Development of InDel Markers for *Gypsophila paniculata* Based on Genome Resequencing

Bin Liu[1,2], *Chunlian Jin*[2], *Fan Li*[2], *Jihua Wang*[2*]

[1] College of Agriculture, Yunnan University [2] Floriculture Research Institute, Yunnan Academy of Agricultural Science

Correspondence:wjh@yaas.org.cn

1 Background

Gypsophila paniculata is usually used as a filler in flower arrangements, which making it an important cut flower in the global market. However, to our knowledge, the genetic research focused on *G. paniculata* is scarce, which hinders the improvement of the cultivars to some extent. Therefore, the construction of a genome-wide InDel marker system will boom the genetic studies *G. paniculata* and promotes its cultivar innovation and improvement.

2 Contents

1.In this study, we re-sequenced the whole genome of a wild-type accession of *G. paniculata* with white flower (WT-W) using next-generation sequencing technology. By comparing with the reference genome of *G. paniculata* (WT-P), a wild-type accession with pink flower, a series of molecular markers distributed genome-wide were identifie.

2. We developed a set of InDel markers with a high level of polymorphism using the information generated by genome sequencing of two wild-type accessions.

3 Result 1

The successful mapping of QTLs relies on the genetic maps with high density of molecular markers between the accessions.To develop sufficient molecular markers for *G. paniculata* genetic research, we detected sequence polymorphisms between WT-P (Figure 1A) and WT-W (Figure 1B). Different kinds of natural genetic variations were detected between the reference and resequencing genome, including 2,377,499 SNPs, 1,366,056 InDels, 1403 SVs, and 28 CNVs, whose densities were shown on the circus map (Figure 1C).

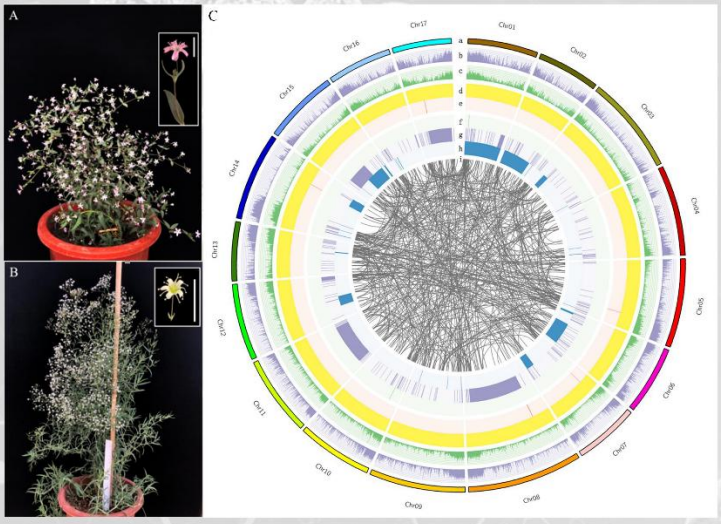

Figure 1. Resequencing of WT-W based on WT-P genome sequence. A. The pink flower wild-type of *G. paniculata* (WT-P). B. The white flower of *G. paniculata* (WT-W). C. Genomic structure variation distribution between the two *G. paniculata* wild-type accessions. a: reference sequence. b: SNP density distribution. c: InDel distribution density. d: CNV duplication. e: CNV deletion. f: SV insertion. g: SV deletion. h: SV inversion. i: SV translocation. Abbreviations include SNP: Single Nucleotide Polymorphism; InDel: Insertion/Deletion; CNV: Copy Number Variations; SV: Structure

4 Result 2

To develop InDel markers that can discriminate alleles between WT-P and WT-W, insertions or deletions over 10 bp were chosen as candidates with the interval of the neighbouring markers set as ~2 Mb. In total, 407 pairs of primers were designed for 17 chromosomes (Figure 2).To validate the newly designed markers, PCR analysis was conducted and the products were analysed by gel electrophoresis. Of the 407 markers, 289 markers distinguished the alleles of WT-P and WT-W clearly.

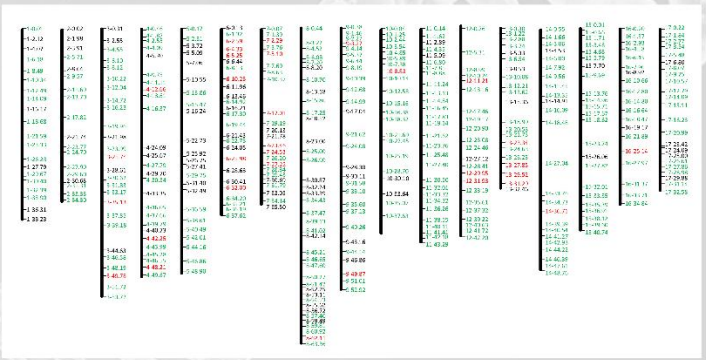

Figure 2. The physical map of 407 InDel markers distributed across all 17 chromosomes of *G. paniculata* genome. The name code of the InDel marker was presented as a chromosome number with the physical distance. Green markers discriminate alleles between WT-P and WT-W. Red markers amplified close bands on gel, and black markers were unavailable.

5 Result 3

To explore the applicability of the InDel markers designed in distinguishing the alleles between wildtype and commercial varieties, PCR amplification was conducted using the genomic DNA of WT-P and four commercial varieties (YX1-4) as templates. Out of the 407 pairs of primers, 191 were able to discriminate alleles between WT-P and commercial cultivars (Figure 3).

Figure 3. Matrix of the polymorphisms using the InDel markers among the five accessions of *G. paniculata*. Blue squares are WT-P bands, green, yellow and orange squares represent bands different from WT-P, and grey squares mean no bands detected.