

Genetic Diversity and Population Structure Analysis of Bigleaf Hydrangea Using Genotyping-by-sequencing

Xingbo Wu and Lisa W. Alexander¹

U.S. Department of Agriculture, Agricultural Research Service, U.S. National Arboretum, Floral and Nursery Plants Research Unit, Otis L. Floyd Nursery Research Center, 472 Cadillac Lane, McMinnville, TN 37110

ADDITIONAL INDEX WORDS. GBS, *Hydrangea macrophylla*, molecular markers, ornamental plant breeding, SNP

ABSTRACT. *Hydrangea macrophylla* (bigleaf hydrangea) is one of the most important floral and nursery crops worldwide. However, breeding of new bigleaf hydrangea cultivars has been hampered by a long breeding cycle and lack of genetic resources. This study investigated the genetic diversity and population structure of 82 bigleaf hydrangea cultivars using single-nucleotide polymorphisms (SNPs) originated from genotyping-by-sequencing. A total of 5803 high-quality SNPs were discovered in a bigleaf hydrangea cultivar panel. A phylogenetic analysis and analysis of molecular variance based on discovered SNPs concluded the taxonomic classification of *H. macrophylla* ssp. *serrata* as a subspecies of *H. macrophylla*. Principal component analysis confirmed ‘Preziosa’ as a hybrid between *H. macrophylla* ssp. *macrophylla* and *H. macrophylla* ssp. *serrata*. In addition, the cultivar Lady in Red was also found to be a hybrid between the two subspecies. The population structure analysis identified three groups among the 82 cultivars. All *H. macrophylla* ssp. *serrata* cultivars belonged to one group, and two groups were revealed within *H. macrophylla* ssp. *macrophylla*. The separation within *H. macrophylla* ssp. *macrophylla* indicated a second gene pool due to breeding efforts that have targeted similar breeding goals for bigleaf hydrangea. The discovered SNPs and the phylogenetic results will facilitate further exploitation and understanding of phylogenetic relationships of bigleaf hydrangea and will serve as a reference for hydrangea breeding improvements.

The genus *Hydrangea* comprises ≈30 species that are widely distributed in America and Asia (McClintock, 1957). Famous for their large inflorescences surrounded by showy sepals, several *Hydrangea* species have been cultivated and are now some of the most popular ornamental flowering shrubs worldwide (Dirr and Dirr, 2004). Hydrangeas are one of the most economically important nursery crops in the United States, with sales exceeding \$120 million in 2014 (U.S. Department of Agriculture, 2014). Five species of *Hydrangea* (*H. macrophylla*, *H. paniculata*, *H. quercifolia*, *H. arborescens*, and *H. anomala* ssp. *petiolaris*) are widely cultivated in the United States as landscape plants (Dirr, 2009). Among them, *H. macrophylla*, also known as bigleaf hydrangea, is the most commonly cultivated species. One of the remarkable traits of bigleaf hydrangea is the aluminum (Al)-dependent flower color; plants may have pink or blue flowers in the presence of Al and a favorable substrate pH (Ma et al., 1997). Bigleaf hydrangeas are popular for their utility as landscape plants, potted plants, and cut flowers; therefore, traits such as flower and foliage color, sepal size and shape, and leaf characteristics have been enhanced through conventional breeding efforts that have led to more than 1000 named cultivars (van Gelderen and van Gelderen, 2004). These breeding efforts have focused largely on novel floral traits rather than disease resistance or environmental tolerance, and much room for improvement exists for a number of environmental and ornamental traits.

Improvements in environmental tolerance and combinations of desirable traits can be achieved through breeding approaches such as crosses within *H. macrophylla* or interspecific hybrid-

ization between *H. macrophylla* and closely related species (Rinehart et al., 2006). Limitations to hybrid breeding of *H. macrophylla* include lack of precise germplasm classification, difficulty recovering favorable trait combinations, and a long juvenile period before selection can occur. The taxonomy classification of the two most important *H. macrophylla* subspecies, ssp. *macrophylla* and ssp. *serrata*, has been disputed due to the lack of strong classification tools and convincing results (Reed and Rinehart, 2007). Most *Hydrangea* species are diploid, with a chromosome number of $2n = 36$ and genome size of 2.0 Gb (Cerbah et al., 2001); however, different ploidy levels among hydrangea species and within bigleaf hydrangea have been reported (Alexander, 2017; Jones et al., 2007; Sax, 1931; Schoennagel, 1931). Cultivars of *H. macrophylla* that are resistant to powdery mildew (*Erysiphe polygonii*) infection have been identified and include ‘Veitchii’, which has been widely used as a parent in crosses (Li et al., 2009; Windham et al., 2011). However, ‘Veitchii’ is not completely resistant to powdery mildew; it has lacecap flowers and ordinary foliage, which often make seedlings undesirable as new cultivars. Hybrids between *Dichroa febrifuga*, which is semi-evergreen and tropical, and *H. macrophylla* showed improved drought and sun tolerance but lacked the showy sepals that appeal to consumers (Jones and Reed, 2006; Reed et al., 2008). Closely related species with ploidy differences have produced fertile interspecific hybrids with *H. macrophylla*, suggesting that this breeding strategy could significantly expand hydrangea genetics over the long term, even when just using the species identified by Rinehart et al. (2006). Hybridizations between *H. macrophylla* and hydrangea species outside the group of closely related taxa have only been accomplished with embryo rescue, and offspring were not vigorous or fertile

Received for publication 8 Mar. 2019. Accepted for publication 21 May 2019.

¹Corresponding author. E-mail: Lisa.Alexander@ars.usda.gov.

(Kudo and Niimi, 1999; Kudo et al., 2002; Reed, 2004; Reed et al., 2001).

Releasing hydrangea cultivars requires screening of hundreds or thousands of seedlings over multiple generations to breed cultivars with the combination of desirable floral traits that consumers expect. As an obligate out-crossing species, *H. macrophylla* has high heterozygosity, and breeding is further hampered by long generation times typical of woody plants (Rinehart et al., 2016). Modern breeding methods that use molecular markers to identify desirable traits during the plant seedling stage could effectively reduce the breeding time, investment, and effort of cultivar release. However, unlike agronomic crops, little is known about the *H. macrophylla* genetic background because prior breeding efforts were mostly based on open-pollinated hybridizations rather than marker development. Simple sequence repeats markers (SSRs) have been developed and used to describe germplasm resources, parentage/pedigree, and intraspecific and interspecific hybrids (Rinehart et al., 2006, 2010). A phylogenetic tree of *Hydrangea* species was established to study the possibility of interspecific hybridization by using 13 highly polymorphic chloroplast markers (Mendoza et al., 2013). Recently, Waki et al. (2018) screened 768 SSR primer pairs in 93 F₂ progeny to identify markers linked to double-flower and hortensia (mophead) traits in *H. macrophylla*; they found that both traits are each controlled by a single recessive gene. Many SSRs were developed through transcriptome sequencing and used for genetic mapping of powdery mildew resistance, remontancy, and flower type in hydrangea, but they were insufficient for linkage mapping due to the lack of marker density (Rinehart et al., 2018).

Accelerated breeding of new *H. macrophylla* cultivars directly helps the nursery industry because consumer interest is driven, in part, by the release of new and novel plants. Conventional breeding can be vastly improved by investigating genetic availability in germplasm collections and increasing breeding accuracy and efficiency using molecular markers. In this regard, genotyping-by-sequencing can be used because of its utility for detecting large numbers of single-nucleotide polymorphism loci and rapidly genotyping diverse accessions. The goals of this research were to apply genotyping-by-sequencing to discover SNPs for bigleaf hydrangea cultivars, investigate the genetic diversity and population structure of *H. macrophylla* cultivars, and determine the disputed taxonomic classification of *H. macrophylla* ssp. *serrata*.

Materials and Methods

PLANT MATERIAL AND DNA EXTRACTION. A total of 87 bigleaf hydrangea cultivars representing two subspecies were included in this study (Table 1). Plants were collected from public or commercial sources and maintained in 14.6-L or 23.0-L containers at the Otis L. Floyd Nursery Research Center in McMinnville, TN. Plants were grown under 56% shade and micro-irrigated using spray stakes. Growing media consisted of pine bark amended with 6.6 kg·m⁻³ 19N-2.1P-7.4K controlled-release fertilizer (Osmocote Pro; ICL Specialty Fertilizers, Dublin, OH), 0.6 kg·m⁻³ micronutrients (Micromax; ICL Specialty Fertilizers), 0.6 kg·m⁻³ iron sulfate, and 0.2 kg·m⁻³ Epsom salts.

Leaf tissues were collected directly from plants into a 2-mL Eppendorf tube with lysis buffer and ground in a homogenizer

(FastPrep-24 5G; MP Biomedicals, Santa Ana, CA). Genomic DNA was isolated from the lysed tissue using DNeasy Plant Mini Kit (Qiagen, Valencia, CA), followed by RNeasy treatment. Nucleic acid was evaluated in 1% agarose gel and quantified using a spectrometer (NanoDrop 2000; Thermo Scientific, Wilmington, DE). Enzyme activity was evaluated using 6 U of *EcoRI* on 300 ng of eight randomly selected samples for 2 h at 37 °C and separated using gel electrophoresis on a 1% agarose gel. A lyophilized aliquot of 1.5 µg DNA was prepared for each sample and submitted to the Biotechnology Center of University of Wisconsin-Madison for library preparation and sequencing.

LIBRARY PREPARATION, SEQUENCING, AND VARIANT CALLING. Library and barcoding methods were performed according to Elshire et al. (2011). To find the best restriction enzyme and adaptor concentration for library preparation, a panel of eight samples was selected to find the optimized reactions that create a library, with the highest concentration at 200 to 250 bp, minimal residual adaptor contamination, and no visible indication of repetitive DNA contamination. Restriction enzyme *ApeKI* was used for DNA digestion, followed by ligation with a barcode and common adaptor. Single-end sequencing of the 95-plex library was performed with a single-lane sequencer (HiSeq. 2500; Illumina, San Diego, CA).

Because there is no available reference genome for the hydrangea species, the Java utility TASSEL UNEAK pipeline (Glaubitz et al., 2014) was used for the genotyping-by-sequencing data analysis. The barcoded sequence reads from each Illumina sequencing file were collapsed into a set of unique sequence tags with counts. The tag count files were filtered for a minimum count threshold of three and merged into the master tag count file. The occupancies of tags for each sample were observed from barcode information in the original FASTQ files. Parameters for the SNP call rate were set, with a maximum call rate of one and minimum call rate of zero to discover the genetic diversity of the diverse bigleaf hydrangea panel. Individual samples containing more than 40% missing data were excluded from the panel. Only SNPs with designated minor allele frequency (≥ 0.05) and few missing data (≤ 0.1) were retained for further diversity analyses.

PHYLOGENETIC AND GENETIC DIVERSITY. The filtered SNP data were summarized and converted to an appropriate format for further analyses using TASSEL V5.2.51 (Bradbury et al., 2007). A phylogenetic tree was drawn for all analyzed cultivars with the dendrogram based on the neighbor-joining option implemented in MEGA 7 (Kumar et al., 2016), with 500 bootstraps for nodal probability estimates. Genetic variation was calculated by an analysis of molecular variance (AMOVA) using Arlequin 3.5.2.2 (Excoffier and Lischer, 2010).

PRINCIPAL COMPONENT ANALYSIS AND POPULATION STRUCTURE. Population structure was achieved by using STRUCTURE 2.3.0 software with the admixture mode (Pritchard et al., 2000). The number of subpopulations, K, was set from 1 to 10 with 100,000 burn-in and 100,000 MCMC with 10 iterations. The output of STRUCTURE was evaluated using STRUCTURE HARVESTER (Earl, 2012) to determine the best K value based on the Evanno test (Evanno et al., 2005). The results were then summarized using CLUMPP (Jakobsson and Rosenberg, 2007) and visualized using DISTRUCT v1.1 (Rosenberg, 2004). The decomposition of the matrix was calculated using TASSEL and then plotted in R using the *autoplot()* function for principal component analysis (PCA).

Table 1. List of 87 hydrangea cultivars evaluated for genetic diversity using single-nucleotide polymorphisms detected via genotyping-by-sequencing.

| Cultivar | Subspecies ^z | Structure group |
|-------------------------------|----------------------------|-----------------|
| All Summer Beauty | <i>macrophylla</i> | II |
| Ayesha | <i>macrophylla</i> | III |
| Beaute Vendomoise | <i>macrophylla</i> | III |
| Big Daddy | <i>macrophylla</i> | III |
| Blauer Prinz | <i>macrophylla</i> | Mixture |
| Blauling | <i>macrophylla</i> | III |
| Blaumeise | <i>macrophylla</i> | Mixture |
| Blue Danube | <i>macrophylla</i> | III |
| Blue Billow | <i>serrata</i> | I |
| Blushing Bride | <i>macrophylla</i> | Mixture |
| Bridal Bouquet | <i>macrophylla</i> | Mixture |
| David Ramsey | <i>macrophylla</i> | II |
| Diadem | <i>serrata</i> | I |
| Domotoi | <i>macrophylla</i> | Mixture |
| Dooley | <i>macrophylla</i> | II |
| Double Dare | <i>macrophylla</i> | III |
| Elster | <i>macrophylla</i> | III |
| Endless Summer | <i>macrophylla</i> | II |
| Everlasting Amethyst | <i>macrophylla</i> | III |
| Everlasting Ocean | <i>macrophylla</i> | III |
| Everlasting Revolution | <i>macrophylla</i> | III |
| Fasan | <i>macrophylla</i> | III |
| Floralia | <i>macrophylla</i> | II |
| Forever and Ever | <i>macrophylla</i> | Mixture |
| Fuji Waterfall | <i>macrophylla/serrata</i> | III |
| General Vicomtesse de Vibraye | <i>macrophylla</i> | II |
| Glowing Embers | <i>macrophylla</i> | III |
| Grossman | <i>macrophylla</i> | II |
| Hamburg | <i>macrophylla</i> | Mixture |
| Hanabi | <i>macrophylla</i> | III |
| Heinrich Seidel | <i>macrophylla</i> | III |
| Holstein | <i>macrophylla</i> | Mixture |
| Horben | <i>macrophylla</i> | III |
| Kern | <i>macrophylla</i> | II |
| Kiyosumi | <i>serrata</i> | I |
| Koby | <i>macrophylla</i> | Mixture |
| Konigstein | <i>macrophylla</i> | III |
| LA Dreamin | <i>macrophylla</i> | III |
| Lady in Red | <i>macrophylla</i> | Mixture |
| Lemon Zest | <i>macrophylla</i> | Mixture |
| Lemonhoff | <i>macrophylla</i> | Mixture |
| Let's Dance Moonlight | <i>macrophylla</i> | Mixture |
| Libelle | <i>macrophylla</i> | III |
| Libelle White | <i>macrophylla</i> | III |
| Light of Day | <i>macrophylla</i> | III |
| Lilacina | <i>macrophylla</i> | Mixture |
| Little Geisha | <i>serrata</i> | I |
| Maculata | <i>macrophylla</i> | III |
| Madam Emile Mouilliere | <i>macrophylla</i> | Mixture |
| Mariesii Perfecta | <i>macrophylla</i> | III |
| Masja | <i>macrophylla</i> | III |
| Merritt's Supreme | <i>macrophylla</i> | III |
| Miss Hepburn | <i>macrophylla</i> | Mixture |
| Monmar | <i>macrophylla</i> | II |
| Mousseline | <i>macrophylla</i> | III |
| Niedersachsen | <i>macrophylla</i> | Mixture |

Continued next page

Table 1. Continued.

| Cultivar | Subspecies ^z | Structure group |
|------------------------------|----------------------------|-----------------|
| Nigra | <i>macrophylla</i> | III |
| Nikko Blue | <i>macrophylla</i> | II |
| Oregon Pride | <i>macrophylla</i> | III |
| PIA | <i>macrophylla</i> | III |
| Penny Mac | <i>macrophylla</i> | II |
| Pretty Maiden | <i>serrata</i> | I |
| Preziosa | <i>macrophylla/serrata</i> | Mixture |
| Princess Juliana | <i>macrophylla</i> | Mixture |
| Regula | <i>macrophylla</i> | III |
| Rhonda | <i>macrophylla</i> | II |
| Sabrina | <i>macrophylla</i> | Mixture |
| Sadie Ray | <i>macrophylla</i> | II |
| Sharona | <i>macrophylla</i> | III |
| Sol | <i>macrophylla</i> | Mixture |
| Souvenir du President Doumer | <i>macrophylla</i> | Mixture |
| Stella | <i>macrophylla</i> | Mixture |
| Sybilla | <i>macrophylla</i> | III |
| Taube | <i>macrophylla</i> | Mixture |
| Todi | <i>macrophylla</i> | Mixture |
| Trophee | <i>macrophylla</i> | III |
| VanHoose White | <i>macrophylla</i> | III |
| Veitchii | <i>macrophylla</i> | Mixture |
| Wayne's White | <i>macrophylla</i> | III |
| Weidler's Blue | <i>macrophylla</i> | III |
| White Wave | <i>macrophylla</i> | Mixture |
| Woodlander | <i>serrata</i> | I |
| Blue Wave | <i>macrophylla</i> | na ^y |
| Nachtigall | <i>macrophylla</i> | na |
| Eisvogel | <i>macrophylla</i> | na |
| La France | <i>macrophylla</i> | na |
| Kluis Superba | <i>macrophylla</i> | na |

^zmacrophylla = *H. macrophylla* ssp. *macrophylla*; serrata = *H. macrophylla* ssp. *serrata*; macrophylla/serrata = proposed hybrids between the two subspecies.

^yna = cultivar removed due to missing data.

Results and Discussion

GENOME-WIDE DISCOVERY OF SNPs. Sequencing of the *ApeKI* genomic complexity reduction libraries generated 24.8 Gb of raw sequence data; 26.3 million raw reads were produced through the sequencing. A total of 393,041 raw SNPs were discovered in the bigleaf hydrangea cultivar panel. SNPs were further filtered by removing any SNP locus with minor allele frequency that was less than 5% and excluding SNPs exceeding more than 10% missing data across the population. Five cultivars possessing high levels of missing data (>80%) were removed from the analysis. Therefore, a total of 5803 high-quality SNPs were identified in 82 bigleaf hydrangea cultivars for future study. Most of the SNPs identified in this study (62.54%) were A/G or C/T transition mutations, with the most common substitution type (31.42%) being A/G transitions (Fig. 1). Transversion-type SNPs, including A/C, A/T, C/G, and G/T conversions, represented the other 37.46% of total SNPs discovered. The least common substitution type was the C/G transversion (7.27%).

PHYLOGENETIC AND GENETIC DIVERSITY. A neighbor-joining phylogenetic tree based on the 5803 discovered SNPs classified the bigleaf hydrangea panel into two major groups (Fig. 2). The two major phylogenetic clades were separated according to

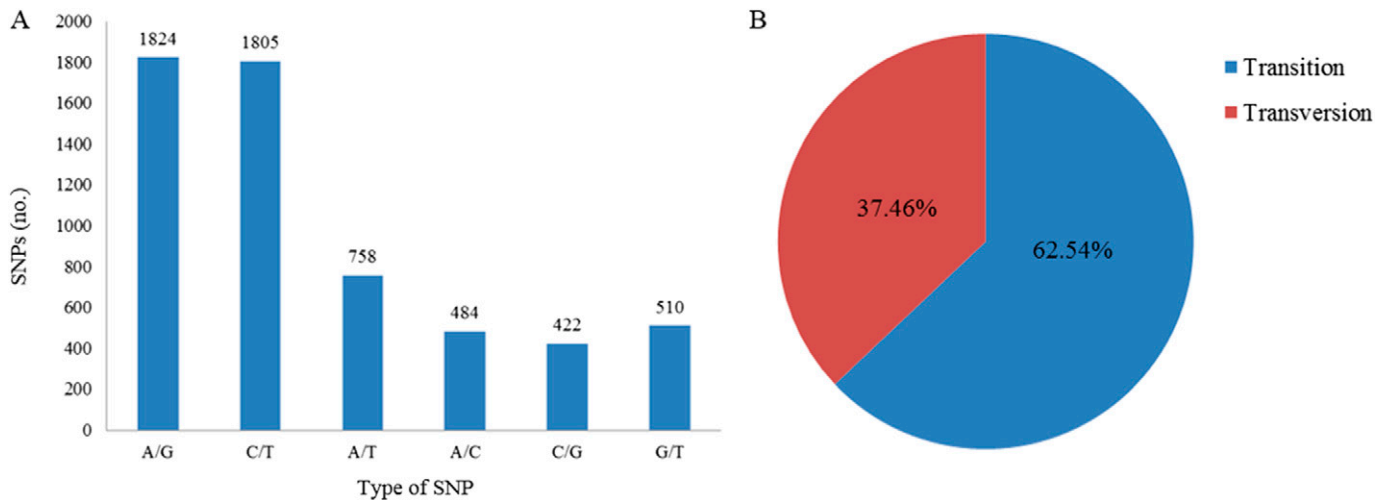


Fig. 1. Numbers and percentages of polymorphic single-nucleotide polymorphism (SNP) loci generated by genotyping-by-sequencing (GBS) of 82 *Hydrangea macrophylla* cultivars. (A) Numbers of nucleotide substitutions for the total 5803 SNPs. (B) Observed percentage of transition and transversion SNP types in *H. macrophylla* cultivars.

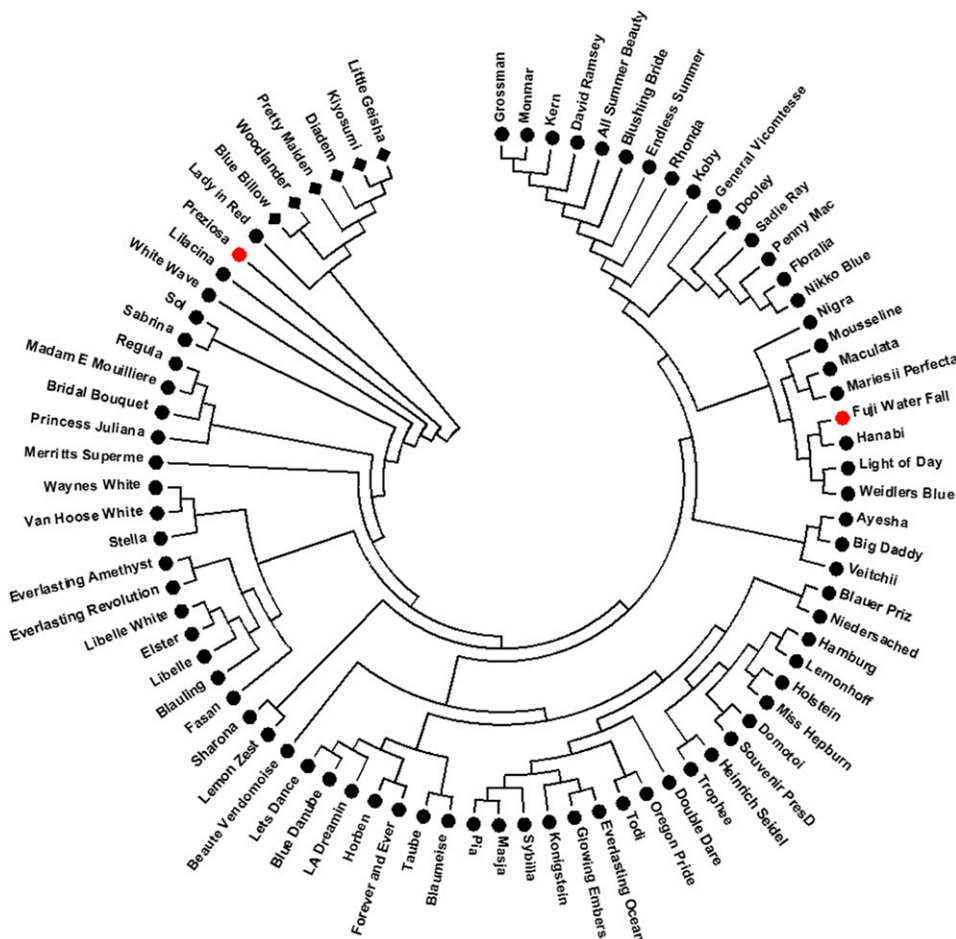


Fig. 2. Dendrogram of 82 *Hydrangea macrophylla* cultivars from two subspecies based on similarity calculated from 5803 genome-wide single-nucleotide polymorphisms. Subspecies *macrophylla* is represented by diamonds, subspecies *serrata* is represented by circles. Two proposed hybrids are colored red.

clustered into two different clades. One cultivar, Lady in Red, is commonly listed as *H. macrophylla* ssp. *macrophylla*, but it is clustered with the ssp. *serrata* group.

'Preziosa' and 'Fuji Waterfall' were proposed to be hybrids by Reed and Rinehart (2007). The genetic and population studies herein confirmed that 'Preziosa' is a hybrid between *H. macrophylla* ssp. *macrophylla* and *H. macrophylla* ssp. *serrata*. 'Preziosa' is a popular season-long bigleaf hydrangea cultivar with small, round mophead flowers and white, pale green, pink, or intense burgundy red color. However, the hypothesis that 'Fuji Waterfall' is a hybrid between the two subspecies was not supported by the current study. 'Fuji Waterfall' is a double-flower lacecap cultivar of ssp. *macrophylla* that was discovered near Mount. Fuji, Japan (Yamamoto, 2000) and recently imported to the United States (Reed and Rinehart, 2007). The genetic similarities between 'Lady in Red' and *H. macrophylla* ssp. *serrata* cultivars indicate that it is a hybrid between two subspecies. 'Lady in Red' is a pinkish white lacecap summer blooms cultivar with deep red stems and leaf veins that complement the lush green foliage that becomes reddish purple during the fall. It is often used as a foundation plant in many gardens.

Hybridization between these two subspecies was also observed in previous studies using a limited number of molecular markers but different cultivars (Reed and Rinehart, 2007). Interspecies crosses between

their subspecies category, ssp. *serrata* and ssp. *macrophylla*, respectively. Two proposed hybrids, 'Preziosa' and 'Fuji Waterfall', were both categorized as ssp. *macrophylla* but were

H. macrophylla × *H. angustipetala* and *H. macrophylla* × *H. arborescens* were proven to be successful (Cai et al., 2015; Kardos et al., 2009; Reed, 2000), with more hydrangea species (*H. sargentiana*, *H. integrifolia*, and *H. seemannii*) proposed to be potential candidates for future bridge crosses (Mendoza et al., 2013). Due to the restricted geographic distribution and similar breeding goals of various hydrangea breeding programs, the genetic diversity of hydrangea is limited and needs to be broadened to expand its cold hardiness and handle the challenges of abiotic and biotic stresses. Cold hardiness variations have been observed in different hydrangea species (Adkins et al., 2002). Therefore, hybridization of hydrangea at subspecies and species levels can be a great way to introduce elite variations (elite traits) for hydrangea genetic improvements toward cold hardiness breeding.

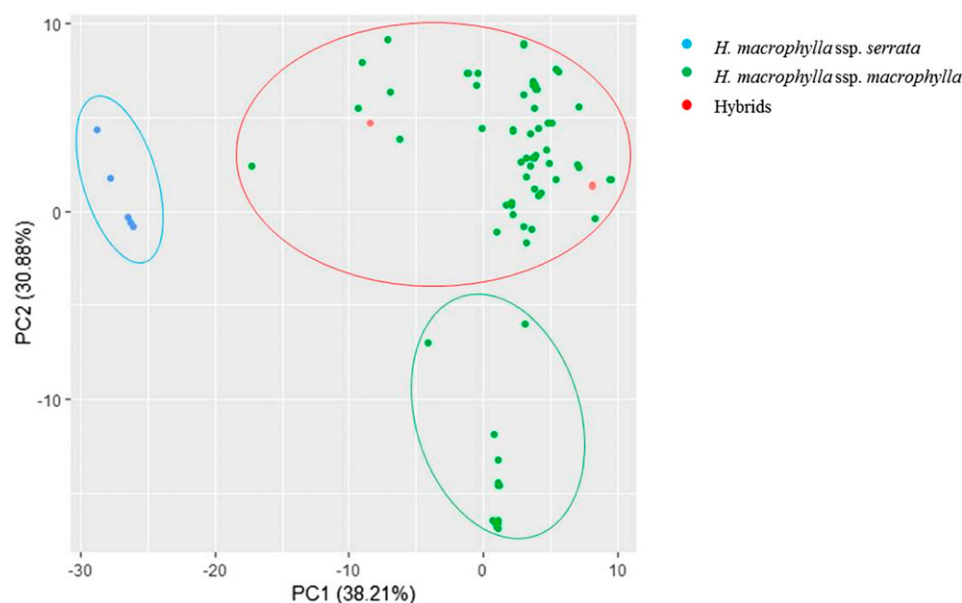


Fig. 3. Principal component (PC) analysis based on 5803 single-nucleotide polymorphisms from 82 *Hydrangea macrophylla* cultivars. Subspecies are color-coded; the three detected clusters are indicated by circles.

PCA AND POPULATION STRUCTURE. The PCA based on the genetic similarity of all samples confirmed separation of the two subspecies and revealed two distinct groups within the subspecies of *macrophylla* (Fig. 3). The first two PCs accounted for 69.09% of the total variation. Principal component 1 (PC1) grouped cultivars into two clusters that correspond to their subspecies, with cluster I representing *H. macrophylla* ssp. *serrata*. Principal component 2 (PC2) revealed the variability within ssp. *macrophylla* by the formation of two clusters (cluster II and cluster III). Cluster II contained 14 *macrophylla* cultivars that were all mopheads. Interestingly, the majority of cultivars in cluster II were of the remontant type. Cluster III included the remainder of the cultivars with a mix of mophead and lacecap flower types. However, lacecap flower cultivars such as Preziosa, Veitchii, Lilacina, Sol, White Wave, and Sabrina were grouped more closely than the other cultivars in cluster III. A lacecap cultivar, Lady in Red, commonly named ssp. *macrophylla*, showed a closer genetic relationship to ssp. *serrata*.

Structure analysis identified three populations in the cultivar panel based on the Evanno test, which corresponded to the both phylogeny and PCA results (Fig. 4). The three populations (groups I, II, and III) consisted of 6, 14, and 35 cultivars, respectively. Group I included all six cultivars from *H. macrophylla* ssp. *serrata*, corresponding to cluster I. Group II and group III consisted of all cultivars from *H. macrophylla* ssp. *macrophylla*. Of which, 14 cultivars that consisted of group II corresponded to cluster II and 35 cultivars from cluster III were in group III. The other ssp. *macrophylla* cultivars, including two proposed hybrids, Preziosa and Fuji Waterfall, were assigned to a mixture group.

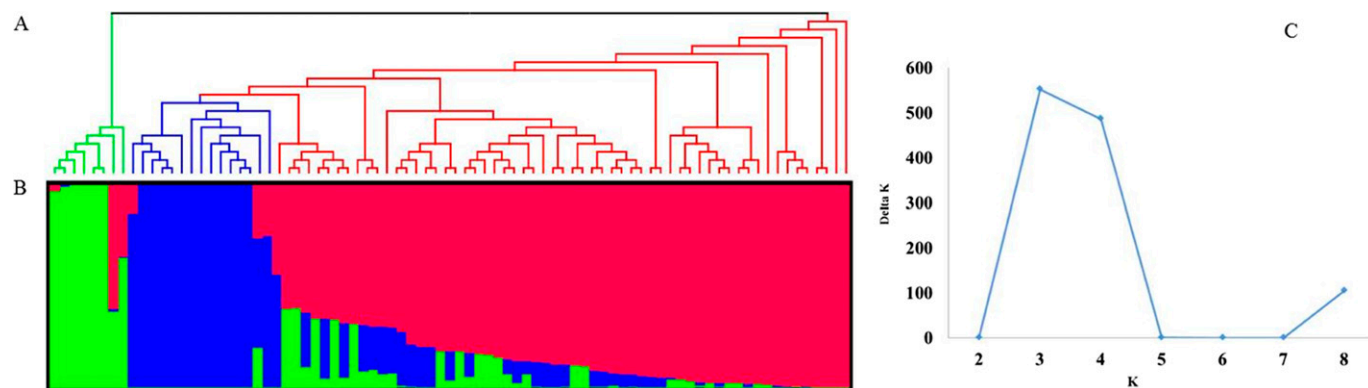


Fig. 4. Population structure analysis of 82 *Hydrangea macrophylla* cultivars based on 5803 single-nucleotide polymorphisms developed using genotyping-by-sequencing. Each individual is represented by a vertical bar reflecting assignment probabilities in each of the three groups. Group I: green bars; group II: blue bars; group III: red bars. (A) Phylogenetic tree based on the neighbor-joining method in MEGA 7 (Kumar et al., 2016). (B) Population structure based on a mixed-model analysis using STRUCTURE software (Pritchard et al., 2000). (C) Plot of the ΔK value with the number of subpopulations (K) from 2 to 10 based on the Evanno test (Evanno et al., 2005). Phylogenetic studies corresponded to population structure groups; ssp. *serrata* (green) agreed with group I, ssp. *macrophylla* was separated into two clades (blue and red) and corresponded to group II and group III, respectively.

Similar results from both population structure and multivariate analysis (PCA) indicated the formation of a second gene pool within *H. macrophylla* ssp. *macrophylla*, even though there is no clear phenotypic evidence. The grouping of famous *macrophylla* cultivars on “the best of best” list of garden hydrangea (Dirr and Dirr, 2004) and the desirability of blue-flowering mophead cultivars might have produced a second gene pool due to breeding efforts that targeted similar goals for bigleaf hydrangea. Hydrangeas are usually maintained and propagated clonally to preserve the genetic composition. Specific trait-driven breeding and selection programs of bigleaf hydrangea may narrow the gene pool, thereby causing a secondary *H. macrophylla* group with a specific genetic makeup that leads to novel traits. Flower color and shape are the other two important traits when breeding ornamental plants. In the floriculture industry, however, blue flowers of desirable plants have been difficult to breed. Many ornamental plants with a high production volume, such as rose (*Rosa* sp.) and chrysanthemum (*Chrysanthemum* sp.), lack the key genes for producing the blue delphinidin pigment or do not have an intracellular environment suitable for developing the blue color (Noda, 2018). However, in bigleaf hydrangea, the internal detoxification mechanism for Al stress develops blue coloration due to the interaction of delphinidin 3-glucoside, 5-caffeoyl-quinic acid and Al^{3+} (Ma et al., 1997; Yoshida et al., 2003), which gives it a great advantage compared with other nursery crops. Furthermore, the mophead inflorescence type, which is composed of many decorative sepals, adds more ornamental value compared with the lacecap inflorescence type (Uemachi and Okumura, 2012). Therefore, blue flowering and mophead inflorescences became the most important breeding targets for bigleaf hydrangeas, which led to a secondary gene pool for *H. macrophylla* ssp. *macrophylla*.

Regarding the taxonomic treatment of *H. macrophylla* subsection *Macrophyllae*, it has long been disputed whether ssp. *serrata* should be elevated to the species level. Morphological differences in flower shape and findings of genome size differences were the supporting evidence of separate species (van Gelderen and van Gelderen 2004; Zonneveld, 2004). Multiple methods, such as ecological tests, morphological traits, cytogenetics, and genetic studies, were used for the classification and reclassification of these two subspecies; however, they failed to reach a consensus due to the lack of comprehensive genetic studies (Reed and Rinehart, 2007). In nature, both subspecies are native to southern China and Japan, but they grow in different environments (McClintock, 1957). *Hydrangea macrophylla* ssp. *macrophylla* is found near the coast, whereas ssp. *serrata* grows in high-elevation areas and is often referred to as mountain hydrangea. To date, confusion that still persists needs to be eliminated with a comprehensive genetic study. In the present study, the phylogenetic tree based on our high-quality SNPs distinguished two geographically isolated hydrangea subspecies, and a new hybrid between them was identified along with other confirmed hybrids (Fig. 2). In breeding practices, crosses between ssp. *macrophylla* and ssp. *serrata* are helpful for making comparisons between interspecies hybridizations of other hydrangea species. The AMOVA test between the species and two subspecies levels indicated that genetic variation between all individuals within the *H. macrophylla* species (91.52%) is higher than that between these two subspecies (5.50%) (Table 2). The results of the current 5803 SNP markers agreed with those of a previous study using only 30 SSR markers (Reed and Rinehart,

Table 2. Analyses of molecular variance of two subspecies of *Hydrangea macrophylla* (ssp. *macrophylla* and ssp. *serrata*) based on 5803 single-nucleotide polymorphisms detected via genotyping-by-sequencing.

| | df | Sum of squares | Variance component | Proportion of variation (%) |
|--------------------------------------|-----|----------------|--------------------|-----------------------------|
| Among populations | 1 | 7,987.71 | 85.75 | 5.50 |
| Among individuals within populations | 81 | 124,844.95 | 47.26 | 3.03 |
| Within individuals | 82 | 119,951.04 | 1,427.99 | 91.52 |
| Total | 164 | 252,783.70 | 1,560.37 | |

2007). The conclusions of these results indicate that ssp. *serrata* is a subspecies of *H. macrophylla*.

A lack of genomic information hampers both the breeding of new hydrangea cultivars and the understanding of genetic mechanisms of aluminum tolerance in hydrangea. Various markers such as AFLPs, RAPDs, ISSRs, SRAPs, and SSRs have been developed for hydrangea, but they were limited by low numbers (Lee and Hyun, 2007; Mortreau et al., 2003; Reed et al., 2001; Rinehart et al., 2006). Sequence-based genotyping is a powerful tool for plant breeding and is especially helpful for plant species without reference genomes (He et al., 2014). In the present study, we discovered 5803 high-quality SNPs by genotyping a set of hydrangea cultivars with reduced-representation sequencing technology. This is the first SNP panel discovered in hydrangea, and it will greatly improve the hydrangea breeding program worldwide.

In summary, a panel of 5803 SNP markers was discovered in a bigleaf hydrangea cultivar panel using genotyping-by-sequencing. A detailed phylogenetic study based on newly discovered SNPs concluded the dispute regarding taxonomic classification of two well-known bigleaf hydrangea subspecies. Proposed hybrids between these two subspecies were verified and a new hybrid was identified by genetic studies. The population structure analysis and PCA based on large-scale SNP makers corresponded to each other and revealed a second gene pool within the *H. macrophylla* ssp. *macrophylla* cultivar, likely due to breeding efforts. Future studies will be focused on using the discovered SNPs to perform genome-wide association studies to identify SNPs associated with valuable agronomic traits such as inflorescence shape, remontancy, and the blue flowering mechanism of bigleaf hydrangea.

Literature Cited

- Adkins, J.A., M.A. Dirr, and O.M. Lindstrom. 2002. Cold hardiness potential of ten *Hydrangea* taxa. *J. Environ. Hort.* 20:171–174.
- Alexander, L. 2017. Production of triploid *Hydrangea macrophylla* via unreduced gamete breeding. *HortScience* 52:221–224.
- Cai, M., K. Wang, L. Luo, H.T. Pan, Q.X. Zhang, and Y.Y. Yang. 2015. Production of interspecific hybrids between *Hydrangea macrophylla* and *Hydrangea arborescens* via ovary culture. *HortScience* 50:1765–1769.
- Cerbah, M., E. Mortreau, S. Brown, S. Siljak-Yakovlev, H. Bertrand, and C. Lambert. 2001. Genome size variation and species relationships in the genus *Hydrangea*. *Theor. Appl. Genet.* 103(1):45–51.
- Dirr, M.A. and B.L. Dirr. 2004. *Hydrangeas for American gardens*. Timber Press, Portland OR.
- Dirr, M.A. 2009. *Manual of woody landscape plants: Their identification, ornamental characteristics, culture, propagation and uses*. Stipes Publ., Champaign, IL.

- Earl, D.A. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359–361.
- Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6(5):e19379.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14:2611–2620.
- Excoffier, L. and H.E. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564–567.
- Glaubitz, J.C., T.M. Casstevens, F. Lu, J. Harriman, R.J. Elshire, Q. Sun, and E.S. Buckler. 2014. TASSEL-GBS: A high capacity genotyping-by-sequencing analysis pipeline. *PLoS One* 9:e90346.
- He, J., X. Zhao, A. Laroche, Z.X. Lu, H. Liu, and Z. Li. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant Sci.* 5:484.
- Jakobsson, M. and N.A. Rosenberg. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806.
- Jones, K.D. and S.M. Reed. 2006. Production and verification of *Hydrangea arborescens* ‘Dardom’ × *H. involucrata* hybrids. *HortScience* 41:564–566.
- Jones, K.D., S.M. Reed, and T.A. Rinehart. 2007. Analysis of ploidy level and its effects on guard cell length, pollen diameter, and fertility in *Hydrangea macrophylla*. *HortScience* 42:483–488.
- Kardos, J.H., C.D. Robacker, M.A. Dirr, and T.A. Rinehart. 2009. Production and verification of *Hydrangea macrophylla* × *H. angustipetala* hybrids. *HortScience* 44:1534–1537.
- Kudo, N., Y. Kimura, and Y. Niimi. 2002. Production of interspecific hybrid plants by crossing *Hydrangea macrophylla* f. *hortensia* (Lam.) Rehd. and *H. quercifolia* Bartr. through ovule culture. *Engeigaku Kenkyuu* 1:9–12.
- Kudo, N. and Y. Niimi. 1999. Production of interspecific hybrid plants through cotyledonary segment culture of embryos derived from crosses between *Hydrangea macrophylla* f. *hortensia* (Lam.) Rehd. and *H. arborescens* L. *J. Jpn. Soc. Hort. Sci.* 68:803–809.
- Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33:1870–1874.
- Lee, J.H. and J.O. Hyun. 2007. The Use of AFLP markers for cultivar identification in *Hydrangea macrophylla*. *J. Korean For. Soc.* 96:125–130.
- Li, Y., M.T. Windham, R.N. Trigiano, S.M. Reed, J.M. Spiers, and T.A. Rinehart. 2009. Bright-field and fluorescence microscopic study of development of *Erysiphe polygoni* in susceptible and resistant bigleaf hydrangea. *Plant Dis.* 93:130–134.
- Ma, J.F., S. Hiradate, K. Nomoto, T. Iwashita, and H. Matsumoto. 1997. Internal detoxification mechanism of Al in hydrangea (identification of Al form in the leaves). *Plant Physiol.* 113:1033–1039.
- McClintock, E. 1957. A monograph of the genus *Hydrangea*. *Proc. Calif. Acad. Sci.* 14:147–256.
- Mendoza, C.G., S. Wanke, P. Goetghebeur, and M.S. Samain. 2013. Facilitating wide hybridization in hydrangea cultivars: A phylogenetic and marker-assisted breeding approach. *Mol. Breed.* 32:233–239.
- Mortreau, E., H. Bertrand, C. Lambert, and J. Lallemand. 2003. Collection of hydrangea: Genetic resources characterisation. *Acta Hort.* 623:231–238.
- Noda, N. 2018. Recent advances in the research and development of blue flowers. *Breed. Sci.* 68:79–87.
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Reed, S., G. Riedel, and M. Pooler. 2001. Verification and establishment of *Hydrangea macrophylla* ‘Kardinal’ × *H. paniculata* ‘Brussels Lace’ interspecific hybrids. *J. Environ. Hort.* 19:85–88.
- Reed, S.M. 2000. Development of an *in ovo* embryo culture procedure for *Hydrangea*. *J. Environ. Hort.* 18:34–39.
- Reed, S.M. 2004. Floral characteristics of a *Hydrangea macrophylla* × *H. paniculata* hybrid. *Proc. Southern Nursery Assn. Res. Conf* 49:580–582.
- Reed, S.M., K.D. Jones, and T.A. Rinehart. 2008. Production and characterization of intergeneric hybrids between *Dichroa febrifuga* and *Hydrangea macrophylla*. *J. Amer. Soc. Hort. Sci.* 133:84–91.
- Reed, S.M. and T.A. Rinehart. 2007. Simple sequence repeat marker analysis of genetic relationships within *Hydrangea macrophylla*. *J. Amer. Soc. Hort. Sci.* 132:341–351.
- Rinehart, T.A., P.A. Wadl, and M.E. Staton. 2018. An update on *Hydrangea macrophylla* breeding targets and genomics. *Acta Hort.* 1191:217–224.
- Rinehart, T.A., B.E. Scheffler, and S.M. Reed. 2006. Genetic diversity estimates for the genus *Hydrangea* and development of a molecular key based on SSRs. *J. Amer. Soc. Hort. Sci.* 131:787–797.
- Rinehart, T.A., B.E. Scheffler, and S.M. Reed. 2010. Ploidy variation and genetic diversity in *Dichroa*. *HortScience* 45:208–213.
- Rosenberg, N.A. 2004. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* 4:137–138.
- Sax, K. 1931. Chromosome numbers in the ligneous Saxifragaceae. *J. Arnold Arbor.* 12:198–205.
- Schoennagel, E. 1931. Chromosomenzahl und Phylogenie der Saxifragaceen. *Bot. Jahrb.* 64:266–308.
- Uemachi, T. and A. Okumura. 2012. The inheritance of inflorescence types in *Hydrangea macrophylla*. *J. Jpn. Soc. Hort. Sci.* 81:263–268.
- U.S. Department of Agriculture. 2014. Census of horticulture specialties. 18 July 2018. <<http://www.agcensus.usda.gov/Publications/2012>>.
- van Gelderen, C.J. and D.M. van Gelderen. 2004. *Encyclopedia of hydrangeas*. Timber Press, Portland, OR.
- Waki, T., M. Kodama, M. Akutsu, K. Namai, M. Ligo, T. Kurokura, T. Yamamoto, K. Nashima, M. Nakayama, and M. Yagi. 2018. Development of DNA markers linked to double-flower and hortensia traits in *Hydrangea macrophylla* (Thunb.). *Ser. Hort. J.* 87:264–273.
- Windham, M.T., S.M. Reed, M.T. Mmbaga, A.S. Windham, Y. Li, and T.A. Rinehart. 2011. Evaluation of powdery mildew resistance in *Hydrangea macrophylla*. *J. Environ. Hort.* 29:60–64.
- Yamamoto, T. 2000. *The Japanese hydrangeas color guide book*. Association Shamrock, Varengeville-sur-Mer, France.
- Yoshida, K., Y. Toyama-Kato, K. Kameda, and T. Kondo. 2003. Sepal color variation of *Hydrangea macrophylla* and vacuolar pH measured with a proton-selective microelectrode. *Plant Cell Physiol.* 44:262–268.
- Zonneveld, B.J.M. 2004. Genome size in *Hydrangea*, p. 245–251. In: C.J. van Gelderen and D.M. van Gelderen (eds.). *Encyclopedia of hydrangeas*. Timber Press, Portland, OR.