

# 绣球花型全基因组关联分析

唐琪琦<sup>1</sup>, 李金帅<sup>2</sup>, 邱 帅<sup>3</sup>, 张宪权<sup>4</sup>, 高 凯<sup>3</sup>, 秦 俊<sup>4</sup>, 刘群录<sup>1</sup>

(<sup>1</sup>上海交通大学设计学院, 上海 200240; <sup>2</sup>上海勘测设计研究院有限公司, 上海 200335; <sup>3</sup>杭州市园林绿化股份有限公司, 杭州 310020;

<sup>4</sup>上海辰山植物园, 上海 201602)

**摘要:** 花序类型是绣球(*Hydrangea macrophylla*)的重要观赏性状之一。为了定位控制花序类型的关键基因, 对 125 株绣球单株进行了基因组重测序, 获得了大量的单核苷酸多态性(SNP)。基于 SNP 分子标记和花序类型, 分别采用 Fastlmm 和 GEMMA 软件的一般线性模型(GLM)和混合线性模型(MLM)进行全基因组关联分析(GWAS), 以筛选出与花序类型密切相关的 SNP。GWAS 分析共获得了 285 个与花序表型显著相关的位点( $-\log_{10}(P) > 9$ ), 表型变异解释率为 11.80% ~ 60.54%。从 285 个位点中筛选出 12 个 SNP 位点, 采用测序法和竞争性等位基因特异性 PCR 法(KASP)对 52 个绣球品种及 9 个杂交后代群体进行了基因分型验证, 其中 SNP 位点 Hma1.2p1\_0060F.1\_796640、Hma1.2p1\_0060F.1\_1540773、Hma1.2p1\_0653F.1\_868484、Hma1.2p1\_0669F.1\_949341 经检验与花序类型紧密关联, 表明控制绣球花序的基因可能位于 Hma1.2p1\_0060F.1、Hma1.2p1\_0653F.1 和 Hma1.2p1\_0669F.1 Scaffold 中。以 285 个 SNP 位点为中心寻找其上下游的基因, 共标注到 1287 个基因。根据基因功能注释等信息预测了 4 个与绣球花序类型相关的编码转录因子的候选基因, 分别为 2 个 MYB, 1 个 MADS 和 1 个 bHLH。该结果为后续的基因克隆和绣球分子标记辅助育种提供了依据。

**关键词:** 绣球; 花序类型; 全基因组关联分析; 基因分型

## Genome-wide Association Study of *Hydrangea macrophylla* Inflorescence Type

TANG Qi-qi<sup>1</sup>, LI Jin-shuai<sup>2</sup>, QIU Shuai<sup>3</sup>, ZHANG Xian-quan<sup>4</sup>, GAO Kai<sup>3</sup>, QIN Jun<sup>4</sup>, LIU Qun-lu<sup>1</sup>

(<sup>1</sup>School of Design, Shanghai Jiao Tong University, Shanghai 200240; <sup>2</sup>Shanghai Investigation Design & Research Institute Co., Ltd., Shanghai 200335; <sup>3</sup>Hangzhou Landscaping Incorporated Company, Hangzhou 310020; <sup>4</sup>Shanghai Chen Shan Botanical Garden, Shanghai 201602)

**Abstract:** Inflorescence type is one of the important ornamental traits of *Hydrangea macrophylla*. To localize the key genes controlling inflorescence types, 125 *H. macrophylla* individual plants were re-sequenced to obtain the single nucleotide polymorphisms (SNPs). Genome-wide association study (GWAS) was carried out using general linear model (GLM) and mixed linear model (MLM) of Fastlmm and GEMMA software, to identify the SNPs that strongly associated with inflorescence types. respectively. 285 SNPs significantly associating with floral phenotypes were found (cutoff:  $-\log_{10}(P) > 9$ ), which contributed to the phenotypic variation from 11.80% to 60.54%. Twelve SNPs from 285 loci were further genotyped by Sanger sequencing and kompetitive allele-specific PCR (KASP) in 52 hydrangea cultivars, and 9 hybrid populations. Four SNPs including Hma1.2p1\_0060F.1\_796640, Hma1.2p1\_0060F.1\_1540773, Hma1.2p1\_0653F.1\_868484 and Hma1.2p1\_0669F.1\_949341 were validated closely associating with inflorescence type, which indicated that the genes controlling hydrangea inflorescence type were probably allocated in three Scaffolds Hma1.2p1\_0060F.1, Hma1.2p1\_0653F.1 and Hma1.2p1\_0669F.1. According to the DNA sequence surrounding the 285 SNPs

收稿日期: 2023-02-08 修回日期: 2023-03-20 网络出版日期: 2023-04-06

URL: <https://doi.org/10.13430/j.cnki.jpgr.20230208004>

第一作者研究方向为风景园林植物研究, E-mail: tqqqq77@sjtu.edu.cn

通信作者: 刘群录, 主要从事风景园林植物研究, E-mail: liuql@sjtu.edu.cn

秦 俊, 主要从事园林植物及应用研究, E-mail: qinjun03@126.com

基金项目: 上海市绿化和市容管理局科学技术攻关项目 (G232406, G212410)

Foundation project: Science and Technology Research Project of Shanghai Greening and City Appearance Administration (G232406, G212410)

associating with floral phenotypes, a total of 1287 genes were annotated. In these genes, four genes putatively encoding for transcription factors, including two *MYBs*, one *MADS*, and one *bHLH*, were predicted to be associated with hydrangea inflorescence type. These findings provided a foundation for future gene cloning and molecular marker-assisted breeding for hydrangea.

**Key words:** *Hydrangea macrophylla*; inflorescence type; genome-wide association study; genotyping

绣球 (*Hydrangea macrophylla* (Thunb.) Ser.) 是虎耳草科绣球属的落叶灌木,其花色艳丽,花序硕大,可做切花、盆花、庭院绿化,具有极高的观赏价值,是绣球属中最具有经济价值的物种之一<sup>[1-2]</sup>。绣球在国内适生范围广,是国内绣球属中最主要的栽培种,市场需求大。但是我国种植的栽培品种多为国外引进,国内种质资源开发不足,需要加强绣球新品种的培育。传统育种如杂交、田间突变筛选等方法存在耗时长、种间杂交不亲和、结实率低等现象,耗费大量人力和财力<sup>[3]</sup>。而分子标记辅助育种的发展为育种开辟了一条新的途径。分子标记可以帮助育种人员更好地了解种质资源的遗传背景,关联特定性状,辅助筛选出符合育种目标的个体,大大缩短了新品种的选育时间。

花型育种是绣球花育种的一个重要方向。绣球有两种花型,一种为平顶型(Lacecap),其中央为非装饰花,周围为装饰花<sup>[4]</sup>,花序顶部较平;另一种花型为圆球型(Mophead),主要由装饰花组成,整体呈圆球形。由于绣球花中的装饰花观赏价值更高,在育种中以增加绣球装饰花的数量作为重要目标。而绣球兼具平顶型与圆球型两种花型,是研究绣球花型形成的良好材料<sup>[5]</sup>。

目前对绣球花型形成进行了初步研究。Reed等<sup>[6]</sup>基于SSR标记对绣球的遗传关系进行了研究,结果表明两种花型的绣球中SSR标记差异很大,圆球型和平顶型两种花型分别聚类于不同类群。Uemachi等<sup>[7]</sup>以绣球和粗齿绣球(*H. serrata* (Thunb.) Ser)为材料,研究发现在平顶型花型中,花序结构、花序轴节间长度都会影响装饰花的数量。Uemachi等<sup>[8]</sup>在1个平顶型花序群体中发现了一株圆球型的突变体,没有介于两种花型之间的中间型,将两种花型的绣球进行杂交, $F_1$ 皆为平顶型花序;自交 $F_2$ 中平顶型与圆球型的比例为3:1,符合孟德尔分离定律,由此推测花型由1个单基因控制。Waki等<sup>[9]</sup>将平顶型绣球分别与两个圆球型绣球品种杂交,后代均为平顶型,由此推测绣球花圆球型性状为隐性基因控制。

全基因组关联分析(GWAS, genome-wide

association study)是以连锁不平衡(LD, linkage disequilibrium)为基础,将分布在基因组中大量的分子标记与表型变异联系起来,是剖析植物复杂性状的有力工具<sup>[10-11]</sup>。GWAS基于高密度分子遗传标记,可以直接研究全基因组水平的DNA变异,在研究之前不需要构建任何假设<sup>[12-13]</sup>。利用GWAS技术可识别物种重要的SNP,并将其转化为易于使用的分子标记<sup>[10]</sup>,大大降低了大样本全基因组水平的分析成本,加快了植物分子育种的研究进程<sup>[12]</sup>。Morris等<sup>[14]</sup>通过GWAS分析定位了几个与高粱长花序分支显著相关的基因。Chong等<sup>[15]</sup>对菊花切花的107个花序相关性状进行了GWAS分析,最终筛选出了影响花序大小的15个关键基因,包括*FT*、*GBF1*等。孟歌等<sup>[16]</sup>也通过GWAS对桃花的花型性状定位了与其相关的SNP和基因,如基因*Prupe.8G118100*,该基因含有B3-DNA结构域,推测其可能参与调控桃花花型。Wu等<sup>[17]</sup>通过简化基因组测序对82份绣球属植物的花序类型与DNA序列之间的关系进行研究,发现了与花序类型有关的SNPs位点。

在前期对于绣球花型的相关研究中,已通过杂交实验、SSR标记遗传分析以及GWAS等方法进行分析,并推测该性状为单个隐性基因控制。但是根据观察,绣球还存在着介于平顶型和圆球型的中间花型。因此推测绣球花型的形成可能存在着更为复杂的机制。2020年6月发布了绣球品种青之岛(*H. macrophylla* 'Aogashima')的基因组,为利用重测序进行绣球GWAS分析提供了条件。相比于简化基因组测序,重测序基因组覆盖度高,变异信息更丰富,并且可以分析SNP位点关联区域内的基因信息。本研究将绣球花型性状与重测序开发的SNP进行全基因组关联分析,定位与花型显著关联的SNP位点,随后利用不同绣球群体对初步鉴定的SNP位点进行验证分析,为后续绣球品种目标花型选育与分子育种提供参考。

## 1 材料与方法

### 1.1 试验材料

(1)GWAS供试材料:本研究所选群体为 $F_1$ 杂

交群体(图1),母本蒂沃利(Tivoli)为圆球型花序,父本塔贝(Taube)为平顶型花序。2020年5月选取群体中125个处于开花期的绣球植株作为研究材料,在杭州画境种业有限公司青山湖研发基地(119°48'N,30°14'W)进行花型观测,在开花期记录每个植株花序类型。本研究选用的绣球群体共有平顶型与圆球型两种花序类型,125株绣球植株中平顶型花序共71株,圆球型花序共54株。于2021年4月采摘新鲜无病虫害嫩叶,立即置于液氮,进行全基因组重测序和数据分析,获得SNP标记数据。



A1~D6: 平顶型; E1~H6: 圆球型  
A1~D6: Laccap inflorescence; E1~H6: Mophead inflorescence  
图1  $F_1$  绣球的两种花型

Fig.1 The inflorescence types of the *H. macrophylla* hybrids

(2) SNP 验证材料: 验证材料包括3部分, 第1部分为GWAS群体中的34个单株, 包括平顶型花序16株和圆球型花序18株; 第2部分为52个绣球品种, 其中平顶型花序和圆球型花序各26个品种(表1)。第3部分包括7个 $F_1$ 和1个 $F_2$ 群体。两个 $F_1$ 群体植株为圆球型花序, 另两个 $F_1$ 群体植株为平顶型花序, 另外3个 $F_1$ 群体和1个 $F_2$ 群体则表现为既有圆球型花序又有平顶型花序(表2)。其中绣球品种与杂交群体分别于2019-2022年间种植于上海辰山植物园资源苗圃和杭州画境种业有限公司青山湖研发基地内。

表1 52个绣球品种列表

Table 1 The list of the 52 hydrangea cultivars

花序 Inflorescence	品种名称 Cultivars name
圆球型 Mophead	史欧尼、蓝色妈妈、望月、汉堡、阿耶莎、奇迹、马雷夏尔、你我的灵感、美丽鲍茨、无尽夏、塞布丽娜、魔幻紫水晶、国产绣球、玫瑰女王、蒂沃利、魔幻珊瑚、拉维布兰、雪球、八重、甜蜜幻想、康士坦、蒂亚娜、水晶绒球、魔玉、火红、蓝宝石
平顶型 Laccap	城崎、中间泽、鸽子、柠檬波、粉色回忆、水平、棉花糖、白花、阿耶莎、三原八重、普利贾、爱你的吻、玫瑰花束、泉鸟、塔贝、银边、森乃妖精、初霜、狐狸、落日、拉娜白、完美玛丽斯、蓝山雀、耐尤斯、初恋、姑娘

表2  $F_1$ 和 $F_2$ 亲本及花型表型

Table 2 The parents of the  $F_1$  and  $F_2$  and the inflorescence types of the hybrids

杂交群体 Hybrid population	亲本 Parents	表型 Phenotypes	数量 Quantity
$F_1$	魔玉×卡米拉	圆球型	13
	花火×蒙娜丽莎	圆球型	24
	星知樱×卡米拉	平顶型	10
	泉鸟×卡米拉	平顶型	13
	蒂沃利×塔贝、 梦幻蓝×塔贝、 塔贝×白塔贝	圆球型、平顶型	26
$F_2$	$F_1$ 混合播种	圆球型、平顶型	22

于2022年6-8月分别采集目标群体植株新鲜、健康的叶片, 其中每个绣球品种取3株植株叶片作为生物学重复。同时在开花时记录花序类型。使用多糖多酚植物基因组提取试剂盒(天根生化科技(北京)有限公司)提取基因组总DNA, 使用NanoDrop 2000/2000c分光光度计和1.0%的琼脂糖凝胶电泳测定DNA浓度和纯度。

## 1.2 重测序与SNP数据集构建

提取125株绣球花植物基因组DNA后, 样本DNA经琼脂糖凝胶电泳和紫外分光光度计检测, 采用机械方法(超声)打断, 通过片段纯化、末端修复、3'端加A、连接测序接头、PCR扩增形成测序文库, 质控(双端测序)后在Illumina平台上测序。对测序得到的原始数据进行过滤, 去除含有接头序列、N含量超过10%的序列以及含有超过序列长度50%的低质量(Q≤10)碱基的序列。以绣球品种青之岛(HMA\_r1.2\_1为二倍体共36条染色体)为参考基因组(<https://plantgarden.jp/en/list/t23110>), 采用BWA<sup>[18]</sup>软件与绣球花参考基因组进行比较, 采用SAMTOOLS去除重复。平均映射率为99.20%。对比结果正常, 后续SNP采集可靠。

由于参考基因组的不完整性,采用GATK<sup>[19]</sup>软件对SNP数据进行进一步筛选,去除Phred值低于30(错误率>1%)的SNP位点,去除测序深度过低或过高的SNP位点,保留最小等位基因频率(MAF, minor allele frequency)≥0.05的位点<sup>[20]</sup>。5'端前几个碱基的错误率较高,因为在测序固定过程中,DNA分子之间的物理位置会有极少数重叠。此外,随着测序序列长度的增加,错误率也会增加,因此样本前4个碱基和后12个碱基的质量值会相对较低,中间测序碱基的质量值会较高。总体质量值高于Q30,即正确率高于99.9%。这些位点均匀分布在染色体上,经过质量控制和筛选,最终共获得1281302个SNP用于后续关联分析。

### 1.3 群体结构与亲缘关系分析

利用基因组重测序检测到的SNPs,设置MAF≥0.05,去除错误率大于1%的位点进行全基因组关联分析。采用Admixture软件,预先设定亚群数目(K值),K值设定为1至10,Length of burning period设定为10000,根据极大似然值原理确定最优分群数,并根据最大似然值绘制对应K值的群体结构分析结果,对聚类结果进行交叉验证,根据交叉验证错误率的谷值确定最优分群数。使用EIGENSOFT软件基于SNP数据,进行主成分分析(PCA, principal component analysis)。使用GCTA version 1.92.1软件,计算样品间亲缘关系K(Kinship)。

### 1.4 全基因组关联分析

基于重测序筛选的SNP标记,结合表型数据,利用Fastlmm和GEMMA软件,以主成分分析中的主成分分析值(PC)和亲缘关系分析矩阵为协变量,分别采用一般线性模型(GLM, general linear model)和混合线性模型(MLM, mixed linear model)对花型性状的表型值与基因型值进行关联分析得到关联

值。其中,一般线性模型以主成分分析值(PC)为协变量,而混合线性模型则以主成分分析值(PC)+亲缘关系矩阵(K)为协变量以降低结果中出现假阳性的概率。通过关联分析得到与花型显著关联的SNP标记,并计算不同SNP位点对表型的贡献率,另外为了避免多重比较中潜在的假阳性,通过Bonferroni检验调整全基因组显著性阈值,设置显著关联阈值 $-\log_{10}(P \text{ value})=9.0$ ,大于该阈值的SNP被认为是与花型相关联的位点。利用R语言包绘制SNP位点的曼哈顿图。

### 1.5 SNP标记筛选

对于全基因组关联分析得到的SNP位点根据重测序基因分型数据进行初步筛选,选择每个Scaffold中 $-\log_{10}(P)$ 值高的前10~20个SNP进行重测序基因分型与观测到的表型比对,筛选出在GWAS群体中分型结果好的位点进行进一步验证分析。

### 1.6 SNP标记分型

(1)测序法:根据青之岛参考基因组,针对验证位点的DNA序列设计一对特异性引物,使SNP位点位于设计的PCR产物序列之间,通过PCR扩增特定目标片段,扩增产物大小为250~450 bp,再将PCR产物经胶回收纯化后进行测序,获得结果后与参考基因组序列进行比对确认基因型。同时对文献[17]中发现的位点进行分型验证,记为Hma.wu。引物序列见表3。PCR反应体系为20 μL,含有10 μL 2×Taq PCR Mix预混液(生工生物工程(上海)股份有限公司),DNA模板(浓度80 ng/μL)1 μL,10 μmol/L正反向引物各1 μL,浓度为25 mmol/L的MgCl<sub>2</sub>溶液1 μL以及6 μL的ddH<sub>2</sub>O。反应程序为95 °C预变性3 min;95 °C变性30 s,51~66 °C退火30 s,72 °C延伸30 s,共35个循环;72 °C延伸10 min,4 °C保存。

表3 测序法的引物序列

Table 3 The primers used for sequencing

SNP位点 SNP Loci	正向引物(5'-3') Forward primers (5'-3')	反向引物(5'-3') Reverse primers (5'-3')	产物长度(bp) Product length
Hma1.2p1_0653F.1_460462	ATTCTGATCCGTTCCTGCTG	TCTGTGGCACTTTACCCCTT	399
Hma1.2p1_0669F.1_220821	ACATGGGACAACCTGGAAT	TTGTATAAGGGGGTATGTTT	379
Hma1.2p1_0669F.1_949341	ATGCGTGTATGTTTGTGTTTGA	TGGGGATTCTTTGATTGTACC	417
Hma1.2p1_0744F.1_396693	CACTGTTGAGGCATTCTG	GGAATCAAATGATACGCA	388
Hma1.2p1_0744F.1_678323	TACGGAATTTGGTTGTTACCTG	CCTTTCAGGCTCTGACTATG	301
Hma1.2p1_0744F.1_875534	CGGTGTAGTTCCTCCAAA	TTGTTGCTTCTAAAGGGTT	384
Hma1.2p1_1283F.1_78611	TCGGCAACCAACACCAATA	CGGCAGCTCCACAATCAC	259
Hma1.2p1_2149F.1_31384	ACCAACGAATGGGTCAATG	CATAACTGGGAGATGTGGGAT	367
Hma.wu	GCTACAGCATACTGATTATCTCC	TGGAGGTCTTAATGCTCATAGAA	285

(2)竞争性等位基因特异性PCR法(KASP):从参考基因中获得SNP位点上下游各200 bp的DNA序列,针对每个SNP位点设计一套引物,包括两条特异性正向引物和一条通用的反向引物,两条特异性正向引物3'端碱基分别与SNP位点突变和未突变碱基互补,两种碱基类型分别用VIC和FAM荧光标记。引物序列由LGC公司合成(表4)。反应体系共

10  $\mu$ L,包括1  $\mu$ L DNA模板(浓度60 ng/ $\mu$ L),5  $\mu$ L 2 $\times$  KASP Master mix, 0.14  $\mu$ L Primer assay mix(3条引物的混合物),以及4  $\mu$ L ddH<sub>2</sub>O。在ABI StepOne Plus 荧光定量PCR仪上进行扩增反应和KASP分析,反应程序为94  $^{\circ}$ C 预变性15 min;94  $^{\circ}$ C 变性20 s,61  $^{\circ}$ C 延伸1 min,之后每个循环降低0.6  $^{\circ}$ C,共10个循环;94  $^{\circ}$ C 变性20 s,55  $^{\circ}$ C 延伸1 min,共30个循环。

表4 用于KASP的引物序列

Table 4 The primers used for KASP

SNP位点 SNP Loci	特异性正向引物(5'-3') Specific forward primers(5'-3')	通用反向引物(5'-3') Universal reverse primers(5'-3')
Hma1.2p1_0060F.1_796640	CAAAAAGTTTTGACAAAGAAAGGTGTTTC CTCAAAAAGTTTTGACAAAGAAAGGTGTTTC	GGAAAAGAAAACGTTGAGAAAATCACCTTT
Hma1.2p1_0060F.1_1540773	AGAAAAGAAAGTCCATGCAGAGATAAG AGAAAAGAAAGTCCATGCAGAGATAAC	GGAAACATTTGCATTCAACCTGGTATT
Hma1.2p1_0653F.1_868484	CAAACACCTCTCTGTGAGTATTAGA CAAACACCTCTCTGTGAGTATTAGT	TGCGTGACCATTTGTCATATACGAAGAAA

引物中加下划线部分为SNP位点;实线为FAM荧光,虚线为VIC荧光

The underlined part of the primer is SNP site; Full line was FAM fluorescence and dotted line was VIC fluorescence

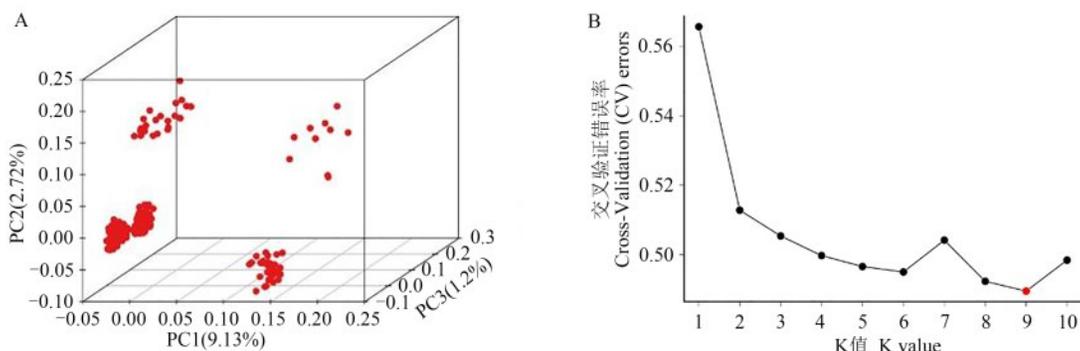
## 1.7 候选基因预测及分析

在获得与表型性状显著相关的SNP位点后,利用绣球品种青之岛的参考基因组对SNP位点附近DNA序列的功能进行注释,在获得基因注释信息后,与GO、KEGG、KOG数据库进行比对,获得基因功能注释。进一步筛选可能与绣球花型形成相关的基因。同时对筛选的基因进行生物信息学分析:从绣球基因组数据库(<https://plantgarden.jp/en/list/t23110>)中下载预测基因的CDS区核苷酸序列,使用NCBI-CDD(<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)和Pfam蛋白数据库(<https://pfam.xfam.org/>)进行保守结构域预测;通过NCBI的blast功能以及模式植物拟南芥数据库([\[arabidopsis.org/\]\(http://arabidopsis.org/\)\)进行同源性比对,并利用MAFFT进行蛋白序列比对,通过MEGA 11构建系统发育树。](https://www.</a></p>
</div>
<div data-bbox=)

## 2 结果与分析

### 2.1 群体结构与亲缘关系分析

通过主成分分析对绣球的群体结构进行评估。结果表明,125份重测序绣球材料在前3个主成分因子中形成了4个不同的簇(图2A),表明该群体积累了足够数量的遗传变异。而通过遗传聚类算法ADMIXTURE<sup>[21]</sup>得出当K=9时,样本具有最优的亚群体数量(图2B),该算法可用于找到一组K个祖先种群及其对当前样本的相对贡献,也证实了由主成分分析得到的结果。



A: 主成分分析, 图上的每个点对应一个样本; B: 根据不同K值计算的CV值, 红点表示最佳的K值

A: Principal component analysis. Each point on the graph corresponds to a sample; B: The calculated CV value using different K value.

The red point showed the best K value

图2 绣球群体结构与亲缘关系

Fig.2 The population structure and genetic relationship of hydrangea

## 2.2 花型性状全基因组关联分析

通过两种方式进行关联分析共筛选获得 369 个显著位点 ( $-\log_{10}(P) > 9$ ), 其中通过 Fastlmm 软件共筛选到 294 个显著位点, 通过 GEMMA 的混合线性模型共筛选到 360 个显著位点。两种模型共同筛选获得 285 个显著位点, 这些位点对表型变异的解释率在 11.80% ~ 60.54% 之间, 分布在 23 个 Scaffold 上

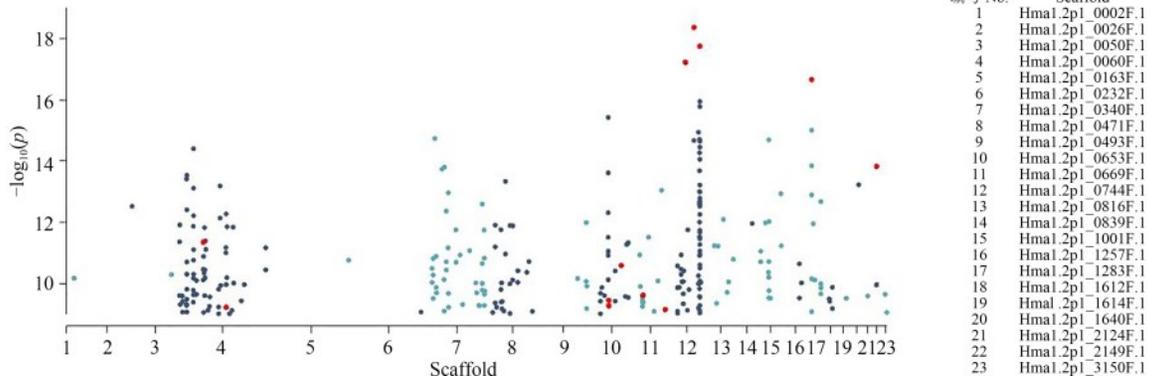


图3 花型性状全基因组关联分析

Fig.3 Genome-wide association study of the floral traits of hydrangea

## 2.3 花型 SNP 筛选

选择 GWAS 结果中每个 Scaffold 中  $-\log_{10}(P)$  值高的前 10~20 个 SNP 位点; 再根据重测序结果分析每个位点的基因型。结果表明位点 Hma1.2p1\_0060F.1\_796640、Hma1.2p1\_0060F.1\_1540773、Hma1.2p1\_0653F.1\_868484、Hma1.2p1\_0669F.1\_949341、Hma1.2p1\_0744F.1\_396693 与绣球花序类型显著相关。这 5 个 SNP 位点可分别将参与重测序的 125 个绣球植株分为 3 种基因型, 即两种纯合基因型和一种杂合基因型, 其中圆球形花型对应一种纯合基因型, 平顶型花型对应另一种纯合基因型以及杂合基因型。如圆球形花序在 Hma1.2p1\_0744F.1\_396693 位点的基因型主要为 G/G, 而该位点上圆球形花型植株中 95.83% 为 G/G 基因型。在位点 Hma1.2p1\_0060F.1\_796640、Hma1.2p1\_0060F.1\_1540773、Hma1.2p1\_0653F.1\_868484、Hma1.2p1\_0669F.1\_949341 的圆球形基因型分别为 T/T (85.71%)、C/C (95.11%)、T/T (86.96%) 及 G/G (86.96%)。

另外 5 个位点 Hma1.2p1\_0653F.1\_460462、Hma1.2p1\_0653F.1\_460468、Hma1.2p1\_0669F.1\_220821、Hma1.2p1\_1283F.1\_78611、Hma1.2p1\_2149F.1\_31384 同样也可分别将 125 个绣球植株分为 3 种基因型, 其中圆球形花型为其中的一种纯合基因型, 但是平顶型花型则有 3 种基因型。圆球形花序在

(图 3)。其中在 Scaffold Hma1.2p1\_0060F.1 上检测到了 67 个 SNP 位点, 在 Hma1.2p1\_0774F.1 上有 64 个 SNP 位点, Hma1.2p1\_0340F.1 有 39 个, Hma1.2p1\_0471F.1 有 24 个, Hma1.2p1\_0653F.1 有 21 个。在 Hma1.2p1\_0744F.1\_396693 标记处的 SNP 关联位点对表型变异的解释率最高, 为 60.54%。

Hma1.2p1\_1283F.1\_78611 位点的基因型为 C/C, 而该位点上圆球形花型植株中 98.87% 为 C/C 基因型; 在位点 Hma1.2p1\_2149F.1\_31384、Hma1.2p1\_0653F.1\_460462、Hma1.2p1\_0669F.1\_220821、Hma1.2p1\_0653F.1\_460468 的圆球形花序基因型分别为 G/G (100%)、G/G (100%)、G/G (100%)、A/A (95.12%)。

此外, 还选择了  $-\log_{10}(P)$  值较高的两个位点 Hma1.2p1\_0744F.1\_678323 和 Hma1.2p1\_0744F.1\_875534, 共 12 个 SNP 位点进行进一步验证分析。这 12 个 SNP 位点的详细信息如表 5 所示。

## 2.4 花型 SNP 分型

### 2.4.1 GWAS 群体样本及亲本分型

从原有 GWAS 群体中随机选择了 34 个单株, 并以两个亲本为材料, 分别利用测序法和竞争性等位基因特异性 PCR 法 (KASP) 两种方法对 SNP 位点 Hma1.2p1\_0060F.1\_796640、Hma1.2p1\_0060F.1\_1540773、Hma1.2p1\_0653F.1\_868484 进行验证。结果显示, KASP (图 4) 和测序法均能准确地将 36 个样本分为不同的基因型, 图中每个圆点代表一个绣球植株, 靠近横坐标轴的植株为纯合基因型, 位于对角线附近的植株则为杂合基因型。其中 Hma1.2p1\_0060F.1\_796640 结果为 17 个平顶型花序植株中 15 株为 C/T 基因型 (2 株结果缺失), 19 个圆球形花序植株全为 T/T 基因型, 其中亲本蒂沃利和塔贝分别为 T/T、C/T, 即杂合型基因型为平顶型花序, T/T

型纯合基因型为圆球型花序。位点 Hma1.2p1\_0060F.1\_1540773 和 Hma1.2p1\_0653F.1\_868484 与其

结果相似(表6)。3个位点的KASP验证结果与重测序分型结果一致,表明重测序结果的可靠性。

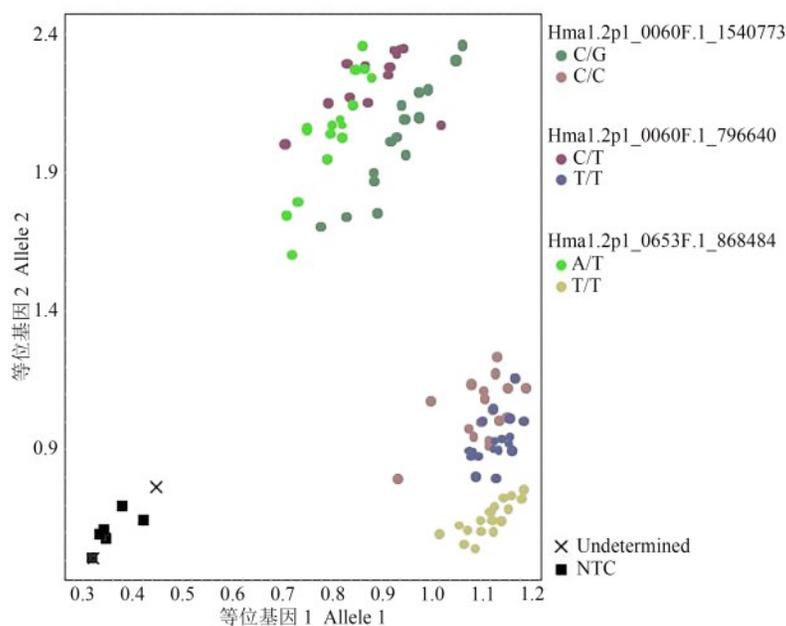
表5 绣球圆球型/平顶型花型性状显著关联位点信息

Table 5 The information of the loci significantly associated with the mophead/lacecap floral phenotypes of hydrangea

基因片段 Scaffold	显著关联SNP位置 Leading SNP position	$-\log_{10}(P)$ 值 $-\log_{10}(P)$	主要等位基因 Major allele	次要等位基因 Minor allele	解释表型变异率(%) Explain phenotypic variation rate
Hma1.2p1_0060F.1	796640	11.35	C	T	30.24
	1540773	9.22	G	C	51.48
Hma1.2p1_0653F.1	460462	9.26	G	A	20.24
	460468	9.44	A	G	20.32
	868484	10.58	A	T	50.09
Hma1.2p1_0669F.1	220821	9.61	G	A	37.04
	949341	9.14	G	T	45.11
Hma1.2p1_0744F.1	396693	17.23	G	A	60.54
	678323	18.37	G	T	25.85
	875534	17.76	G	A	25.18
Hma1.2p1_1283F.1	78611	16.67	C	T	29.91
Hma1.2p1_2149F.1	31384	13.83	G	A	25.30

主要等位基因为参考基因组中该位点核苷酸类型,次要等位基因为重测序中出现的另一种核苷酸类型

The primary allele is the nucleotide type at this site in the reference genome, and the secondary allele is another nucleotide type that appears in resequencing



NTC: 阴性对照, 用 ddH<sub>2</sub>O 代替模板 DNA; 下同

NTC: Negative control, ddH<sub>2</sub>O was used instead of template DNA; The same as below

图4 3个SNP位点KASP分型图

Fig.4 KASP genotyping plot of the three SNP locus

表6 KASP基因分型结果

Table 6 The genotyping results of KASP

花序类型 Floral types	SNP位点 SNP locus					
	Hma1.2p1_0060F.1_1540773		Hma1.2p1_0060F.1_796640		Hma1.2p1_0653F.1_868484	
	C/G	C/C	C/T	T/T	A/T	T/T
圆球型 Mophead	1	18	0	19	0	18
平顶型 Lacecap	15	1	15	0	16	1

**2.4.2 绣球品种及 F<sub>1</sub>、F<sub>2</sub> 杂交后代分型** 以 50 个绣球品种为材料,利用测序法和竞争性等位基因特异性 PCR 法对前期筛选的 12 个 SNP 都进行了基因分型验证。结果表明,每个品种的 3 次生物学重复结果均一致。其中除 Hma1.2p1\_0744F.1\_678323、Hma1.2p1\_0744F.1\_875534、Hma1.2p1\_1283F.1\_78611 以外,其他 9 个位点均在 50 个绣球品种中表现出了多态性。与 GWAS 群体中结果不同的是

SNP 位点 Hma1.2p1\_0060F.1\_796640、Hma1.2p1\_0060F.1\_1540773、Hma1.2p1\_0653F.1\_868484、Hma1.2p1\_0669F.1\_949341 的两种纯合基因型分别是圆球型和平顶型花序,而杂合基因型既有圆球型花序又有平顶型花序(图 5),且比例接近 1:1(表 7)。其余 SNP 位点虽表现出多态性,但与花序类型这一性状关联性不强。

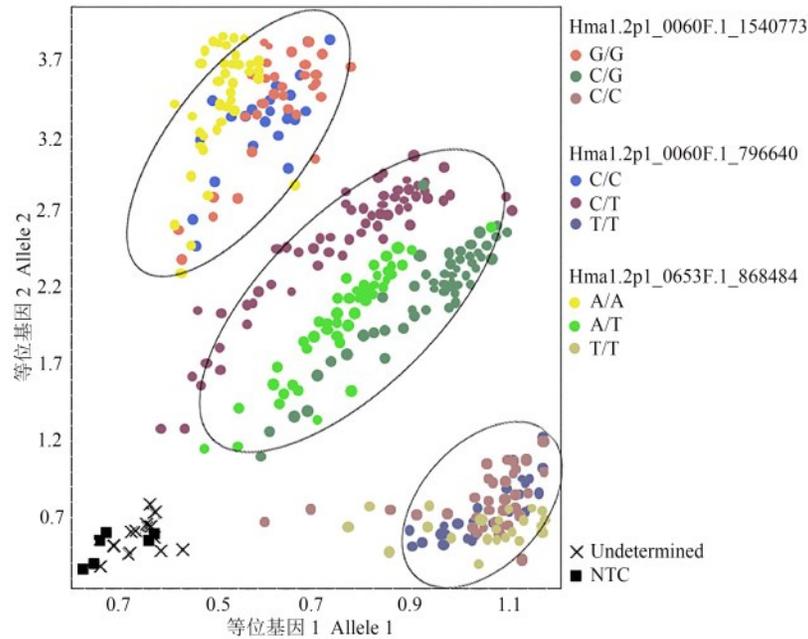


图 5 不同 SNP 分型结果

Fig.5 The genotyping results of SNP

表 7 50 个绣球品种分型结果

Table 7 The genotyping results of the 50 hydrangea cultivars

位点名称 Loci name	圆球型等位基因 Alleles of mophead			平顶型等位基因 Alleles of lacecap		
	A	H	B	A	H	B
Hma1.2p1_0060F.1_796640	3	13	9	14	11	1
Hma1.2p1_0060F.1_1540773	1	11	13	13	10	2
Hma1.2p1_0653F.1_868484	3	12	9	16	9	0
Hma1.2p1_0669F.1_949341	5	16	13	1	11	4
Hma.wu	0	0	26	7	16	3

A: 纯合基因型 1; B: 纯合基因型 2; H: 杂合基因型; 下同

A: Homozygous genotype1; B: Homozygous genotype 2; H: Heterozygous genotype; The same as below

此外根据 Wu 等<sup>[17]</sup>的引物序列通过测序法对其发现的 SNP 位点进行验证,记为 Hma.wu,结果如表 7、表 8 所示。绣球品种的分型结果与文献<sup>[17]</sup>一致,圆球型花序表现为单一类型的基因型即 T/T 纯合基因型,平顶型花序总体为 C/C、C/T 基因型(表 7)。在杂交群体的验证结果中圆球型花序依然为 T/T 基因型,平顶型花序则对应 3 种基因型

(表 8)。尽管对于该点的验证结果与本研究发现的位点分型结果不一致,但依然可以看出该点与绣球花序之间的相关性。

在 108 个杂交后代的验证中,12 个位点的整体结果与在品种验证中结果一致,Scaffold Hma1.2p1\_0060F.1、Hma1.2p1\_0653F.1 和 Hma1.2p1\_0669F.1 上的位点表现出与绣球花序的相关性(表 8),如位点

Hma1.2p1\_0060F.1\_1540773,圆球型花序中,3株为G/G基因型、21株为C/G基因型、19株为C/C基因型,而平顶型花序中,20株为G/G基因型、20株为C/G基因型、5株为C/C基因型,即纯合基因型C/C、G/G分别为圆球型、平顶型花序的概率更高,而杂合基因型C/G中既包含圆球型花序又包含平顶型花序。另外 Scaffold Hma1.2p1\_0653F.1 中的 460462 位点

在品种及杂交后代中表现出丰富的多态性(表9);除频率最高的G/G基因型,还包括A/G、A/A、G/T、T/T四种基因型;除G/G以外的基因型都表现为平顶型花序,而G/G基因型中则既有圆球型花序又有平顶型花序。在52个绣球品种及部分杂交群体分型结果中,共有12个平顶型绣球单株为除G/G外的基因型(表9)。

表8 F<sub>1</sub>、F<sub>2</sub>绣球植株分型结果

Table 8 The genotyping results of F<sub>1</sub>、F<sub>2</sub> generation hydrangea plants

位点名称 Loci name	圆球型等位基因 Alleles of mophead			平顶型等位基因 Alleles of lacecap		
	A	H	B	A	H	B
Hma1.2p1_0060F.1_796640	7	23	15	12	21	4
Hma1.2p1_0060F.1_1540773	3	21	19	20	20	5
Hma1.2p1_0653F.1_868484	8	20	16	19	16	3
Hma1.2p1_0669F.1_949341	9	11	10	12	9	4
Hma. wu	1	0	53	3	20	26

表9 位点 Hma1.2p1\_0653F.1\_460462 分型结果

Table 9 The genotyping results of Hma1.2p1\_0653F.1\_460462

位点名称 Loci name	圆球型等位基因 Alleles of mophead			平顶型等位基因 Alleles of lacecap		
	G/G	G/G	A/A	A/G	G/T	T/T
Hma1.2p1_0653F.1_460462	56	36	4	2	2	4

根据基因分型结果可知,与绣球圆球型/平顶型花序类型显著相关的位点可能位于 Scaffold Hma1.2p1\_0060F.1、Hma1.2p1\_0653F.1 和 Hma1.2p1\_0669F.1 上,且可能由多个位点共同控制。

## 2.5 候选基因预测

以绣球青之岛基因组为参考,在285个关联位点上下游100 kb范围内进行基因功能标注,共获得1287个基因,其中1114个功能标注基因,173个功能未知基因。基因功能涉及到DNA复制、信号转导机制、脂质和氨基酸的转运与代谢及物质运输等过程。根据前期对于SNP位点的验证情况,重点关注 Scaffold Hma1.2p1\_0060F.1、Hma1.2p1\_0653F.1 和 Hma1.2p1\_0669F.1 上的SNP位点和基因。根据基因功能注释,预测获得4个可能与花发育相关的候选基因,包括转录调控基因 *MYB*、*MADS* 和 *bHLH* (表10)。根据已有的研究,这三类转录因子家族中许多基因与花发育密切相关。其中有两个基因功能注释为转录因子 *MYB*,在参考基因组中的基因编号分别为 Hma1.2p1\_0653F.1\_g196210.gene 和 Hma1.2p1\_0669F.1\_g200310.gene。Hma1.2p1\_0653F.1\_g196210.gene 包含9个SNP位点,对表型变异解释率在17.43%~56.46%之间。Hma1.2p1\_0669F.1\_g200310.gene 内部包含1个SNP位点,对表型变异解释率为

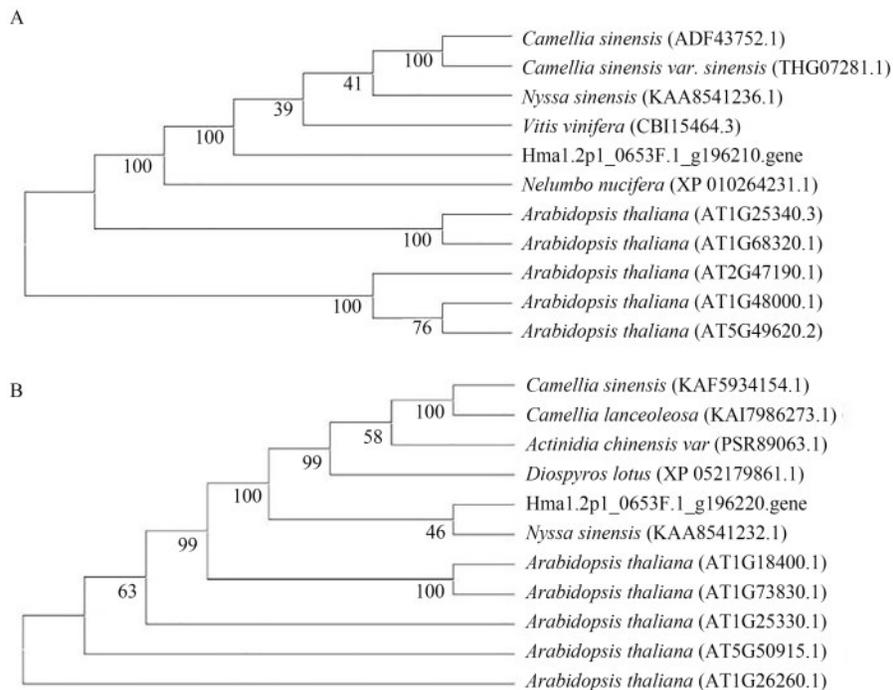
45.11%。基因编号为 Hma1.2p1\_0060F.1\_g032780.gene 的基因功能注释为 *MADS* 转录因子,包含11个SNP位点,对表型变异的解释率在36%~51.83%之间。基因编号为 Hma1.2p1\_0653F.1\_g196220.gene 的基因功能注释为 *bHLH* 转录因子,包含9个SNP位点,对表型变异的解释率为17.43%~56.46%之间。

由于绣球参考基因组序列不完整,目前只获得了基因 Hma1.2p1\_0653F.1\_g196210.gene 和 Hma1.2p1\_0653F.1\_g196220.gene 的完整CDS序列。其中 *MYB* 转录因子 Hma1.2p1\_0653F.1\_g196210.gene 基因的CDS全长912 bp,推测编码303个氨基酸;通过NCBI和Pfam在线工具发现 Hma1.2p1\_0653F.1\_g196210.gene 含有两个 *MYB* 保守结构域,预测其为典型的R2R3型 *MYB* 转录因子。将NCBI和拟南芥数据库中blast匹配度最高的前5个物种或基因进行蛋白序列比对,再用MEGA11构建系统发育树,可以看出 Hma1.2p1\_0653F.1\_g196210.gene 与葡萄 (*Vitis vinifera*) 中编号为 CBI15464.3 基因亲缘关系最近(图6A)。在拟南芥中与 Hma1.2p1\_0653F.1\_g196210.gene 亲缘关系最近的基因为 AT1G25340.3,其功能注释为在开花阶段表达于花瓣和萼片,参与DNA的转录调控。

表 10 GWAS 调控花序类型的候选基因预测

Table 10 Candidate genes prediction for inflorescence types regulation at GWAS

基因编号 Gene ID	功能注释 Gene annotation	SNP 位置 SNP loci	包含 SNP Containing SNP
Hma1.2p1_0653F.1_g196210.gene	MYB 转录因子	基因间隔区	Hma1.2p1_0653F.1_435347、Hma1.2p1_0653F.1_435301、 Hma1.2p1_0653F.1_435314、Hma1.2p1_0653F.1_436635、 Hma1.2p1_0653F.1_439519、Hma1.2p1_0653F.1_435224、 Hma1.2p1_0653F.1_439417、Hma1.2p1_0653F.1_460468、 Hma1.2p1_0653F.1_460462
Hma1.2p1_0669F.1_g200310.gene	MYB 转录因子	内含子区	Hma1.2p1_0669F.1_949341
Hma1.2p1_0060F.1_g032780.gene	MADS 转录因子	基因间隔区	Hma1.2p1_0060F.1_469796、Hma1.2p1_0060F.1_468584、 Hma1.2p1_0060F.1_469453、Hma1.2p1_0060F.1_460979、 Hma1.2p1_0060F.1_469557、Hma1.2p1_0060F.1_469349、 Hma1.2p1_0060F.1_473598、Hma1.2p1_0060F.1_469871、 Hma1.2p1_0060F.1_460851、Hma1.2p1_0060F.1_472680、 Hma1.2p1_0060F.1_465691、
Hma1.2p1_0653F.1_g196220.gene	bHLH 转录因子	基因间隔区	Hma1.2p1_0653F.1_435347、Hma1.2p1_0653F.1_435301、 Hma1.2p1_0653F.1_435314、Hma1.2p1_0653F.1_436635、 Hma1.2p1_0653F.1_439519、Hma1.2p1_0653F.1_435224、 Hma1.2p1_0653F.1_439417、Hma1.2p1_0653F.1_460468、 Hma1.2p1_0653F.1_460462



中华猕猴桃: *Actinidia chinensis* var; 拟南芥: *Arabidopsis thaliana*; 狭叶油茶: *Camellia lanceoleosa*; 茶: *Camellia sinensis*; 茶(原变种): *Camellia sinensis* var. *sinensis*; 君迁子: *Diospyros lotus*; 莲: *Nelumbo nucifera*; 蓝果树: *Nyssa sinensis*; 葡萄: *Vitis vinifera*

图 6 绣球 Hma1.2p1\_0653F.1\_g196210.gene 和 Hma1.2p1\_0653F.1\_g196220.gene 与其他物种进化树分析

Fig.6 Phylogenetic trees of Hma1.2p1\_0653F.1\_g196210.gene and Hma1.2p1\_0653F.1\_g196220.gene in hydrangea and other species

bHLH 转录因子 Hma1.2p1\_0653F.1\_g196220.gene 基因 CDS 区全长 774 bp, 推测编码 257 个氨基酸, 含有一个显著的 HLH 保守结构域。从进化树中可以看出该基因与蓝果树 (*Nyssa sinensis*) 中 KAA8541232.1 基因亲缘关系最近, 且与拟南芥中 AT1G18400.1 基因同源(图 6 B), 该基因在拟南芥花

序分生组织以及开花期间的花瓣及萼片中都有表达。

### 3 讨论

本研究在对 125 个绣球单株重测序的基础上, 对其花型性状进行了 GWAS 分析, 获得了 4 个与

绣球花序类型紧密关联的 SNP 位点,分别为 Hma1.2p1\_0060F.1\_796640、Hma1.2p1\_0060F.1\_1540773、Hma1.2p1\_0653F.1\_868484 和 Hma1.2p1\_0669F.1\_949341。由此认为绣球圆球型花序可能由多个基因控制,而非单基因控制。这与前人的研究结果不同。Uemachi 等<sup>[8]</sup>通过杂交育种试验发现后代性状分离比符合孟德尔分离定律,认为花型性状为单个主基因控制。Waki 等<sup>[9]</sup>通过 DNA 测序技术,设计 SSR 引物构建绣球遗传图谱,根据连锁分析结果认为绣球的重瓣和圆球型花序性状均由单个隐性基因控制。这些结论的差异可能与对绣球花序类型的划分有关。Uemachi 等<sup>[8]</sup>认为在圆球型花序和平顶型花序间没有中间性状<sup>[7]</sup>。但是根据观测,两种花序类型之间存在一些过渡的形态。如部分平顶型绣球花序中央,非装饰花中也会出现装饰花,且外围装饰花也并非规则的一轮(图 7),或是存在部分平顶型花序中有多轮的装饰花,而非装饰花只占中间很小的一部分,这种结构与普通的平顶型花序明显不同。Waki 等<sup>[9]</sup>也发现 F<sub>2</sub> 群体的圆球型与平顶型花型性状分离比约为 1:10,而不是预期的 1:3。因此本研究推测绣球的花序类型可能是一个复杂的数量性状,并且由多个位点多个基因共同控制。



图 7 平顶型和圆球型间过渡形态的绣球花序

Fig.7 *Hydrangea* inflorescences of transitional form between lacecap and mophead

根据分型结果,与绣球花型相关的基因可能在 Scaffold Hma1.2p1\_0060F.1、Hma1.2p1\_0653F.1 和 Hma1.2p1\_0669F.1 中。Wu 等<sup>[17]</sup>对 82 株绣球进行基于简化基因组测序的全基因组关联分析,获得了 1 个与绣球花序性状关联的 C/T 型 SNP 位点,但未明确该位点在基因组中的具体位置。根据文献<sup>[17]</sup>中提供的引物序列,进行 PCR 扩增及测序获得 DNA 片段序列,其中包含内切酶位点。将该 DNA 片段序列在绣球参考基因组中 blast 比对,根据文献<sup>[17]</sup>提到的酶切位点进一步确定该位点同样位于

Hma1.2p1\_0060F.1 上。

在这 3 个 Scaffold 中,选择与花型关联的 SNP 位点上下游 100 kb 范围内进行基因功能注释,发现了 4 个可能与绣球花型相关的基因,分别编码 MYB、MADS 和 bHLH 转录因子。MYB 转录因子在花的发育过程中起着重要作用,包括花粉发育和花色形成等方面。而在对桂花 R2R3-MYB 基因研究中,发现在不同组织中 *Of MYB5*、*Of MYB17* 两个基因在花芽中表达量最高,认为其可能参与了花开放过程的调控<sup>[22]</sup>;而在观赏凤梨中筛选出的 *AcMYB* 基因在催花 5d 后表达升高,其他时间均降低,推测其可能在花芽分化早期起作用<sup>[23]</sup>。MYB 转录因子还涉及花蕾发育的整个过程。以石斛为例,*DhMYB1* 基因的沉默可以使表皮细胞形态由锥形变为扁平,从而改变花瓣细胞形态<sup>[24]</sup>。在 MADS-box 转录因子家族中,*SEP (SEPAL-LATA)* 类基因通常在促进花芽萌发中发挥作用。在拟南芥中 SEP 类 MADS-box 转录因子调控花序发育,*SEP1*、*SEP2* 和 *SEP3* 是调控花器官发育的基因,这些基因在功能上是冗余的,是正常花瓣、心皮和雄蕊发育所必需的,缺乏所有 SEP 基因活性的三重突变拟南芥植株开花后所有花器官都表现为萼片<sup>[25]</sup>。bHLH 转录因子也在植物生长发育、花形态发生等过程中发挥重要作用。bHLH 家族 *PIF4* 蛋白在温度升高时与 *FT* 基因启动子结合,激活 *FT* 基因表达,促进植物开花<sup>[26]</sup>。此外,bHLH 基因还常常会与 MYB 等基因形成转录调控复合体进而调控植物的生长发育<sup>[27]</sup>,如在金鱼草中 MYB-bHLH-WD40 复合体会参与调控花瓣表皮圆锥形细胞的形成<sup>[28]</sup>。本研究鉴定结果可作为进一步研究植物生物学功能和花发育机制的基础。

#### 参考文献

- [1] 乔谦,王江勇,陶吉寒.绣球属植物研究进展.农学学报,2020,10(4):60-64  
Qiao Q, Wang J Y, Tao J H. Research progress of *Hydrangea*. Journal of Agriculture, 2020, 10(4): 60-64
- [2] Trankner C, Kruger J, Wanke S, Naumann J, Wenke T, Engel F. Rapid identification of inflorescence type markers by genotyping-by-sequencing of diploid and triploid F<sub>1</sub> plants of *Hydrangea macrophylla*. BMC Genetics, 2019, 20: 12
- [3] Reed S M. Compatibility studies in *Hydrangea*. Journal of Environmental Horticulture, 2000, 18: 29-33
- [4] Kitamura Y, Hosokawa M, Uemachi T, Yazawa S. Increase in the number of decorative florets in the inflorescence of *Hydrangea* through phytoplasma infection. Scientia Horticulturae, 2010, 124(1): 134-138

- [5] Jones K A, Reed S M. Production and verification of *Hydrangea arborescens* 'Dardom' x *H. involucrata* Hybrids. HortScience, 2006, 41:564-566
- [6] Reed S M, Rinehart T A. Simple-sequence repeat marker analysis of genetic relationships within *Hydrangea paniculata*. HortScience A Publication of the American Society for Horticultural Science, 2009, 44(1): 27-31
- [7] Uemachi T, Nishio T. Relationship between the inflorescence structure and setting of decorative flowers on Lacecap cultivars of *Hydrangea macrophylla* (Thunb.) Ser. and *H. serrata* (Thunb.) Ser. Horticultural Research, 2005, 4(4): 435-438
- [8] Uemachi T, Okumura A. The inheritance of inflorescence types in *Hydrangea macrophylla*. Journal- Japanese Society for Horticultural Science, 2012, 81(3): 263-268
- [9] Waki T, Kodama M, Akutsu M, Namai K, Iigo M, Kurokura T, Yamamoto T, Nashima K, Nakayama M, Yagi M. Development of DNA markers linked to double-flower and hortensia traits in *Hydrangea macrophylla* (Thunb.) Ser. Horticulture Journal, 2017, 87(2): 264-273
- [10] Myles S, Peiffer J, Brown P J, Ersoz E S, Zhang Z W, Costich D E, Buckler E S. Association mapping: Critical considerations shift from genotyping to experimental design. Plant Cell, 2009, 21(8): 2194-2202
- [11] Zhang D, Kong W Q, Robertson J, Goff V H, Epps E, Kerr A, Mills G, Cromwell J, Lugin Y, Phillips C, Paterson A H. Genetic analysis of inflorescence and plant height components in sorghum (Panicoidae) and comparative genetics with rice (Oryzoidae). BMC Plant Biology, 2015, 15:107
- [12] 张雁明, 邢国芳, 刘美桃, 刘晓东, 韩渊怀. 全基因组关联分析:基因组学研究的机遇与挑战. 生物技术通报, 2013(6): 1-6
- Zhang Y M, Xing G F, Liu M T, Liu X D, Han Y H. Genome wide association study: Opportunities and challenges in genomic research. Biotechnology Bulletin, 2013(6): 1-6
- [13] Aranzana M J, Kim S, Zhao K Y, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C L. Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. PLoS Genetics, 2005, 1(5): e60
- [14] Morris G P, Ramu P, Deshpande S P, Hash C T, Shah T, Upadhyaya H D, Riera-Lizarazu O, Brown P J, Acharya C B, Mitchell S E, Harriman J, Glaubitz J C, Buckler E S, Kresovich S. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proceedings of the National Academy of Sciences, 2013, 110(2):453-458
- [15] Chong X R, Su J S, Wang F, Wang H B, Song A P, Guan Z Y, Fang W M, Jiang J F, Chen S M, Chen F D, Zhang F. Identification of favorable SNP alleles and candidate genes responsible for inflorescence-related traits via GWAS in chrysanthemum. Plant Molecular Biology, 2019, 99: 407-20
- [16] 孟歌, 朱更瑞, 方伟超, 陈昌文, 王新卫, 王力荣, 曹珂. 桃的花型性状相关 SNP 位点挖掘与候选基因分析. 植物遗传资源学报, 2022, 23(2): 505-517
- Meng G, Zhu G R, Fang W C, Chen C W, Wang X W, Wang L R, Cao K. Genome-wide association study identified SNP alleles and candidate genes for flower shape trait in Peach (*Prunus persica*). Journal of Plant Genetic Resources, 2022, 23(2): 505-517
- [17] Wu X, Alexander L W. Genome-wide association studies for inflorescence type and remotancy in *Hydrangea macrophylla*. Horticulture Research, 2020, 7:27
- [18] Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics, 2009, 25 (14) : 1754-1760
- [19] Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 2010, 20(9): 1297-1303
- [20] Sunseri F, Lupini A, Mauceri A, De Lorenzis G, Araniti F, Brancadoro L, Dattola A, Gullo G, Zappia R, Mercati F. Single nucleotide polymorphism profiles reveal an admixture genetic structure of grapevine germplasm from Calabria, Italy, uncovering its key role for the diversification of cultivars in the Mediterranean Basin. Australian Journal of Grape and Wine Research, 2018, 24(3): 345-359
- [21] Alexander D H, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 2009, 19(9): 1655-1664
- [22] 姚亦凡, 董彬, 冯成庸, 杨丽媛, 赵宏波. 桂花 R2R3-MYB 家族基因鉴定及其在花开放过程中的表达分析. 园艺学报, 2020, 47(10):2027-2039
- Yao Y F, Dong B, Feng C Y, Yang L Y, Zhao H B. Identification of the R2R3-MYB family of *Osmanthus fragrans* and its expression in the process of flower opening. Acta Horticulturae Sinica, 2020, 47(10):2027-2039
- [23] 周璐. 观赏凤梨花芽分化的转录组学研究. 太原:山西大学, 2021
- Zhou L. Transcriptomic analysis of flower bud differentiation in ornamental pineapple. Taiyuan: Shanxi University, 2021
- [24] Lau S E, Schwarzacher T, Othman R Y, Ann J, Harikrishna. dsRNA silencing of an R2R3-MYB transcription factor affects flower cell shape in a *Dendrobium* hybrid. BMC Plant Biology, 2015, 15:194
- [25] Pelaz S, Ditta G S, Baumann E, Wisman E, Yanofsky M F. B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature, 2000, 405(6783): 200-203
- [26] Vinod K S, Doris L, Katja J, Enriqueta A, Elizabeth A, Nicholas P H, Philip A W. Transcription factor PIF4 controls the thermosensory activation of flowering. Nature, 2012, 484 (7393):242-245
- [27] 于冰, 田焯, 李海英, 吕笑言, 王宇光, 端木慧子. 植物 bHLH 转录因子的研究进展. 中国农学通报, 2019, 35(9):75-80
- Yu B, Tian Y, Li H Y, Lv X Y, Wang Y G, Duan-Mu H Z. Research progress of plant bHLH transcription factor. Chinese Agricultural Science Bulletin, 2019, 35(9):75-80
- [28] Glover B J, Perez-Rodriguez M, Martin C. Development of several epidermal cell types can be specified by the same MYB related plant transcription factor. Development, 1998, 125: 3497-3508