

# A Phylogenomic Study of Korean *Hosta* Species Through Target Enrichment & Genome Skimming

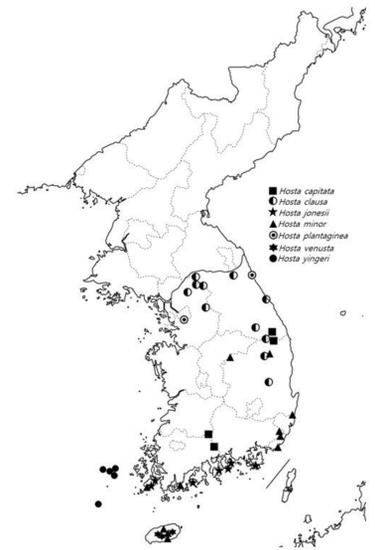
Mi-Jeong Yoo<sup>1</sup>, Byoung-Yoon Lee<sup>2</sup>, Sangtae Kim<sup>3</sup>, Suhyeon Park<sup>3,4</sup>, Dong-Pil Jin<sup>2</sup>, Chae Eun Lim<sup>2P\*</sup>

<sup>1</sup> Clarkson University, Potsdam, NY 13699, U.S.A.; <sup>2</sup> National Institute of Biological Resources, Incheon, Korea; <sup>3</sup> Sungshin Women's University, Seoul, Korea;

<sup>4</sup> InfoBoss Co., Ltd. and InfoBoss Research Center, Seoul, Korea.

## Abstract & Background

The genus *Hosta* (Agavoideae; Asparagaceae) is one of the most popular landscaping and ornamental plants native to temperate East Asia. In Korea, five to eleven species have been described (Chung and Kim, 1991), but the lack of taxonomically distinct traits makes it hard to reach a consensus regarding *Hosta* species entity. Here, we aimed to infer Korean *Hosta* species entity, their phylogenetic relationships, and estimate divergence time. To do so, we captured 251 single-copy nuclear gene sequences and plastid genomes from 51 accessions of seven Korean *Hosta* species (see Figure 1) using the Hyb-Seq method. Phylogenetic analyses of the captured sequences retrieved six species groups of *Hosta* in Korea: 1) *H. minor*/*H. venusta*, 2) *H. jonesii*, 3) *H. capitata*, 4) *H. clausa*, 5) *H. yingeri*, and 6) *H. plantaginea*. Based on the low sequence divergence and overlapping morphological characteristics, *H. venusta* was treated as *H. minor* var. *venusta*. Divergence time estimated here showed that *Hosta* diverged 38.38 million years ago, while Korean *Hosta* species rapidly diversified during the late Miocene. Last, we explored the possibility of these genomic data for inferring the origin of cultivated species. This study provides the most comprehensive genomic resources which can be used in the development of potential molecular markers for population and conservation studies of *Hosta*.

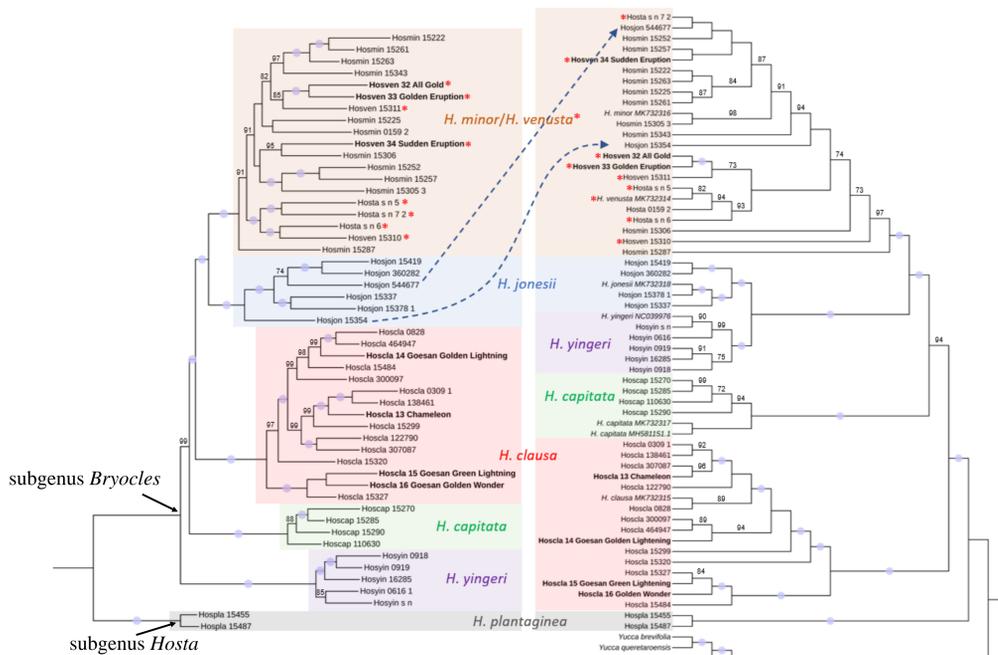


▲ Figure 1. Distribution of Korean *Hosta* species used in this study.

## Contents

### I. Phylogenomic analysis of Korean *Hosta* species

- ▶ On both nuclear and chloroplast trees (Figure 2), Korean *Hosta* species exhibited monophyly.
- ▶ Korean *Hosta* species were divided to six group with strong bootstrap value (100 %): 1) *H. minor*/*H. venusta*, 2) *H. jonesii*, 3) *H. capitata*, 4) *H. clausa*, 5) *H. yingeri*, and 6) *H. plantaginea*.
- ▶ *Hosta plantaginea* (subgenus *Hosta*) was revealed as sister to remained species (subgenus *Bryocles*), and it is thought to reflect distinctiveness of *H. plantaginea* such as geographic origin (from China) and unique flora characteristics.
- ▶ Position of *H. yingeri* (isolated in Heuksando Island) differs according to nuclear and chloroplast, but similarity of the genome size of *H. yingeri* (19.1±0.06 pg), *H. capitata* (19.3±0.06 pg) and *H. clausa* (19.2±0.18 pg) excluding *H. jonesii* (17.5±0.09 pg) supports the result of nuclear ML tree.
- ▶ Disparity of placement of *H. capitata* and *H. clausa* was also detected in both nuclear and chloroplast ML tree, which could be explained their rapid diversification. This was confirmed by short branch length of two species.
- ▶ Two accession of *H. jonesii* was nested in *H. minor*/*H. venusta* clade, implying the introgression from *H. minor* to *H. jonesii*.



▲ Figure 2. Comparison of maximum likelihood trees based on concatenated supercontigs (251 loci, 328,094 bp) with a model of GTR+F+R3 (the left tree) and plastome sequence matrix (159,451 bp) with a model of TPM2u+F+I (the right tree). The numbers on the node represent posterior probabilities (PP) from 1,000 bootstrap replicates. Purple circle on the node indicates 100% of PP. Bold taxa shows cultivated species, while red asterisks indicate *H. venusta*. The branch lengths in the chloroplast ML tree were ignored (see Figure S2 for the ML tree with branch lengths).

## Conclusions

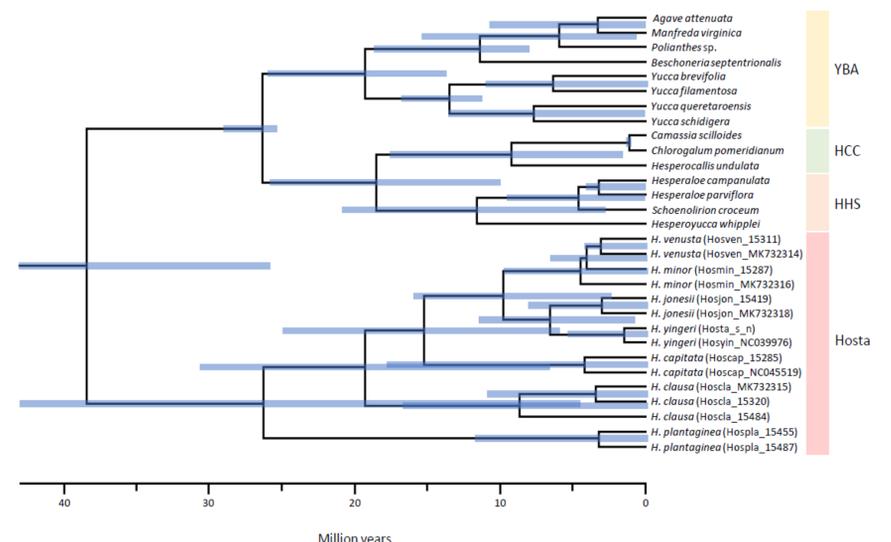
The Hyb-Seq method allows us to capture several hundreds of low or single-copy nuclear genes as well as obtain an organelle genome sequence. Due to low genetic diversities, commonly used nuclear and chloroplast markers have shown inconsistent results with low resolution in the phylogeny of *Hosta*. Here, the genomic data obtained with Hyb-Seq successfully identified six species groups of Korean *Hosta* along with their well-resolved and supported phylogenetic relationships. In addition, we first detected several introgressive hybridization events which have never been detected in the previous studies. Also, we demonstrated the applicability of Hyb-Seq for tracing the origin of cultivated *Hosta* species. This study provides the most comprehensive genomic resources which can be used in the development of potential molecular markers for population and conservation studies of *Hosta*.

## Reference

1. Chung, M. G., & Kim, J. W. (1991). The genus *Hosta* Tratt. (Liliaceae) in Korea. *SIDA*, 14, 411-420.
2. McKain, M. R., Wickett, N., Zhang, Y., Ayyampalayam, S., McCombie, W. R., Chase, M. W., . . . Leebens-Mack, J. (2012). Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). *Am J Bot*, 99(2), 397-406.
3. McKain, M. R., McNeal, J. R., Kellar, P. R., Eguarte, L. E., Pires, J. C., & Leebens-Mack, J. (2016). Timing of rapid diversification and convergent origins of active pollination within Agavoideae (Asparagaceae). *Am J Bot*, 103(10), 1717-1729.

### II. Divergence time estimation on Korean *Hosta* species

- ▶ Our divergence time analysis on *Hosta* species inferred that this genus was diverged during the late Eocene (38.38 Ma) (Figure 3), which is earlier than measurement of previous studies (during Oligocene (20~27 Ma)) (McKain et al., 2012, 2016).
- ▶ The 95% highest posterior density (HPD) intervals were very wide for *Hosta* clades, which might be due to low sequence divergence among *Hosta* species as shown their short branch lengths and low nucleotide diversities.
- ▶ Considering the divergence time (~ 10 Ma, late Miocene), the limited gene flow between Korean and Japanese *Hosta* species might have accelerated the speciation in Korea. It is because the Korean Peninsula and Japanese Islands were completely separated by East Sea during the late Miocene.



▲ Figure 3. Divergence time estimations for Korean *Hosta* species estimated from 79 protein-coding genes of the chloroplast genome. Blue bars represent 95% highest posterior density (HPD) intervals. YBA = *Yucca-Beschorneria-Agave*, HHS = *Hesperoyucca-Hesperaloe-Schoenolirion*, HCC = *Hesperocallis-Camassia-Chlorogalum*.

### III. Application of Hyb-seq results for identifying the origin of cultivar species

- ▶ We examined the seven cultivar species (three originated from *H. venusta* and four originated from *H. clausa*).
- ▶ In both ML trees, two *H. clausa* cultivars (Hoscla\_15 and Hoscla\_16) were close to Hoscla\_15327 collected from southern Korea. The other two cultivars were clustered with *H. clausa* species occurring in Northern South Korea (Figure 2), showing both nuclear and chloroplast genes can be used for tracing their potential parental species.
- ▶ The possibility was also proved with *H. venusta* cultivars which were included in *H. minor*/*H. venusta* clades with *H. venusta* (Hosven\_15311) and *H. minor* (Hosmin\_15257), which were collected from Jeju Island (Figure 1).