/TTC,与龙井 43 一致;四碱基的优势重复基元为 AAAT/ATTT,与龙井 43、舒茶早、云抗 10 号、紫鹃等茶树一致;五碱基的优势重复基元为 AAAAT/ATTTT,与舒茶早、云抗 10 号一致;六碱基的优势重复基元为 AAAAAT/ATTTTT,与云抗 10 号一致(表 6)。表明基于转录组测序得到的 SSR 位点中可能普遍大量存在这类重复基元,推测可能是某些特定的 DNA 家族或者辅助合成某些优势蛋白^[23]。此外,在搜索到的紫芽六堡茶 SSR 位点中还发现了在高等植物转录组中较为少见的 CG/GC 重复基元 11 个,以及在双子叶植物中不常见的 CCG/CGG 重复基元 45 个,这些位点可能与一些特定的生理功能,如抗逆性、信号传导、代谢调控等有关,表明紫芽六堡茶的 SSR 位点具有一定的特异性。

的占 13.42%。当 SSR 长度在 12 ~ 20 bp 之间时,多态性表现一般;当 SSR 长度 > 20 bp 时,则具有较高的多态性。紫芽六堡茶转录组的 SSR 平均长度为 20.80 bp,其中序列长度 > 20 bp 的占 44.05%,高于

四球茶(33.22%)、龙井 43(19.77%)以及舒茶早(8.65%)。表明紫芽六堡茶转录组 SSR 具有很高多态性和利用价值。

综上所述,本研究结果表明,紫芽六堡茶的 SSR 种类丰富、可用性高、多态性高。可为进一步开发紫芽六堡茶功能基因 SSR 分子标记奠定基础,对六堡茶种质资源的保护及茶树特定性状的辅助选育、基因组学的研究等方面都具有重要的意义。

参考文献:

- [1] Mrázek J, Guo X X, Shah A. Simple sequence repeats in prokaryotic genomes[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(20): 8472–8477.
- [2] 毛娟, 江鸿键, 杨如兵, 等. 野生和栽培大理茶居群的遗传多样性与群体结构[J]. 茶叶科学, 2021, 41(4): 454–462.
- [3] 陈春林, 田易萍, 陈林波, 等. 基于荧光标记的紫娟茶树转录组 EST-SSR 标记开发[J]. 江苏农业学报, 2018, 34(4): 747–753.
- [4] 班秋艳, 潘宇婷, 潘铖, 等. 基于 EST-SSR 标记的陕西茶树种质资源遗传多样性研究[J]. 西北农林科技大学学报(自然科学版), 2021, 49(11): 69–78.
- [5] Chen Y X, Chen Y S, Shi C M, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data[J]. Gigascience, 2018, 7(1): 1–6.
- [6] 郭聪, 陈燕, 王莹, 等. 美国红枫转录组 SSR 序列分析[J]. 中南林业科技大学学报, 2021, 41(7): 132–141.
- [7] 杨华, 陈琪, 韦朝领, 等. 茶树转录组中 SSR 位点的信息分析[J]. 安徽农业大学学报, 2011, 38(6): 882–886.
- [8] 李炎林, 杨星星, 张家银, 等. 南方红豆杉转录组 SSR 挖掘及分子标记的研究[J]. 园艺学报, 2014, 41(4): 735–745.
- [9] 黄海燕, 杜红岩, 乌云塔娜, 等. 基于杜仲转录组序列的 SSR 分子标记的开发[J]. 林业科学, 2013, 49(5): 176–181.
- [10] 戴亚平. 短丝木犀转录组测序及 SSR 分子标记开发研究[D].

南京:南京林业大学, 2015.

- [11] 范雪娇, 汪贵斌, 杨晓明, 等. 银杏转录组 SSR 位点分布及其序列特征分析[J]. 经济林研究, 2021, 39(2): 63–73.
- [12] 杜明凤, 丁贵杰. 基于马尾松干旱转录组的抗旱功能 SSR 位点分析[J]. 林业科学研究, 2018, 31(5): 9–19.
- [13] 蔡年辉, 许玉兰, 徐杨, 等. 云南松转录组 SSR 的分布及其序列特征[J]. 云南大学学报(自然科学版), 2015, 37(5): 770–778.
- [14] 黄立飞, 房伯平, 陈景益, 等. 甘薯 EST 资源的 SSR 信息分析[J]. 植物生理学通讯, 2009, 45(1): 23–27.
- [15] 许明, 曾莉, 黄昕颖, 等. 藤茶转录组 SSR 分子标记开发与初步验证[J]. 分子植物育种, 2020, 18(19): 6441–6448.
- [16] 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.
- [17] 彭靖茹, 檀业维, 黄寿辉, 等. 茶树‘云抗 10 号’全基因组 SSR 信息分析及 4CL 基因引物开发[J]. 分子植物育种, 2020, 18(16): 5411–5418.
- [18] 黎瑞源, 邢辉, 申铁. 四球茶转录组 SSR 位点信息分析[J]. 安徽农业大学学报, 2017, 44(4): 558–562.
- [19] 陈琪, 杨华, 韦朝领, 等. 基于茶树芽叶转录组序列的 EST-SSR 分布特征研究[J]. 安徽农业大学学报, 2016, 43(2): 170–175.
- [20] Fraser L G, Harvey C F, Crowhurst R N, et al. EST-derived microsatellites from *Actinidia* species and their potential for mapping[J]. Theoretical and Applied Genetics, 2004, 108(6): 1010–1016.
- [21] Jung S, Abbott A, Jesudurai C, et al. Frequency, type, distribution and annotation of simple sequence repeats in Rosaceae ESTs[J]. Functional & Integrative Genomics, 2005, 5(3): 136–143.
- [22] Varshney R K, Graner A, Sorrells M E. Genic microsatellite markers in plants: features and applications[J]. Trends in Biotechnology, 2005, 23(1): 48–55.
- [23] 郝广婧, 祁银燕, 张得芳, 等. 基于转录组的黑果枸杞 SSR 分布特征分析及引物设计[J]. 分子植物育种, 2019, 17(13): 4342–4350.

(上接第 35 页)

- [21] Young M D, Wakefield M J, Smyth G K, et al. Gene ontology analysis for RNA-seq: accounting for selection bias[J]. Genome Biology, 2010, 11(2): R14.
- [22] Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment[J]. Nucleic Acids Research, 2007, 36(S1): D480–D484.
- [23] Mao X Z, Cai T, Olyarchuk J G, et al. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary[J]. Bioinformatics, 2005, 21(19): 3787–3793.
- [24] Trapnell C, Williams B A, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation[J]. Nature

Biotechnology, 2010, 28(5): 511–515.

- [25] 车荣波. 基于转录组数据的鲑鱼分子标记筛选及基因差异表达分析[D]. 舟山:浙江海洋学院, 2015: 45–60.
- [26] 全俊龙. 日本七鳃鳗(*Lampetra japonica*)血液、肝脏组织的比较转录组学研究[D]. 大连:辽宁师范大学, 2013: 33–60.
- [27] 崔军. 基于 RNA-Seq 技术的红鳍东方鲀鳃和鳔转录组的初步研究[D]. 大连:大连海洋大学, 2013: 46–67.
- [28] Hinamoto N, Maeda T. Isolation of microsatellite markers in *Neoseiulus womersleyi* Schicha (Acari: Phytoseiidae)[J]. Journal of the Acarological Society of Japan, 2005, 14(1): 25–30.
- [29] 全迎春, 马冬梅, 白俊杰, 等. 大口黑鲈转录组 SNPs 筛选及其与生长的关联分析[J]. 水生生物学报, 2016, 40(6): 1128–1134.
- [30] 龚诗琦, 王志勇, 肖世俊, 等. 黄姑鱼转录组 SSR 的开发与验证[J]. 集美大学学报(自然科学版), 2016, 21(4): 241–246.