



INSTITUTIONEN FÖR BIOLOGI OCH MILJÖVETENSKAP

Polyploid Phylogenetics in Plants: Insights on non-model organisms in Fabaceae and Malvaceae

Jonna Sofia Eriksson

Institutionen för biologi och miljövetenskap
Naturvetenskapliga fakulteten

Opponent: Prof. Jonathan Wendel,
Department of Ecology, Evolution, and Organismal Biology,
Iowa State University, USA

Examinator: Prof. Mari Källersjö
Institutionen för Biologi och Miljövetenskap

Akademisk avhandling för filosofie doktorsexamen i naturvetenskap, inriktning biologi, som med tillstånd från Naturvetenskapliga fakulteten kommer att offentligt försvaras fredagen den 23 maj 2019 kl. 15.00 i Hörsalen, institutionen för biologi och miljövetenskap, Carl Skottsbergs gata 22B, Göteborg.

ISBN: 978-91-7833-241-0 (tryckt version)

ISBN: 978-91-7833-242-7 (pdf)

Tillgänglig via <http://hdl.handle.net/2077/59804>

Abstract

Genome duplication is a common phenomenon in angiosperms and advances in sequencing technologies and bioinformatics is now revealing its prevalence and significance. In a polyploidization event all genes in a genome are doubled. Duplicated genes can take one of three paths, either both of the duplicated are maintained or one of them may be randomly lost or selected against. Polyploid species are challenging for resolving species relationships because of the number of duplicated genes and the different processes leading to genome reduction.

In this thesis I investigate the mode of polyploid origin (i.e. auto- and allopolyploidy) and the role of ancient genome duplication in plant speciation. Here, two families were studied because of their numerous rounds of polyploidization events, Fabaceae and Malvaceae. To discern between the complex processes involving polyploidy, large amounts of data were generated from several nuclear genes using target gene capture and Illumina sequencing.

In *Medicago* (Fabaceae) two modes of polyploidy (autopolyploidy and allopolyploidy) were discovered. In the first case, an autopolyploid mode was identified, because gene comparisons showed two independently evolving species, a cryptic tetraploid species derived from a diploid progenitor. In the second case, two woody tetraploids were found to be hybrids. Since the closest parental lineages associated with the hybridization only have woody roots, the genome duplication for genes related to woodiness may be an instance of transgressive phenotype (extreme morphological character) in the hybrids.

Evidence of genome duplication appears not only in recently formed polyploids, but may also be present through genes that were duplicated in a lineage's history. The traces of ancient genome duplication events may be scarce owing to random mutations and gene losses. *Hibiscus* (Malvaceae) is a plant genus which possesses diverse chromosome numbers among species, indicative of potential polyploidization events. By studying the number of gene copies in diploid species, two ancient genome duplications were identified in *Hibiscus*. Additionally, numerous polyploidization events following the ancient duplications were detected among the extant species, indicating a complex reticulate history.

Hibisceae consists of five major clades: /Calyphylli, /Euhibiscus, /Furcaria, /Trionum and /Megistohibiscus but with inconsistent genus naming conventions (e.g. *Hibiscus* occurs across all of the clades). In this study, phylogenetic analysis of HTS data supported the classification of the group into the major clades; all were found to be monophyletic and no hybrid polyploidization events were found to have occurred between them. Additionally, each major clade's taxa were found to have common base chromosome numbers. Given the results, a taxonomic renaming, based on base chromosome number, of the major clade's genera is recommended.

This thesis demonstrates a rich history of polyploidizations both recent and ancient – highlighting the important role this phenomenon has played in the evolution of two distantly-related plant families. Polyploidy may explain the underlying causes of when classical taxonomy (classification before DNA sequences) is not enough and may potentially lead to underestimation of the true number of species. Due to the unique patterns across lineages, polyploidizations allows for no generalizations; despite its ubiquity, it remains mysterious. Polyploidy is, at least in part, reversible and leads to a smaller genome size over time.

Keywords: Diploidization, genome duplication, haplotypes, Hibisceae, *Hibiscus*, homoeologues, high-throughput sequencing, *Pavonia*, Malvaceae, polyploidization