A DISSERTATION FOR THE DEGREE OF MASTER OF SCIENCE

Impact of K-Pg Event on Evolution of Apocynaceae Inferred from 224 Plastomes

BY

JISEOK KIM

FEBRUARY, 2022

MAJOR IN CROP SCIENCE AND BIOTECHNOLOGY

DEPARTMENT OF AGRICULTURE, FORESTRY AND BIORESOURCES

THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

Impact of K-Pg Event on Evolution of Apocynaceae Inferred from 224 Plastomes

UNDER THE DIRECTION OF PROF. TAE-JIN YANG

SUBMITTED TO THE FACULTY OF
GRADUATE SCHOOL OF AGRICULTURE AND LIFE SCIENCES,
SEOUL NATIONAL UNIVERSITY

BY
JISEOK KIM

MAJOR IN CROP SCIENCE AND BIOTECHNOLOGY,
SEOUL NATIONAL UNIVERSITY

FEBRUARY, 2022

APPROVED AS A QUALIFIED DISSERTATION OF JISEOK KIM
FOR THE DEGREE OF MASTER OF SCIENCE

BY THE COMMITTEE MEMBERS

CHAIRMAN
Do-Soon Kim, Ph.D.

VICE-CHAIRMAN
Tae-Jin Yang, Ph.D.

MEMBER
Nam-Chon Paek, Ph.D.

ABSTRACT

Apocynaceae is considered a family that evolved successfully and well adapted to various environments due to its wide habitat and diverse morphological traits. Many pharmacological effects have been reported in several Apocynaceae plants. Considering its morphological diversity and successful adaptation to the global environment, phylogenetic analysis of Apocynaceae will give an insight into the evolution of plants and its connection with geological history. The evolutionary study will also be utilized to enhance the industrial potential of Apocynaceae. In this study, I studied 224 plastomes of Apocynaceae including sequence data available on the database and sequences produced from Korean domestic Apocynaceae plants. Phylogenetic analysis and divergence time estimation were conducted based on coding sequences of the whole plastomes. As a result, the first divergence time among three major subfamilies of Apocynaceae was calculated as 71.5 million years ago, which is approximately close to the K-Pg extinction event. It is investigated that subfamilies diverged after the K-Pg event have a tendency to include higher proportions of herbaceous species and non-polyploid species, which supports the impact of the event on the evolution of Apocynaceae. I also investigated unique structural variations and gene mutation of plastome which might reflect correlation with the divergence of an unique subfamily. A 4.6 kb segment inserted from mitochondrial DNA sequences and deletion of 3 amino acids in PsaI proteins are unique in Asclepiadoideae subfamily which have prevalent herbaceous species.

Key words: Apocynaceae; K-Pg extiction event; plastome; evolutionary study; phylogenetic analysis; divergence time estimation; *psaI*.

Student number: 2020-25881

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

K-Pg	Cretaceous–Paleogene
*dna*LCW	*De novo* assembly of low coverage WGS
nrDNA	Nuclear ribosomal DNA
MYA	Million years ago
MCMC	Monte Charlo Markov Chain
nrDNA	Nuclear ribosomal DNA
SNP	Single nucleotide polymorphism
LSC	Long sing copy
IR	Inverted repeat
SSC	Short single copy
ITS	Intergenic spacer
PtMt	plastid DNA derived from mitochondrial genome
MtPt	mitochondrial DNA derived from plastid genome
NAD	Nicotinamide adenine dinucleotide
ESS	Effective sample size
HPD	Highest posterior density

APPENDIX

Introduction

Apocynaceae consists of 378 genera and about 5,350 species and is distributed in tropical, dry, temperate and continental climate regions. It is considered a family that evolved successfully and well adapted to various environments due to its wide habitat and diverse morphological traits: woody climbers, vines, perennial herbs, trees and shrubs. Floral organs, fruits and seeds of Apocynaceae are anatomically complex, which resulted in multiple research elucidating the mechanism of its floral organ development and pollination. Most species of Rauvolfioideae are trees or shrubs, while Apocynoideae and Asclepiadoideae are predominantly herbaceous (Endress *et al.,* 2018).

Apocynaceae underwent several re-classifications due to its morphological diversity. After the first establishment of Apocineae (Apocynaceae) by Jussieu (1789), Asclepiadeae (Asclepiadaceae) was separated from Apocynaceae by Brown (1810a, 1810b). Later, Asclepiadaceae was also divided into two subfamilies, Asclepiadaceae and Periplocaceae, by Schlechter (1905). Endress *et al.* (2000) suggested merging these three families based on morphological evidence and molecular studies. During this re-classification, five subfamilies were recognized. Apocynoideae and Plumeroideae were transferred from the former Apocynaceae after separation by Brown, and Plumeroideae was renamed as Rauvolfioideae. Secamonoideae were transferred from Asclepiadaceae. Asclepiadaceae and Periplocaceae were downgraded to Asclepiadoideae and Periplocoideae, respectively. These changes have been accepted by the APG II system (THE ANGIOSPERM PHYLOGENY GROUP. 2003).

Several plants in Apocynaceae have been reported to have pharmacological effects. *Catharanthus roseus* has been used as an ingredient of vinblastine, an anticancer drug. Genomes of *Rhazya stricta*, *Calotropis gigantea*, and *Asclepias syriaca* are reported to study biosynthesis pathways of mono-indole alkaloids of Apocynaceae. Also, genus *Asclepias* is widely used for horticulture in North America, and genus *Cryptostegia* is utilized as a source of rubber.

Considering its morphological diversity and successful adaptation to the global environment, phylogenetic analysis of Apocynaceae will give an insight into the evolution of plants and its connection with geological history. The evolutionary study will also be utilized to enhance the industrial potential of Apocynaceae.

Multiple studies in which phylogenetic analyses were conducted based on plastomes of Apocynaceae plants have been reported. Fishbein *et al.* (2018) produced a maximum-likelihood phylogeny using plastomes of 103 species of *Asclepias*. Nazar *et al.* (2019) investigated phylogenetic relationships in Apocynaceae based on nuclear PHYA and plastid *trnL-F* sequences of 112 taxa. Plastomes are considered useful materials for estimating evolutionary history between plant species because they are well-conserved due to their high copy number and because of their uniparental inheritance. However, research analyzing phylogeny of Apocynaceae based on all available plastomes and investigating evolutionary meaning from genetic variations in a geological context has not been reported yet. In this study, I studied all plastomes of Apocynaceae available on the database as well as newly assembled plastomes of Korean domestic Apocynaceae plants, and conducted phylogenetic analysis and divergence time estimation. I also suggest evolutionary insights from structural variations of genomes and morphological differences between subfamilies, inferring probable geological events which have driven such divergence.

Materials and methods

Collection of plastomes

Eight Apocynaceae species samples were provided from Medicinal Herb Garden, College of Pharmacy, Seoul National University (Gyeonggi-do, Republic of Korea). A sample of *Trachelospermum asiaticum* was manually collected from Jeju-do, Republic of Korea (Table S2). Each sample was ground into fine powder submerged in liquid nitrogen, and DNA was extracted with Exgene Plant Mini Kit (GeneAll), following the manufacturer's instructions. Extracted DNA was sequenced on the Illumina MiSeq platform, and plastomes were assembled using *dna*LCW (de novo assembly of low coverage WGS) method (Kim *et al.,* 2015).

I also collected 215 plastomes of Apocynaceae species and four plastomes of outgroup on the NCBI GenBank database, regardless of whether they are complete or partial (Table S1). After excluding duplicates, 228 plastomes were collected and analyzed in this study in total.

Annotation of plastomes and inspection of gene sequences

Plastomes of nine Korean domestic Apocynaceae plants were annotated by GeSeq (Tillich *et al.,* 2017). All annotated genes of each plastome were manually inspected to determine the exact state of the genes. Normal genes were categorized as complete genes. Genes including undetermined nucleotides or overlapped with incomplete assembly regions were categorized as partial genes. Genes annotated as pseudogenes or containing frameshifts or truncations are categorized as pseudogenes. Mis-annotated genes with erroneous positions or reading frames were manually corrected.

Phylogenetic analysis

Out of 228 plastomes collected, 122 accessions belong to the genus *Asclepias*. *Asclepias* plastomes were screened to diminish redundancy of coding sequences and bias of sampling size. Seven plastomes of *Asclepias* accessions, introduced as exemplars of four major clades of *Asclepias* by Fishbein *et al.* (2018), were selected. Also, another complete *Asclepias* plastome not included in the dataset of Fishbein *et al.* (2018) was selected. 114 *Asclepias* plastomes other than the eight selected ones were excluded. Common coding sequences included in the remaining 114 accessions were extracted from annotation files, excluding genes annotated as pseudogenes in one or more accessions. Coding sequences extracted from each accession were concatenated in alphabetical order. The concatenated sequences were aligned using MAFFT (Katoh *et al.,* 2019), and conserved blocks were selected using Gblocks (Castresasna. 2000). A maximum-likelihood phylogenetic tree based on the curated sequences was constructed using RAxML (Stamatakis. 2014) with the GTRGAMMA model and 1,000 bootstraps.

Structural comparative analysis of complete plastomes

Out of 228 plastomes collected, 32 complete plastomes were selected. The selected plastomes were ordered following a topology suggested by the phylogenetic tree previously constructed, and the sequences of two adjacent plastomes were compared with each other using BLASTZ (Schwartz *et al.,* 2003).

Calculation and visualization of Ka/Ks values and selection of genes

Out of 228 plastomes collected, plastomes of three major ingroup subfamilies, Asclepiadoideae, Apocynoideae and Rauvolfioideae, were used in dN/dS calculation. One Secamonoideae accession, three Periplocoideae accessions, two polyphyletic Rauvolfioideae accessions and four outgroup accessions were excluded. All coding sequences were extracted from each accession unless the coding sequence is a pseudogene. The dN/dS calculation was conducted on every possible combination of two coding sequences of the same gene, from different accessions, using the CodeML package of PAML (Yang. 2007) with codon frequency F3X4.

The calculated Ka/Ks values were visualized using R (R Core Team. 2021). The values were first grouped into intra-subfamily and inter-subfamily values, and the two density plots were generated from the values of each group. The density plot generated from inter-subfamily values was divided into twelve overlays following major categories of chloroplast genes. Each overlay was divided into multiple overlays following genes. The inter-subfamily values were then divided into three subgroups according to which comparison pair of subfamilies the accessions used in the calculation belong.

Next, a boxplot was generated from all inter-subfamily Ka/Ks values and sorted by genes. Also, a grouped bar plot was generated with the mean of values according to genes and subgroups they belong. Then, the genes with possibilities of playing roles in divergence between subfamilies were selected judging from all generated plots.

Structural comparative analysis of genes

Coding sequences of the selected genes were converted into amino acid sequences. The amino acid sequences were aligned using MAFFT, following the same output order as the phylogenetic tree order. The aligned sequences were manually inspected, searching for significant structural changes accumulated across the divergence.

Divergence time estimation and evolutionary analysis

Divergence time estimation was conducted based on a set of curated sequences used in phylogenetic analysis, using BEAST 2 (Bouckaert *et al.,* 2014). Gamma site model, relaxed lognormal model and Yule model were set as site model, clock model and tree prior, respectively. The topology was set by packing each monophyletic crown into priors following the topology of the maximum-likelihood phylogenetic tree. Tree calibration priors were set on two points: 71.06 MYA for the divergence of Apocynaceae (Magallon *et al.,* 2015) and 47.8 MYA for the fossil evidence of *Asclepiadospermum* (Del Rio *et al.,* 2020). MCMC chain length was set to 20,000,000. The outputs of six independent runs were treated with 15% burn-in and combined using LogCombiner in the BEAST 2 package.

After the estimation, notable genetic variations observed in previous steps and morphological traits were bound to the MCMC tree. Genetic variations include structural variations in gene or genome level. Morphological traits were researched via literature studies. Finally, the integrated information was interpreted in a geological and evolutionary context.

Results

Genome assembly of nine Apocynaceae plants

Approximately 3.5 – 4.9 million reads were sequenced for collected nine Korean domestic Apocynaceae plants. The total amounts of NGS libraries are approximately 1.07 – 1.48 Gbps. Plastomes and 45S nrDNA sequences of the nine plants were assembled by *dna*LCW guideline (Kim *et al.,* 2015) (Table 1). The nine newly assembled plastomes have typical quadripartite architecture (Fig. 1a). Sequence lengths of plastomes are 153,532 – 161,968 bp. The lengths of LSC, IR and SSC are 85,367 – 90,553 bp, 23,876 – 25,765 bp and 17,438 – 23,663 bp, respectively.

The nine plastomes have 0 – 9091 SNPs and 4 – 908 InDels. Three plastomes of *T. asiaticum* collections have 0 – 5 SNPs and 3 – 4 InDels between each other. *Vincetoxicum paniculatum* and two *Cynanchum* species are relatively more similar than their outgroups, having 168 – 764 SNPs and 90 – 110 InDels (Table 2), and this result matches well with a previous study that *Cynanchum* and *Vincentoxicum* were once hard to discriminate, but lately recognized as different taxa (Liede, 1996).

The nine newly assembled 45S nrDNA also have a typical structure with three subunits and two intergenic spacers (Fig. 1b). Sequence lengths of 45S nrDNA sequences range from 5,809 bp to 5,852 bp. The lengths of 18S subunit, ITS1, 5.8S subunit, ITS2 and 28S subunit are 1,807 – 1,809 bp, 227 – 243 bp, 160 – 162 bp, 220 – 250 bp and 3,390 – 3,395 bp, respectively.

Two phylogenetic trees were constructed with plastomes and 45S nrDNA sequences, respectively, and the topology of both trees are identical (Fig. 1c).

Table 1. Raw read data and newly assembled plastomes and 45S rDNA sequences of Korean domestic Apocynaceae plants.

Subfamily	Name	Raw read data		Plastome				45S nrDNA					
		Total reads	**Total bases (bp)**	**Total length (bp)**	**LSC (bp)**	**IR (bp)**	**SSC (bp)**	**Total length (bp)**	**18S (bp)**	**ITS1 (bp)**	**5.8S (bp)**	**ITS2 (bp)**	**28S (bp)**
Asclepiadoideae	*Vincetoxicum paniculatum*	3,705,022	1,115,211,622	159,533	90,360	24,919	19,335	5,834	1,807	227	162	243	3,395
Asclepiadoideae	*Cynanchum nipponicum*	3,956,276	1,190,839,076	159,126	90,341	24,791	19,203	5,834	1,807	228	162	242	3,395
Asclepiadoideae	*Cynanchum ascyrifolium*	3,957,072	1,191,078,672	161,968	90,553	23,876	23,663	5,835	1,807	228	162	243	3,395
Apocynoideae	*Trachelospermum asiaticum* -1	4,913,032	1,478,822,632	155,536	85,831	25,755	18,195	5,809	1,808	229	160	220	3,392
Apocynoideae	*Trachelospermum asiaticum* -2	4,053,492	1,220,101,092	155,529	85,804	25,765	18,195	5,809	1,808	229	160	220	3,392
Apocynoideae	*Trachelospermum asiaticum* -3	4,162,844	1,253,016,044	155,496	85,793	25,754	18,195	5,809	1,808	229	160	220	3,392
Apocynoideae	*Nerium oleander* -1	4,138,288	1,245,624,688	154,860	85,473	25,626	18,135	5,825	1,807	232	160	234	3,392
Rauvolfioideae	*Amsonia elliptica*	4,480,608	1,348,663,008	154,225	85,367	25,710	17,438	5,833	1,808	231	160	242	3,392
Rauvolfioideae	*Vinca major* -1	3,544,728	1,066,963,128	153,523	85,557	25,124	17,718	5,852	1,809	243	160	250	3,390

Table 2. The number of SNPs and InDels between newly assembled plastomes.

	Vincetoxicum paniculatum	*Cynanchum nipponicum*	*Cynanchum ascyrifolium*	*Trachelospermum asiaticum* -1	*Trachelospermum asiaticum* -2	*Trachelospermum asiaticum* -3	*Nerium oleander* -1	*Amsonia elliptica*	*Vinca major* -1
Vincetoxicum paniculatum		90	93	701	701	702	695	874	955
Cynanchum nipponicum	168		110	703	703	702	697	877	955
Cynanchum ascyrifolium	237	764		741	741	741	732	911	980
Trachelospermum asiaticum -1	6202	6225	6308		4	3	410	658	716
Trachelospermum asiaticum -2	6204	6227	6310	0		4	411	658	714
Trachelospermum asiaticum -3	6208	6250	6333	1	5		410	659	716
Nerium oleander -1	6120	6079	6183	2925	2917	2928		632	684
Amsonia elliptica	7896	7865	7942	4999	4995	5008	4224		714
Vinca major -1	9019	9041	9091	6177	6184	6182	5451	5031	

Note: The lower triangle shows the total nucleotide substitutions, while the upper triangle indicates the number of InDels.

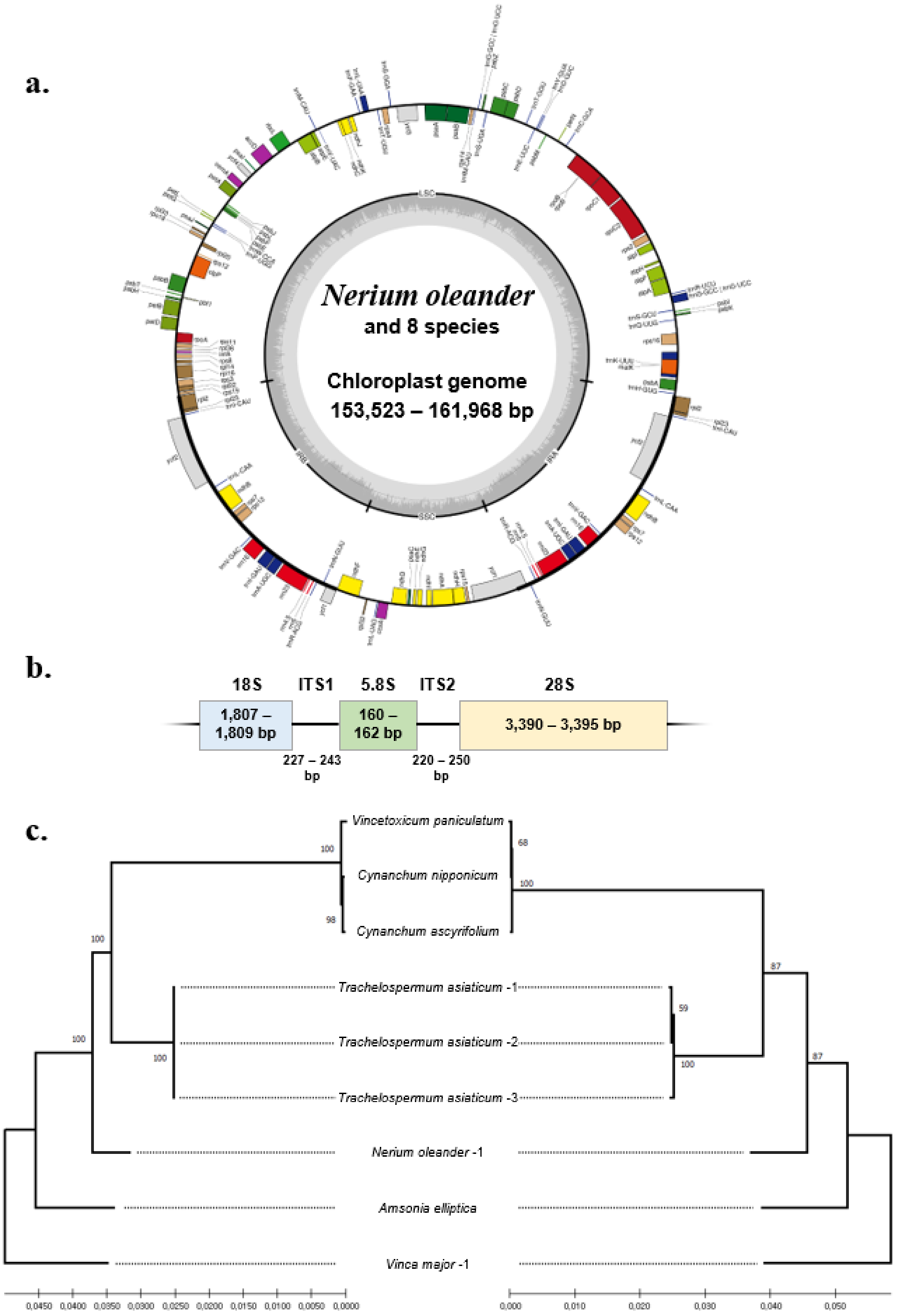

Figure 1. Newly assembled plastomes and 45S nrDNA sequences and their phylogeny. (a) Plastome of nine Apocynaceae plants. (b) 45S nrDNA sequences of nine Apocynaceae plants. (c) Phylogenetic trees of nine plastomes (left) and 45S nrDNA sequences (right).

Retrieval and functional annotation of coding sequences of collected plastomes

Out of 18,012 possible coding sequences of 79 genes and 228 plastomes collected, 361 coding sequences were erroneous (Fig. 2). Out of the 361 coding sequences, 212 coding sequences contained undetermined nucleotides (N) or were truncated with unfinished assembly. Another 147 coding sequences were pseudogenes which contains early termination or critical truncation. The other two coding sequences were completely not found in the plastomes.

Four gene structures were annotated as pseudogenes. Many *Asclepias* species have pseudogenized *clpP*. *Hoya liangii*, *Hoya pottsii*, *Eustegia minuta* and *Pachypodium baronii* have pseudogenized *ndh* genes. Also, two *Hoya* species have pseudogenized *accD* and *ycf* genes.

Phylogenetic analysis

A maximum-likelihood phylogenetic tree based on coding sequences was constructed (Fig. 3). The three major subfamilies, Rauvolfioideae, Apocynoideae and Asclepiadoideae, are finely grouped. *Secamone afzelii*, the only collected accession in Secamonoideae, is located right before the divergence of Asclepiadoideae. However, Periplocoideae stands inside Apocynoideae, and two accessions of Rauvolfioideae, *Prestonia portobellensis* and *Amphineurion marginatum*, are also located within Apocynoideae. Overall, the topology of the phylogenetic tree follows the conventional taxonomical system, but some species, Periplocoideae species and two Rauvolfioideae species, are positioned in different groups.

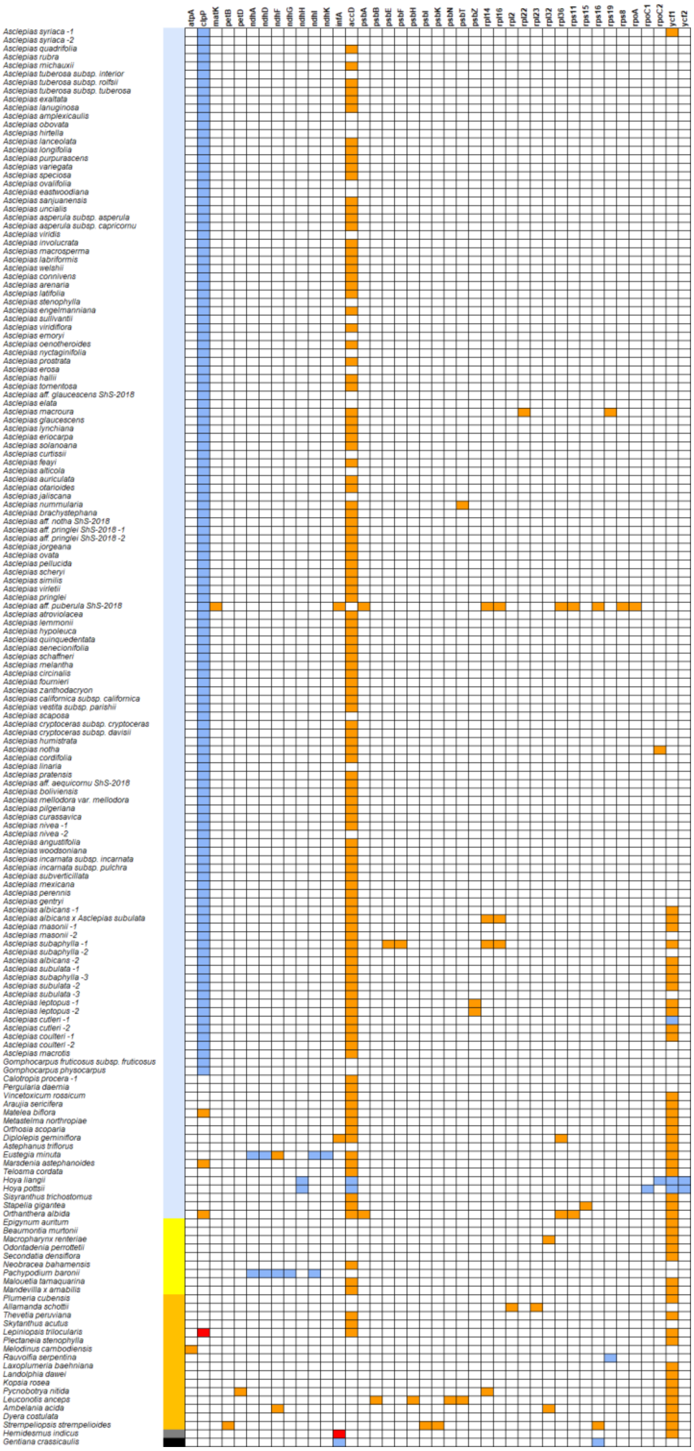

Figure 2. The integrity of coding sequences of collected plastomes. Each cell represents the state of coding sequence of the corresponding gene and accession. (white: normal; orange: partial sequence; blue: pseudogene; red: non-existent) Columns and rows with no erroneous cells are omitted. Subfamilies are denoted as blue, yellow and orange colors for Asclepiadoideae, Apocynoideae and Rauvolfioideae, respectively, on the right side of each species.

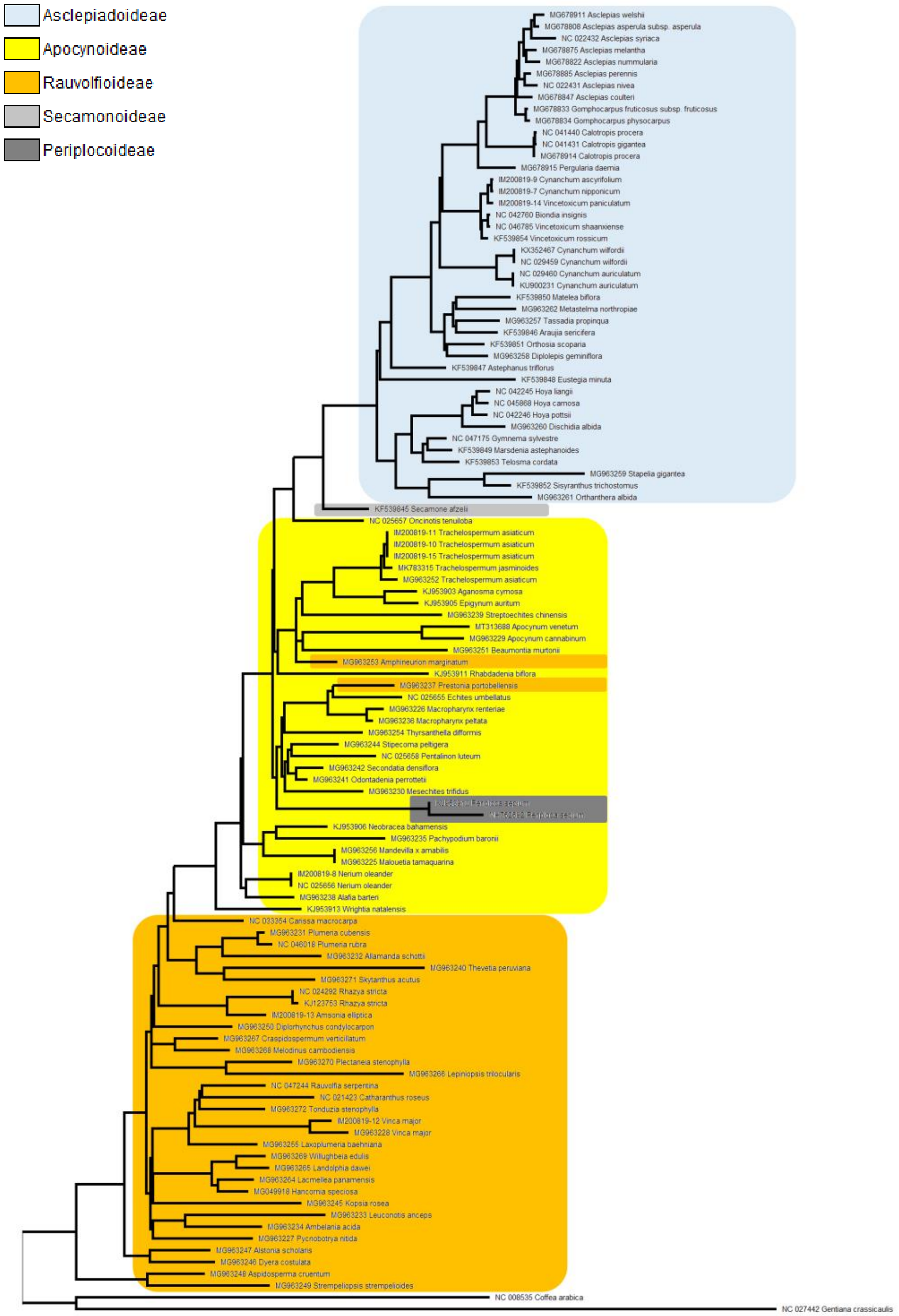

Figure 3. A maximum-likelihood phylogenetic tree based on plastomes of Apocynaceae plants. Subfamilies in which each species belongs are denoted as blue, yellow and orange colors for Asclepiadoideae, Apocynoideae and Rauvolfioideae, respectively. Two Rauvolfioideae species and subfamily Periplocoideae are positioned within Apocynoideae cluster.

Subfamily-unique structural evolution of plastomes

Comparative analysis of complete plastomes based on a phylogenetic tree order do not show dramatic structural variation. However, a plastome size increase was identified in *H. carnosa*, which turned out to be the IR expansion due to partial duplication of SSC regions (Rodda *et al.,* 2021). Also, a large insertion is discovered at a point when Asclepiadoideae diverged from Apocynoideae (Fig. 4). This insertion is approximately 4.6 kb in length and is reported as an intracellular gene transfer region from mitochondrial DNA sequences (plastid DNA derived from mitochondrial genome: PtMt) (Straub *et al.,* 2013). The inserted PtMt is well conserved in most Asclepiadoideae species except for shrinkage of the region of *A. syriaca*. Therefore, it is assumed that the variation occurred when the Asclepiadoideae subfamily diverged from two other subfamilies.

Calculation and visualization of Ka/Ks values and selection of genes

The synonymous mutation rate and non-synonymous mutation rate of each gene were calculated, and the results were visualized into density plots to search positively selected genes. The results of the intra-subfamily comparison and inter-subfamily comparison were indicated in blue and yellow polygons, respectively (Fig. 5a). Judging from the intra-subfamily mutation rate is higher than the inter-subfamily mutation rate, I considered the calculated values reliable. When inter-subfamily values were divided by categories based on the function of genes, values of photosystem and NADH dehydrogenase categories, which involve in photosynthesis and respiration, are relatively low, and values of ribosomal protein, *clpP* and *matK* categories are relatively high (Fig. 5b). Among genes in the photosystem I category, in which most of the members have low variations, *psaI* shows extraordinarily high variations (Fig. 5c). Also, among genes in the ribosomal protein short subunit category, *rps7* shows clear and concentrated Ka/Ks ratios (Fig. 5d). Thus, these two genes are assumed to be positively selected in the divergence progress of Asclepiadoideae and the two other subfamilies. This tendency is also observed on the boxplot and grouped bar plot. On the boxplot, *psaI*, *rps7*, *clpP* and *psbT* have their quartiles exceeding 1 (Fig. 6). On the grouped bar plot, only *psaI* and *rps7* exceed 1 (Fig. 7).

Structural comparative analysis of genes

Structures of *psaI* and *rps7* were observed by comparing the sequences of all accessions. In *psaI*, a deletion event of 9 bp length and substitution event of four amino acids are detected between Asclepiadoideae and the other two subfamilies (Fig. 8a). The deletion event occurred on the region next to the start codon and did not cause a frameshift. Since this region is not homologous with the mitochondrial genome, it is assumed that intracellular gene transfer events do not have a correlation with the deletion event. Also, since there are no repeat regions neighboring this region, and the neighboring intergenic regions are well conserved (Fig. 9), the deletion is not likely to have occurred by slipped strand mispairing.

The coding sequence of *rps7* is well conserved, but Asclepiadoideae has relatively more amino acid substitutions while the other two subfamilies are similar to each other (Fig. 8b). Rps7 of Asclepiadoideae has 4 – 7 subsutitutions compared with Apocynoideae and Rauvolfioideae. There are two substitutions in Rps7 between Apocynoideae and Rauvolfioideae.

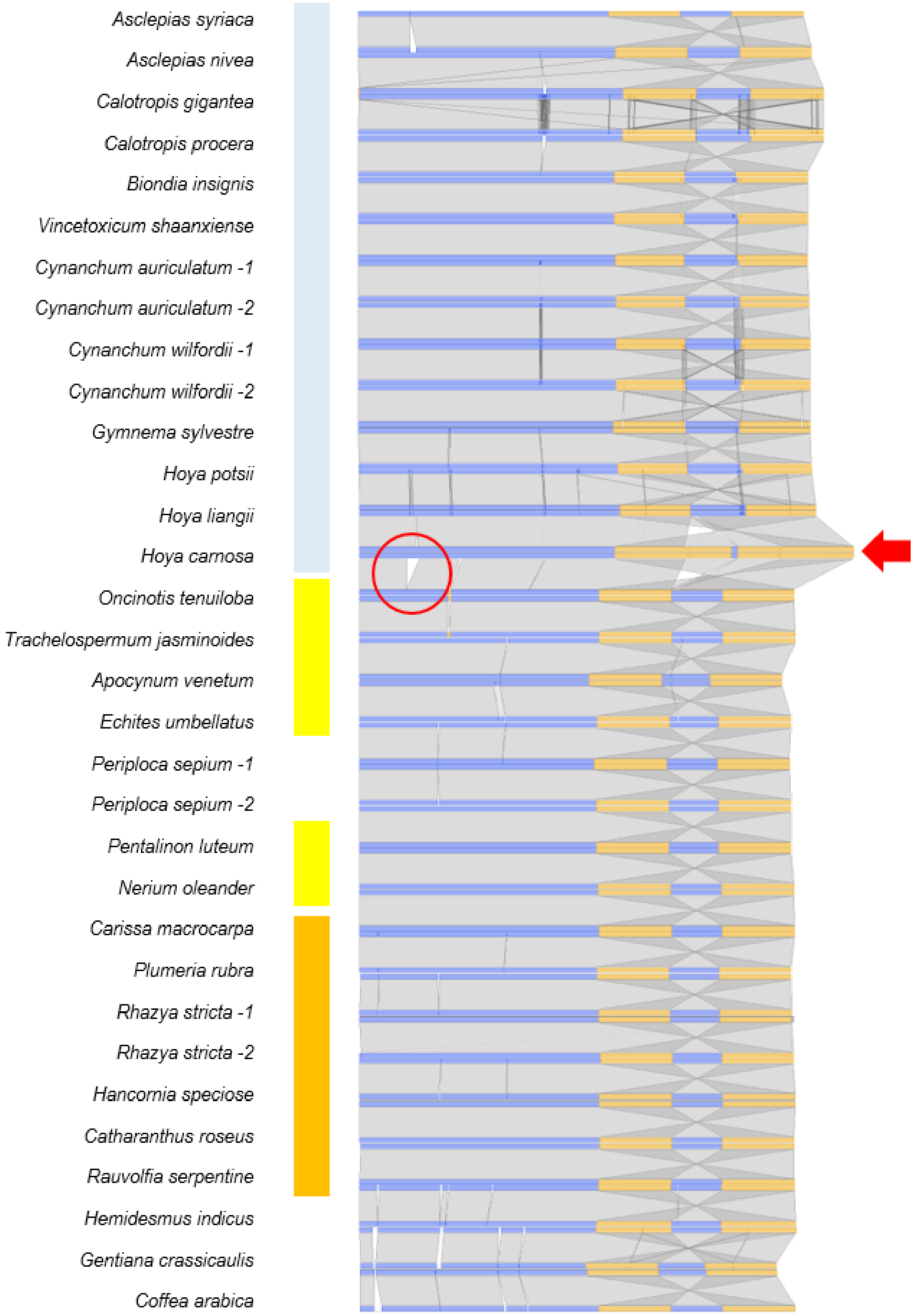

Figure 4. Synteny visualization among 30 complete plastomes in Apocynaceae and two outgroup species. The red circle indicates approximately 4.6 kbp insertion caused by intracellular gene transfer event. The red arrow mark indicates IR expansion of *H. carnosa*. Subfamily names are denoted as different color boxes, blue, yellow, and orange for Asclepiadoideae, Apocynoideae, and Rauvolfioideae, respectively.

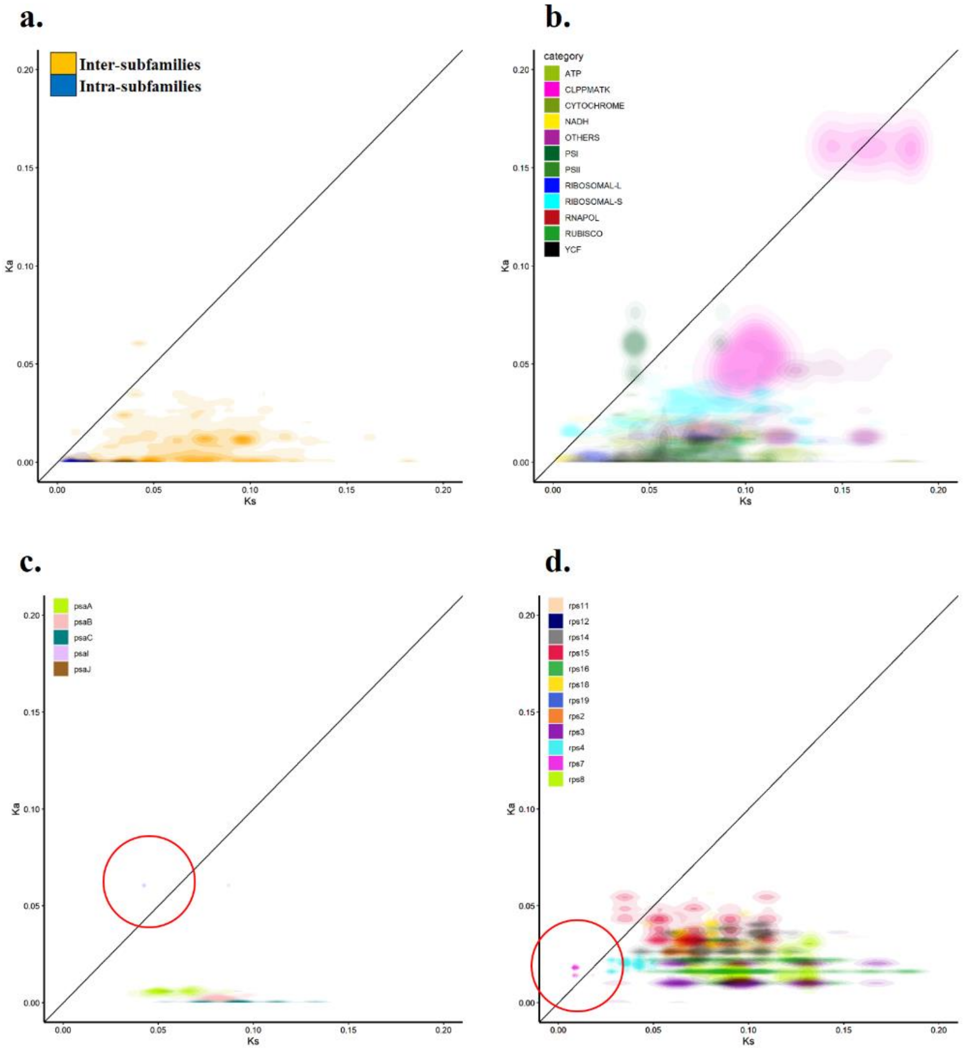

Figure 5. Density plots of Ka/Ks values generated by dN/dS calculation. (a) Overlayed density plots of inter-subfamily and intra-subfamily Ka/Ks values. (b) Overlayed density plots of Ka/Ks values in twelve gene categories. (c) Overlayed density plots of Ka/Ks values in photosystem I gene category. The red circle indicates the values of *psaI*. (d) Overlayed density plots of Ka/Ks values in ribosomal proteins – short subunits gene category. The red circle indicates the values of *rps7*.

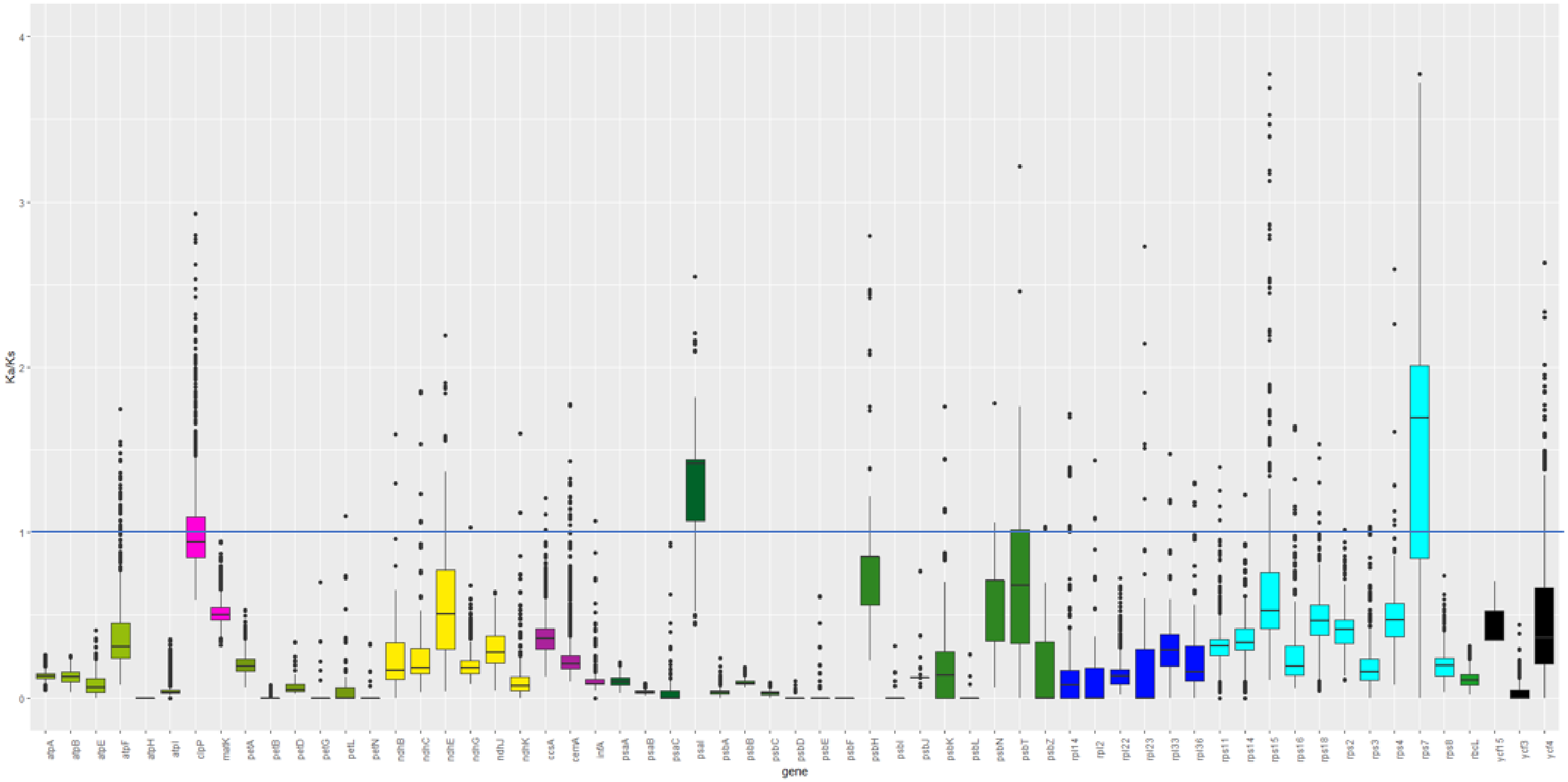

Figure 6. A boxplot of Ka/Ks values generated by dN/dS calculation.

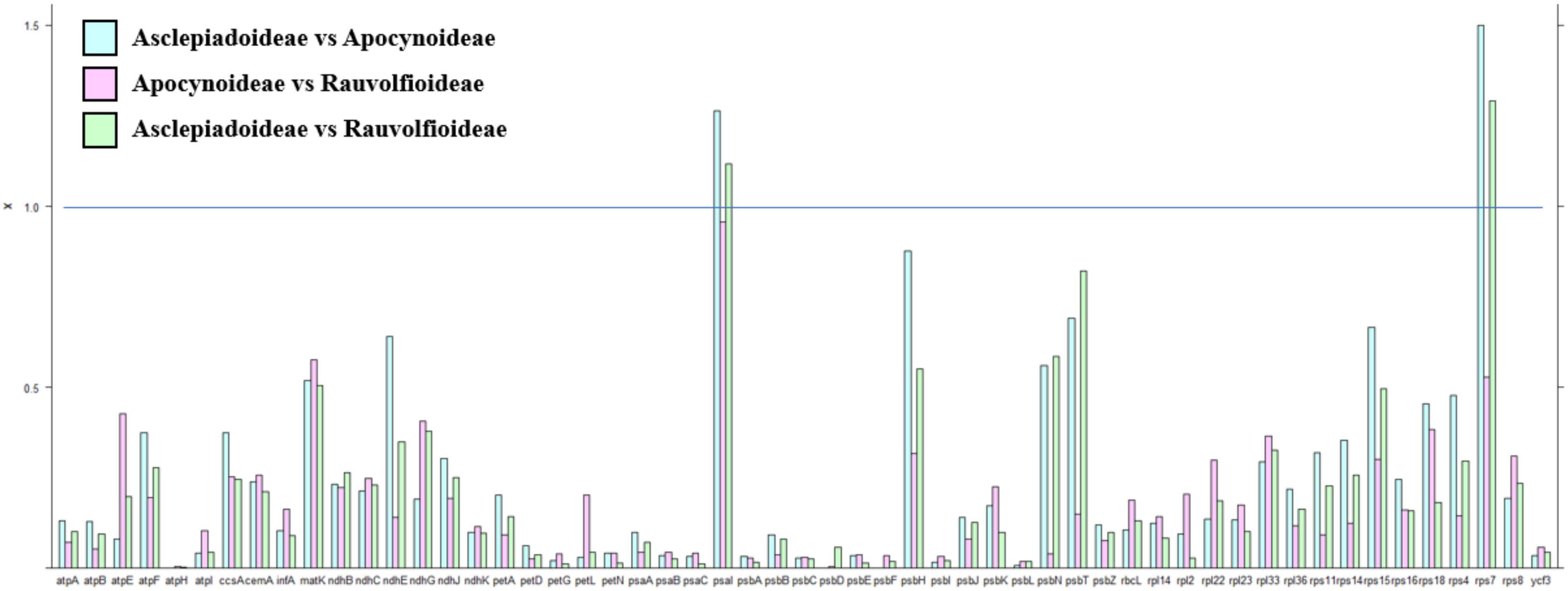

Figure 7. A grouped bar plot of Ka/Ks values generated by dN/dS calculation.

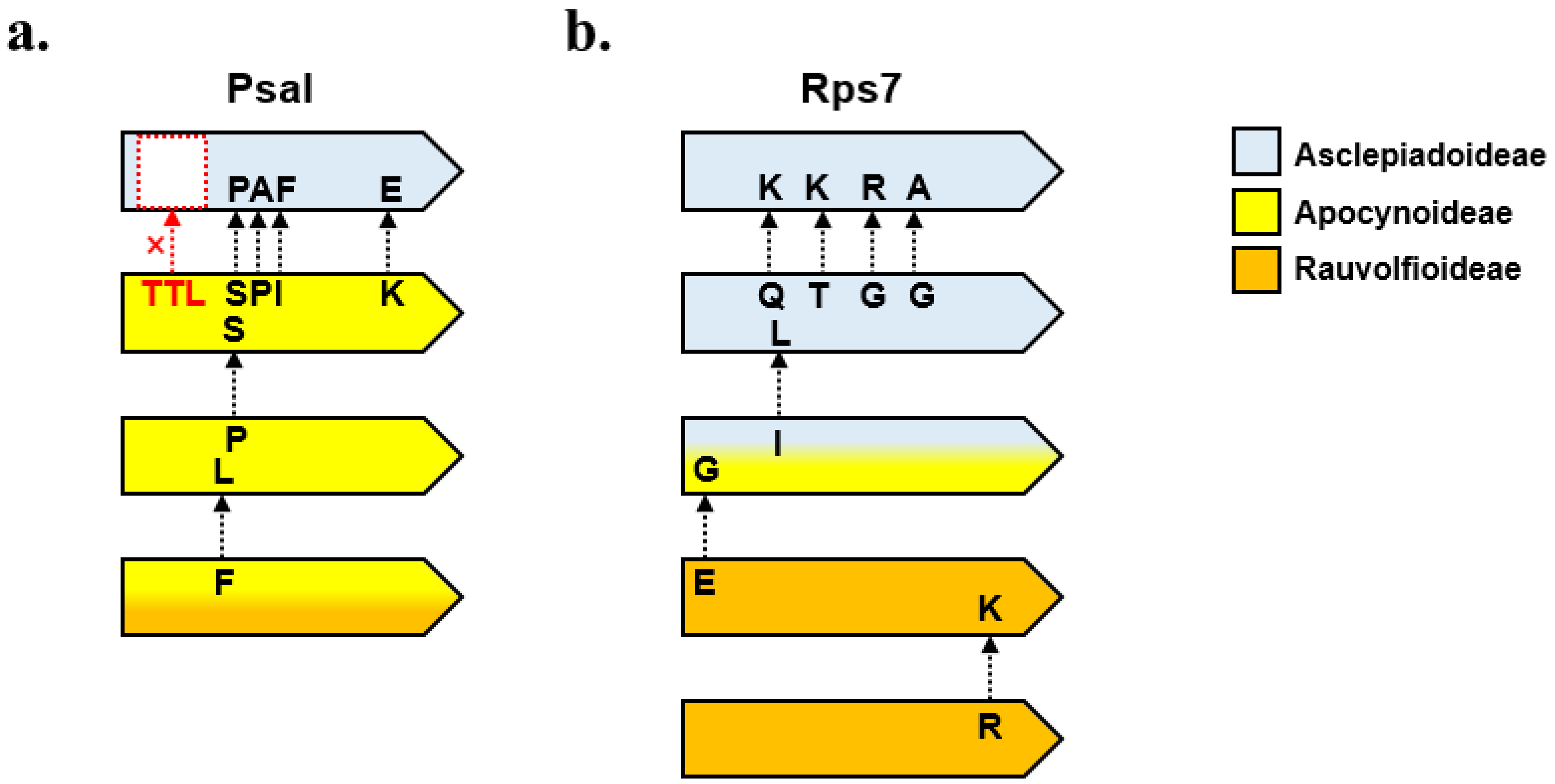

Figure 8. Structural variations of amino acid sequences of PsaI and Rps7 accumulated through divergence of subfamilies. Subfamilies were denoted as different colors in each model boxes, blue, yellow and orange for Asclepiadoideae, Apocynoideae and Rauvolfioideae, respectively.

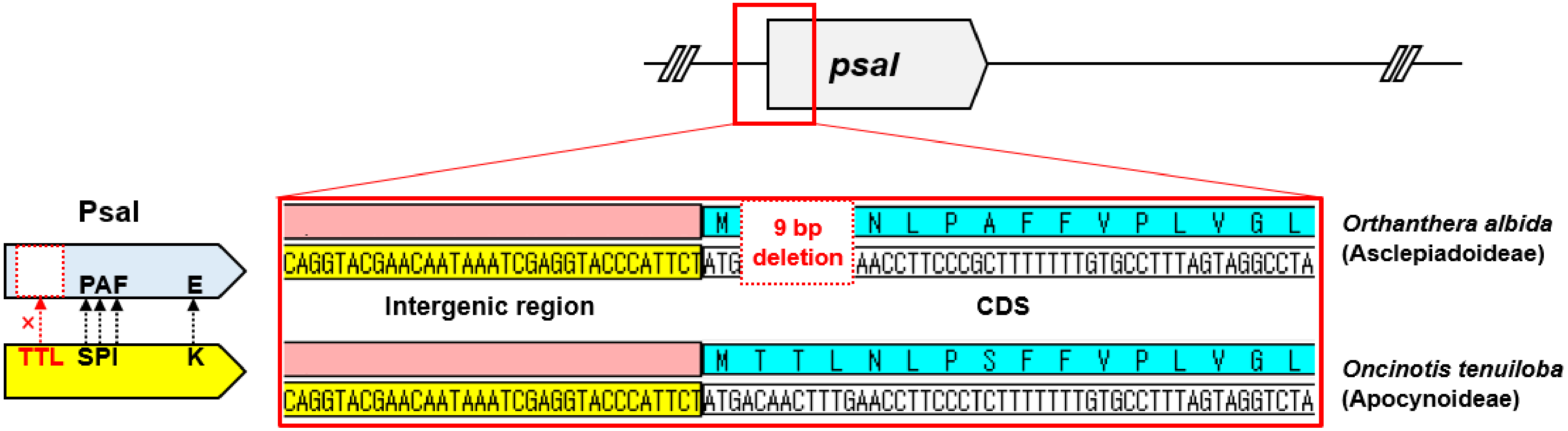

Figure 9. Nucleotide and amino acid sequences of *Orthanthera albida* and *Oncinotis tenuiloba* near *psaI* start codon where 9-bp deletion event occured. Subfamilies were denoted as different colors in PsaI model boxes, blue and yellow for Asclepiadoideae and the others, respectively.

Divergence time estimation and evolutionary analysis

As a result of divergence time estimation using coding sequences of 228 accessions, including an outgroup of four accessions, Rauvolfioideae and Asclepiadoideae are estimated to have diverged from its most recent common ancestor 71.5 and 33.1 MYA, respectively (Fig. 10). The ESS values of major parameters are over 100. The time points approximately accord to Cretaceous–Paleogene (K-Pg) extinction event and Eocene–Oligocene (E-O_G) extinction event.

As a result of investigating karyotype and morphological traits of these accessions by literature studies, herbaceous and woody species are prevalent in Asclepiadoideae and Rauvolfioideae, respectively, and Apocynoideae does not show a one-sided tendency. Also, most Asclepiadoideae species have 11 pairs of chromosomes, while Apocynoideae and Rauvolfioideae species have various numbers of chromosomes. Furthermore, Rauvolfioideae included some polyploid species.

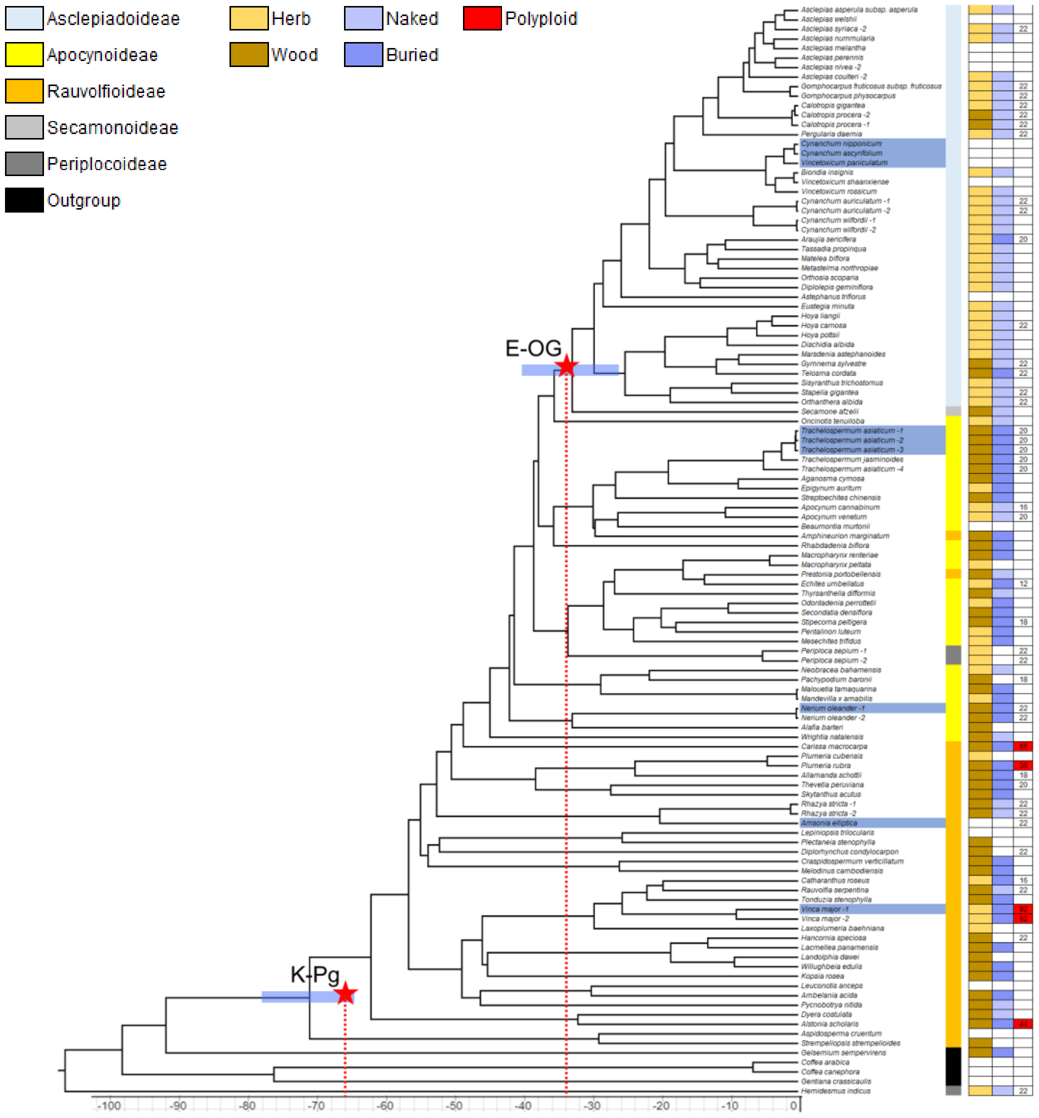

Figure 10. A MCMC tree of Apocynaceae with divergence time estimation. The blue bars overlapping nodes indicate 95% HPD of heights. The two red stars indicate the K-Pg event and the E-O$_{\text{G}}$ event, respectively. Subfamily names in which each species belongs are denoted as blue, yellow, and orange boxes for Asclepiadoideae, Apocynoideae and Rauvolfioideae, respectively, on right side of each species. Two morphological traits, woodiness and flower type, and ploidy levels are indicated as different color boxes in each block, yellow, blue and red, respectively, on the right. The yellow, dark yellow and white boxes indicate herb, woody and unknown plant types, respectively. The blue, dark blue and white boxes indicate naked, buried and unknown flowering types, respectively. The chromosome numbers are denoted in boxes and polyploids are denoted as red boxes.

Discussion

Diverse pseudogenization events in several species

In this study, *clpP* shows high level of polymorphisms including substitutions, elongation and insertion-deletion, especially on exon 3. Pseudogenization of *clpP* detected in most *Asclepias* species is reported to be caused by 28-bp insertion near the end of exon 3 (Straub *et al.,* 2011). However, additional studies are required to validate whether the 28-bp insertion caused the pseudogenization of the *clpP* gene since it is inferred with sequence variations but was not proved for actual function. For example, *clpP* of *Asclepias albicans* (JN710458) is annotated as pseudogene although it does not have the 28-bp insertion. On the other hand, *clpP* of *Gomphocarpus fruticosus* subsp. *fruticosus* (MG678833) is not annotated as pseudogene although it has the 28-bp insertion.

Pseudogenization of genes in the NADH dehydrogenase category are identified in two *Hoya* species, tropical epiphyte, and *E. minuta* and *P. baronii*, native to dry regions (Fig. 2). Those species chronically inhabit the environment with insufficient water supply. Pseudogenization of *accD*, *ycf1* and genes in the NADH dehydrogenase category is a general feature for plants with desiccation tolerance (Xu *et al.* 2018). Thus, the pseudogenization of genes in the NADH dehydrogenase category found in some Apocynaceae species is thought to be individual events that occurred through speciation progress for adaptation to a diverse environment such as surface of stems of woods or dry regions with limited water supply.

Taxonomical mispositioning of three species

As a result of maximum-likelihood phylogenetic analysis with 228 plastomes, most species of Apocynaceae are classified in accordance with the conventional taxonomical system. However, three species are positioned in clades which differ with the original subfamily group (Fig. 3). Two Rauvolfioideae species, *A. marginatum* and *P. portobellensis*, are positioned within Apocynoideae group in this study. Also, another species belonged in Periplocoideae subfamily was grouped inside Apocynoideae submaily. Since these mismatches were also observed in other studies (Del Rio *et al.* 2020), the taxonomical mis-positioning should be considered and examined. The reason for these mismatches could be misclassification by traditional plant classification methods or limitation of plastomes due to their uniparental inheritance. Further studies will be necessary to clarify the taxonomical position of these plant species.

The estimated divergence time of Apocynaceae and geological interpretation

As a result of divergence time estimation by MCMC tree analysis, the divergence time of Apocynaceae was estimated to be 71.5 MYA, and a subfamily, Asclepiadoideae, is estimated to diverge from the others 33.1 MYA (Fig. 10). Many discussions about the divergence time of Apocynaceae are still underway (Magallon *et al.,* 2015; Neupane *et al.,* 2017; Del Rio *et al.,* 2020). However, judging from the analysis method that combined *in silico* plastome data with physical fossil evidence achieved fine parameter scores, it is reasonable to estimate the divergence time of Apocynaceae to 71.5 MYA.

During the K-Pg event, a mass extinction happened in a brief period when a massive meteorite collided on the Yucatan Peninsula in the Gulf of Mexico, creating a Chixulub crater with a diameter of more than 150 km (Hildebrant *et al.* 1991). As a result, more than 70% of all animal and plant species on Earth have been extinct, including dinosaurs (Jablonski, 1994). The botanical phase of the forest, composed of coniferous trees, was turned into a broad-leaved forest during this time (Carvalho *et al.* 2021), and shrubs and herbaceous plants thrived considerably (Levin *et al.* 2018).

Also, it is known that many whole-genome duplication events are concentrated close to the K-Pg boundary (Levin *et al.* 2018). Therefore, it is assumed that subfamilies that diverged around the K-Pg event, including Rauvolfioideae, underwent more whole-genome duplication events compared to other subfamilies. This hypothesis corresponds to the result of literature studies for karyotype, in which most polyploid species belong to Rauvolfioideae.

Genomic trace of Asclepiadoideae evolution

Asclepiadoidae is estimated to have diverged 33.1 MYA when a drastic temperature decrease occurred at the end of the Eocene, which is known as the last major extinction event (Ivany *et al.* 2000). The warm climate of the Paleocene, when the surface temperature of the polar region was close to 15 °C, abruptly cooled to a level similar to that of the Holocene, current geological epoch (Bruke *et al.* 2018.), and tropical rainforests, which were once found all over the world, shrank to equatorial regions, resulting in rapid environmental changes (Buerki *et al.* 2013).

All of the plastomes belonged Asclepiadoideae species have 9-bp deletion near the 5' end of *psaI* gene (Fig. 8a). Although little is known about the function of PsaI, it is known that PsaI is responsible for the stability of photosystem I under high light and low temperature conditions (Schöttler *et al.* 2017). It is assumed that rapid leaf senescence, caused by incomplete function of the PsaI, during the low temperature period of late Eocene caused an advantageous effect on dormancy and survival by senescence mediated by mutation of the *psaI* gene. I assume that the functional change of the *psaI* gene imposed a positive selection pressure on the divergence of Asclepiadoideae in which herbaceous species are dominant.

Also, in comparison among the plastome sequences of Apocynaceae, there was a large insertion of mitochondrial genome fragment (plastid DNA derived from mitochondrial genome: PtMt) into the *rps2 – rpoC2* intergenic region of Asclepiadoideae (Fig. 4). This insertion event is thought to have happened during the subfamily divergence phase. The PtMt is 4.6 kbp in length and common in most of species in the Apocynaceae subfamily. However, considering that the insertion size is 2.5 kbp in the case of *A. syriaca*, the size is thought to be variable by subsequential deletion of this additional PtMt segment.

A total of six amino acid substitutions were found in the *rps7* gene of 224 plastomes. Among the six amino acid changes, four were unique and common in the Asclepiadoideae species. The mutation might be associated with divergence of the Asclepiadoideae subfamily with two other subfamilies. However, more research is required to prove how and what these variations affect for the functional evolution of this subfamily.

Conclusion

The phylogenetic relationships of the Apocynaceae were re-established using huge genome data including 224 plastomes in Apocynaceae family. I carefully calculated their divergence time based on sequence variations and fossil evidence. I also tried to identify unique structure modification and gene mutation in plastomes which are possibly involved in subfamily emergence.

Since introduction of next generation sequencing, massive WGS data have become publicly available. Also, as genome assembly methodologies including *dna*LCW advance, plastome sequence data are being reported in large quantity. Now, it is necessary to consider how this massive data would be used in an efficient way. In this study, I draw a novel conclusion by integrating previously reported public plastome sequence data, which is difficult to achieve with single data. Likewise, the use of plastome data in macroscopic perspective is necessary.

This study could be valuable to understand evolution of the diverse plant species in Apocynaceae family which adapted in global climates and environments.

References

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut and A. J. Drummond (2014). "BEAST 2: a software platform for Bayesian evolutionary analysis." PLoS computational biology **10**(4): e1003537.

Brown, R. (1810a). Prodromus Florae Novae Hollandiae et Insulae van Diemen. London. Reprinted in C. G. D. Nees yon Esenbeck (ed., 1827), Robert Brown's vermischte botanische Schriften 3, issued 1828 as a separate work, Nuremberg.

Brown, R. (1810b). On the Asclepiadeae, a natural order of plants separated from the Apocineae of Jussieu. [Preprint of Mem. Were. Nat. Hist. Soc. 1:12-78 (1811)].

Buerki, S., F. Forest, T. Stadler and N. Alvarez (2013). "The abrupt climate change at the Eocene–Oligocene boundary and the emergence of South-East Asia triggered the spread of sapindaceous lineages." Annals of Botany **112**(1): 151-160.

Burke, K. D., J. W. Williams, M. A. Chandler, A. M. Haywood, D. J. Lunt and B. L. Otto-Bliesner (2018). "Pliocene and Eocene provide best analogs for near-future climates." Proceedings of the National Academy of Sciences **115**(52): 13288-13293.

Carvalho, M. R., C. Jaramillo, F. de la Parra, D. Caballero-Rodríguez, F. Herrera, S. Wing, B. L. Turner, C. D'Apolito, M. Romero-Báez and P. Narváez (2021). "Extinction at the end-Cretaceous and the origin of modern Neotropical rainforests." Science **372**(6537): 63-68.

Castresana, J. (2000). "Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis." Molecular biology and evolution **17**(4): 540-552.

Del Rio, C., T. X. Wang, J. Liu, S. Q. Liang, R. A. Spicer, F. X. Wu, Z. K. Zhou and T. Su (2020). "Asclepiadospermum gen. nov., the earliest fossil record of Asclepiadoideae (Apocynaceae) from the early Eocene of central Qinghai-Tibetan Plateau, and its biogeographic implications." American Journal of Botany **107**(1): 126-138.

Endress, M. E. and P. V. Bruyns (2000). "A revised classification of the Apocynaceae sl." The Botanical Review **66**(1): 1-56.

Endress, M. E., U. Meve, D. J. Middleton and S. Liede-Schumann (2018). Apocynaceae. Flowering Plants. Eudicots, Springer**:** 207-411.

Fishbein, M., S. C. Straub, J. Boutte, K. Hansen, R. C. Cronn and A. Liston (2018). "Evolution at the tips: Asclepias phylogenomics and new perspectives on leaf surfaces." American Journal of Botany **105**(3): 514-524.

Group, A. P. (2009). "An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III." Botanical Journal of the Linnean Society **161**(2): 105-121.

Hildebrand, A. R., G. T. Penfield, D. A. Kring, M. Pilkington, A. Camargo Z, S. B. Jacobsen and W. V. Boynton (1991). "Chicxulub crater: a possible Cretaceous/Tertiary boundary impact crater on the Yucatan Peninsula, Mexico." Geology **19**(9): 867-871.

Ivany, L. C., W. P. Patterson and K. C. Lohmann (2000). "Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary." Nature **407**(6806): 887-890.

Jablonski, D. (1994). "Extinctions in the fossil record." Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **344**(1307): 11-17.

Jussieu, A. L. de. (1789). Genera Plantarum. Herissant, Paris.

Katoh, K., J. Rozewicki and K. D. Yamada (2019). "MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization." Briefings in bioinformatics **20**(4): 1160-1166.

Kim, K., S.-C. Lee, J. Lee, Y. Yu, K. Yang, B.-S. Choi, H.-J. Koh, N. E. Waminal, H.-I. Choi and N.-H. Kim (2015). "Complete chloroplast and ribosomal sequences for 30 accessions elucidate evolution of Oryza AA genome species." Scientific reports **5**(1): 1-13.

Levin, D. A. and D. E. Soltis (2018). "Factors promoting polyploid persistence and diversification and limiting diploid speciation during the K–Pg interlude." Current opinion in plant biology **42**: 1-7.

Liede, S. (1996). "Cynanchum. Rhodostegiella. Vincetoxicum. Tylophora (Asclepiadaceae): New Considerations on an Old Problem." Taxon: 193-211.

Magallón, S., S. Gómez-Acevedo, L. L. Sánchez-Reyes and T. Hernández-Hernández (2015). "A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity." New Phytologist **207**(2): 437-453.

Nazar, N., J. J. Clarkson, D. Goyder, E. Kaky, T. Mahmood and M. W. Chase (2019). "Phylogenetic relationships in Apocynaceae based on nuclear PHYA and plastid trnL-F sequences, with a focus on tribal relationships." Caryologia **72**(1): 55-81.

Neupane, S., P. O. Lewis, S. Dessein, H. Shanks, S. Paudyal and F. Lens (2017). "Evolution of woody life form on tropical mountains in the tribe Spermacoceae (Rubiaceae)." American Journal of Botany **104**(3): 419-438.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.Rproject.org/

Rodda, M. and M. A. Niissalo (2021). "Plastome evolution and organisation in the Hoya group (Apocynaceae)." Scientific Reports **11**(1): 1-13.

Schlechter, F. R. R. (1905). Periplocaceae & Asclepiadaceae. Pp. 351-369 in K. Schumann & K. Lauterbach (eds.), Nachtraege zur Flora des Deutschen Suedseegebiets. Borntraeger, Leipzig.

Schöttler, M. A., W. Thiele, K. Belkius, S. V. Bergner, C. Flügel, G. Wittenberg, S. Agrawal, S. Stegemann, S. Ruf and R. Bock (2017). "The plastid-encoded PsaI subunit stabilizes photosystem I during leaf senescence in tobacco." Journal of experimental botany **68**(5): 1137-1155.

Schwartz, S., W. J. Kent, A. Smit, Z. Zhang, R. Baertsch, R. C. Hardison, D. Haussler and W. Miller (2003). "Human–mouse alignments with BLASTZ." Genome research **13**(1): 103-107.

Stamatakis, A. (2014). "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies." Bioinformatics **30**(9): 1312-1313.

Straub, S. C., M. Fishbein, T. Livshultz, Z. Foster, M. Parks, K. Weitemier, R. C. Cronn and A. Liston (2011). "Building a model: developing genomic resources for common milkweed (Asclepias syriaca) with low coverage genome sequencing." BMC genomics **12**(1): 1-22.

Straub, S. C., R. C. Cronn, C. Edwards, M. Fishbein and A. Liston (2013). "Horizontal transfer of DNA from the mitochondrial to the plastid genome and its subsequent evolution in milkweeds (Apocynaceae)." Genome biology and evolution **5**(10): 1872-1885.

THE ANGIOSPERM PHYLOGENY GROUP. (2003). "An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II." Botanical Journal of the Linnean Society **141**(4): 399-436.

Tillich, M., P. Lehwark, T. Pellizzer, E. S. Ulbricht-Jones, A. Fischer, R. Bock and S. Greiner (2017). "GeSeq–versatile and accurate annotation of organelle genomes." Nucleic acids research **45**(W1): W6-W11.

Xu, Z., T. Xin, D. Bartels, Y. Li, W. Gu, H. Yao, S. Liu, H. Yu, X. Pu and J. Zhou (2018). "Genome analysis of the ancient tracheophyte Selaginella tamariscina reveals evolutionary features relevant to the acquisition of desiccation tolerance." Molecular plant **11**(7): 983-994.

Yang, Z. (2007). "PAML 4: phylogenetic analysis by maximum likelihood." Molecular biology and evolution **24**(8): 1586-1591.

Supplementary data

Table S1. The list of 219 plastomes collected from NCBI genbank database.

Subfamily	Name	Accession no.
Asclepiadoideae	*Asclepias syriaca* -1	JF433943
Asclepiadoideae	*Asclepias syriaca* -2	NC_022432
Asclepiadoideae	*Asclepias quadrifolia*	MG678890
Asclepiadoideae	*Asclepias rubra*	MG678827
Asclepiadoideae	*Asclepias michauxii*	MG678878
Asclepiadoideae	*Asclepias tuberosa* subsp. *interior*	MG678831
Asclepiadoideae	*Asclepias tuberosa* subsp. *rolfsii*	MG678904
Asclepiadoideae	*Asclepias tuberosa* subsp. *tuberosa*	MG678905
Asclepiadoideae	*Asclepias exaltata*	MG678854
Asclepiadoideae	*Asclepias lanuginosa*	MG678818
Asclepiadoideae	*Asclepias amplexicaulis*	MG678807
Asclepiadoideae	*Asclepias obovata*	MG678824
Asclepiadoideae	*Asclepias hirtella*	MG678815
Asclepiadoideae	*Asclepias lanceolata*	MG678817
Asclepiadoideae	*Asclepias longifolia*	MG678820
Asclepiadoideae	*Asclepias purpurascens*	MG678889
Asclepiadoideae	*Asclepias variegata*	MG678907
Asclepiadoideae	*Asclepias speciosa*	MG678899
Asclepiadoideae	*Asclepias ovalifolia*	MG678825
Asclepiadoideae	*Asclepias eastwoodiana*	MG678811
Asclepiadoideae	*Asclepias sanjuanensis*	MG678893
Asclepiadoideae	*Asclepias uncialis*	MG678906
Asclepiadoideae	*Asclepias asperula* subsp. *asperula*	MG678808
Asclepiadoideae	*Asclepias asperula* subsp. *capricornu*	MG678840
Asclepiadoideae	*Asclepias viridis*	MG678832
Asclepiadoideae	*Asclepias involucrata*	MG678865
Asclepiadoideae	*Asclepias macrosperma*	MG678872
Asclepiadoideae	*Asclepias labriformis*	MG678867
Asclepiadoideae	*Asclepias welshii*	MG678911
Asclepiadoideae	*Asclepias connivens*	MG678809
Asclepiadoideae	*Asclepias arenaria*	MG678839
Asclepiadoideae	*Asclepias latifolia*	MG678868
Asclepiadoideae	*Asclepias stenophylla*	MG678829
Asclepiadoideae	*Asclepias engelmanniana*	MG678852
Asclepiadoideae	*Asclepias sullivantii*	MG678830
Asclepiadoideae	*Asclepias viridiflora*	MG678909
Asclepiadoideae	*Asclepias emoryi*	MG678813
Asclepiadoideae	*Asclepias oenotheroides*	MG678881
Asclepiadoideae	*Asclepias nyctaginifolia*	MG678823
Asclepiadoideae	*Asclepias prostrata*	MG678888
Asclepiadoideae	*Asclepias erosa*	MG678814
Asclepiadoideae	*Asclepias hallii*	MG678860
Asclepiadoideae	*Asclepias tomentosa*	MG678903
Asclepiadoideae	*Asclepias aff. glaucescens* ShS-2018	MG678805
Asclepiadoideae	*Asclepias elata*	MG678812
Asclepiadoideae	*Asclepias macroura*	MG678821
Asclepiadoideae	*Asclepias glaucescens*	MG678859
Asclepiadoideae	*Asclepias lynchiana*	MG678871
Asclepiadoideae	*Asclepias eriocarpa*	MG678853
Asclepiadoideae	*Asclepias solanoana*	MG678897
Asclepiadoideae	*Asclepias curtissii*	MG678810
Asclepiadoideae	*Asclepias feayi*	MG678855
Asclepiadoideae	*Asclepias alticola*	MG678806

Asclepiadoideae	*Asclepias auriculata*	MG678842
Asclepiadoideae	*Asclepias otarioides*	MG678882
Asclepiadoideae	*Asclepias jaliscana*	MG678816
Asclepiadoideae	*Asclepias nummularia*	MG678822
Asclepiadoideae	*Asclepias brachystephana*	MG678844
Asclepiadoideae	*Asclepias aff. notha* ShS-2018	MG678898
Asclepiadoideae	*Asclepias aff. pringlei* ShS-2018 -1	MG678836
Asclepiadoideae	*Asclepias aff. pringlei* ShS-2018 -2	MG678837
Asclepiadoideae	*Asclepias jorgeana*	MG678866
Asclepiadoideae	*Asclepias ovata*	MG678883
Asclepiadoideae	*Asclepias pellucida*	MG678884
Asclepiadoideae	*Asclepias scheryi*	MG678895
Asclepiadoideae	*Asclepias similis*	MG678896
Asclepiadoideae	*Asclepias virletii*	MG678910
Asclepiadoideae	*Asclepias pringlei*	MG678887
Asclepiadoideae	*Asclepias aff. puberula* ShS-2018	MG678826
Asclepiadoideae	*Asclepias atroviolacea*	MG678841
Asclepiadoideae	*Asclepias lemmonii*	MG678869
Asclepiadoideae	*Asclepias hypoleuca*	MG678862
Asclepiadoideae	*Asclepias quinquedentata*	MG678891
Asclepiadoideae	*Asclepias senecionifolia*	MG678892
Asclepiadoideae	*Asclepias schaffneri*	MG678894
Asclepiadoideae	*Asclepias melantha*	MG678875
Asclepiadoideae	*Asclepias circinalis*	MG678846
Asclepiadoideae	*Asclepias fournieri*	MG678857
Asclepiadoideae	*Asclepias zanthodacryon*	MG678913
Asclepiadoideae	*Asclepias californica* subsp. *californica*	MG678845
Asclepiadoideae	*Asclepias vestita* subsp. *parishii*	MG678908
Asclepiadoideae	*Asclepias scaposa*	MG678828
Asclepiadoideae	*Asclepias cryptoceras* subsp. *cryptoceras*	MG678849
Asclepiadoideae	*Asclepias cryptoceras* subsp. *davisii*	MG678850
Asclepiadoideae	*Asclepias humistrata*	MG678861
Asclepiadoideae	*Asclepias notha*	MG678880
Asclepiadoideae	*Asclepias cordifolia*	MG678848
Asclepiadoideae	*Asclepias linaria*	MG678819
Asclepiadoideae	*Asclepias pratensis*	MG678886
Asclepiadoideae	*Asclepias aff. aequicornu* ShS-2018	MG678835
Asclepiadoideae	*Asclepias boliviensis*	MG678843
Asclepiadoideae	*Asclepias mellodora* var. *mellodora*	MG678876
Asclepiadoideae	*Asclepias pilgeriana*	MG678856
Asclepiadoideae	*Asclepias curassavica*	MG678851
Asclepiadoideae	*Asclepias nivea* -1	MG678879
Asclepiadoideae	*Asclepias nivea* -2	NC_022431
Asclepiadoideae	*Asclepias angustifolia*	MG678838
Asclepiadoideae	*Asclepias woodsoniana*	MG678912
Asclepiadoideae	*Asclepias incarnata* subsp. *incarnata*	MG678863
Asclepiadoideae	*Asclepias incarnata* subsp. *pulchra*	MG678864
Asclepiadoideae	*Asclepias subverticillata*	MG678902
Asclepiadoideae	*Asclepias mexicana*	MG678877
Asclepiadoideae	*Asclepias perennis*	MG678885
Asclepiadoideae	*Asclepias gentryi*	MG678858
Asclepiadoideae	*Asclepias albicans* -1	JN710457
Asclepiadoideae	*Asclepias albicans* -2	JN710458
Asclepiadoideae	*Asclepias albicans* x *Asclepias subulata*	JN710470
Asclepiadoideae	*Asclepias masonii* -1	JN710465
Asclepiadoideae	*Asclepias masonii* -2	MG678874
Asclepiadoideae	*Asclepias subaphylla* -1	JN710467
Asclepiadoideae	*Asclepias subaphylla* -2	MG678900
Asclepiadoideae	*Asclepias subaphylla* -3	JN710466
Asclepiadoideae	*Asclepias subulata* -1	JN710468

Asclepiadoideae	*Asclepias subulata* -2	JN710469
Asclepiadoideae	*Asclepias subulata* -3	MG678901
Asclepiadoideae	*Asclepias leptopus* -1	JN710462
Asclepiadoideae	*Asclepias leptopus* -2	MG678870
Asclepiadoideae	*Asclepias cutleri* -1	JN710460
Asclepiadoideae	*Asclepias cutleri* -2	JN710461
Asclepiadoideae	*Asclepias coulteri* -1	JN710459
Asclepiadoideae	*Asclepias coulteri* -2	MG678847
Asclepiadoideae	*Asclepias macrotis*	MG678873
Asclepiadoideae	*Gomphocarpus fruticosus* subsp. *fruticosus*	MG678833
Asclepiadoideae	*Gomphocarpus physocarpus*	MG678834
Asclepiadoideae	*Calotropis procera* -1	MG678914
Asclepiadoideae	*Calotropis procera* -2	NC_041440
Asclepiadoideae	*Calotropis gigantea*	NC_041431
Asclepiadoideae	*Pergularia daemia*	MG678915
Asclepiadoideae	*Vincetoxicum rossicum*	KF539854
Asclepiadoideae	*Biondia insignis*	NC_042760
Asclepiadoideae	*Vincetoxicum shaanxiense*	NC_046785
Asclepiadoideae	*Cynanchum auriculatum* -1	KU900231
Asclepiadoideae	*Cynanchum auriculatum* -2	NC_029460
Asclepiadoideae	*Cynanchum wilfordii* -1	KX352467
Asclepiadoideae	*Cynanchum wilfordii* -2	NC_029459
Asclepiadoideae	*Araujia sericifera*	KF539846
Asclepiadoideae	*Tassadia propinqua*	MG963257
Asclepiadoideae	*Matelea biflora*	KF539850
Asclepiadoideae	*Metastelma northropiae*	MG963262
Asclepiadoideae	*Orthosia scoparia*	KF539851
Asclepiadoideae	*Diplolepis geminiflora*	MG963258
Asclepiadoideae	*Astephanus triflorus*	KF539847
Asclepiadoideae	*Eustegia minuta*	KF539848
Asclepiadoideae	*Marsdenia astephanoides*	KF539849
Asclepiadoideae	*Gymnema sylvestre*	NC_047175
Asclepiadoideae	*Telosma cordata*	KF539853
Asclepiadoideae	*Dischidia albida*	MG963260
Asclepiadoideae	*Hoya liangii*	NC_042245
Asclepiadoideae	*Hoya carnosa*	NC_045868
Asclepiadoideae	*Hoya pottsii*	NC_042246
Asclepiadoideae	*Sisyranthus trichostomus*	KF539852
Asclepiadoideae	*Stapelia gigantea*	MG963259
Asclepiadoideae	*Orthanthera albida*	MG963261
Apocynoideae	*Oncinotis tenuiloba*	NC_025657
Apocynoideae	*Trachelospermum asiaticum* -4	MG963252
Apocynoideae	*Trachelospermum jasminoides*	MK783315
Apocynoideae	*Aganosma cymosa*	KJ953903
Apocynoideae	*Epigynum auritum*	KJ953905
Apocynoideae	*Streptoechites chinensis*	MG963239
Apocynoideae	*Apocynum cannabinum*	MG963229
Apocynoideae	*Apocynum venetum*	MT313688
Apocynoideae	*Beaumontia murtonii*	MG963251
Apocynoideae	*Rhabdadenia biflora*	KJ953911
Apocynoideae	*Echites umbellatus*	NC_025655
Apocynoideae	*Macropharynx renteriae*	MG963226
Apocynoideae	*Macropharynx peltata*	MG963236
Apocynoideae	*Thyrsanthella difformis*	MG963254
Apocynoideae	*Mesechites trifidus*	MG963230
Apocynoideae	*Odontadenia perrottetii*	MG963241
Apocynoideae	*Secondatia densiflora*	MG963242
Apocynoideae	*Stipecoma peltigera*	MG963244
Apocynoideae	*Pentalinon luteum*	NC_025658
Apocynoideae	*Nerium oleander* -2	NC_025656

Apocynoideae	*Alafia barteri*	MG963238
Apocynoideae	*Neobracea bahamensis*	KJ953906
Apocynoideae	*Pachypodium baronii*	MG963235
Apocynoideae	*Malouetia tamaquarina*	MG963225
Apocynoideae	*Mandevilla* x *amabilis*	MG963256
Apocynoideae	*Wrightia natalensis*	KJ953913
Rauvolfioideae	*Amphineurion marginatum*	MG963253
Rauvolfioideae	*Prestonia portobellensis*	MG963237
Rauvolfioideae	*Carissa macrocarpa*	NC_033354
Rauvolfioideae	*Plumeria cubensis*	MG963231
Rauvolfioideae	*Plumeria rubra*	NC_046018
Rauvolfioideae	*Allamanda schottii*	MG963232
Rauvolfioideae	*Thevetia peruviana*	MG963240
Rauvolfioideae	*Skytanthus acutus*	MG963271
Rauvolfioideae	*Rhazya stricta* -1	KJ123753
Rauvolfioideae	*Rhazya stricta* -2	NC_024292
Rauvolfioideae	*Diplorhynchus condylocarpon*	MG963250
Rauvolfioideae	*Lepiniopsis trilocularis*	MG963266
Rauvolfioideae	*Plectaneia stenophylla*	MG963270
Rauvolfioideae	*Craspidospermum verticillatum*	MG963267
Rauvolfioideae	*Melodinus cambodiensis*	MG963268
Rauvolfioideae	*Vinca major* -2	MG963228
Rauvolfioideae	*Tonduzia stenophylla*	MG963272
Rauvolfioideae	*Catharanthus roseus*	NC_021423
Rauvolfioideae	*Rauvolfia serpentina*	NC_047244
Rauvolfioideae	*Laxoplumeria baehniana*	MG963255
Rauvolfioideae	*Hancornia speciosa*	MG049918
Rauvolfioideae	*Lacmellea panamensis*	MG963264
Rauvolfioideae	*Landolphia dawei*	MG963265
Rauvolfioideae	*Willughbeia edulis*	MG963269
Rauvolfioideae	*Kopsia rosea*	MG963245
Rauvolfioideae	*Pycnobotrya nitida*	MG963227
Rauvolfioideae	*Leuconotis anceps*	MG963233
Rauvolfioideae	*Ambelania acida*	MG963234
Rauvolfioideae	*Dyera costulata*	MG963246
Rauvolfioideae	*Alstonia scholaris*	MG963247
Rauvolfioideae	*Aspidosperma cruentum*	MG963248
Rauvolfioideae	*Strempeliopsis strempelioides*	MG963249
Secamonoideae	*Secamone afzelii*	KF539845
Periplocoideae	*Periploca sepium* -1	KJ953910
Periplocoideae	*Periploca sepium* -2	MH752592
Periplocoideae	*Hemidesmus indicus*	NC_047471
(Gelsemiaceae)	*Gelsemium sempervirens*	MG963263
(Rubiaceae)	*Coffea arabica*	NC_008535
(Rubiaceae)	*Coffea canephora*	NC_030053
(Gentianaceae)	*Gentiana crassicaulis*	NC_027442

Note. The name of subfamily of an accession not included in Apocynaceae is indicated as the name of its family in parentheses. The duplicate scientific names are distinguished with numbers.

Table S2. The list of nine Korean domestic plants of Apocynaceae.

Subfamily	Name	Source	Accession no.
Asclepiadoideae	*Vincetoxicum paniculatum*	Medicinal Herb Garden	IM200819-14
Asclepiadoideae	*Cynanchum nipponicum*	Medicinal Herb Garden	IM200819-7
Asclepiadoideae	*Cynanchum ascyrifolium*	Medicinal Herb Garden	IM200819-9
Apocynoideae	*Trachelospermum asiaticum* -1	Medicinal Herb Garden	IM200819-10
Apocynoideae	*Trachelospermum asiaticum* -2	Medicinal Herb Garden	IM200819-11
Apocynoideae	*Trachelospermum asiaticum* -3	Jeju-do, Korea	IM200819-15
Apocynoideae	*Nerium oleander* -1	Medicinal Herb Garden	IM200819-8
Rauvolfioideae	*Amsonia elliptica*	Medicinal Herb Garden	IM200819-13
Rauvolfioideae	*Vinca major* -1	Medicinal Herb Garden	IM200819-12

Note. The duplicate scientific names are distinguished with numbers.

ABSTRACT IN KOREAN

협죽도과는 넓은 서식지와 다양한 형태학적 형질을 가지고 있으며, 성공적으로 진화하여 다양한 환경에 적응한 과로 평가된다. 일부 협죽도과 식물에서는 다양한 약리학적 효능이 보고된다. 형태학적 다양성과 전 세계 환경에 성공적으로 적응한 점을 고려할 때, 협죽도과의 계통학적 분석은 식물의 진화 및 지질학적 연혁과의 연관성에 대한 새로운 이해를 제공할 수 있다. 또한 협죽도과에 대한 진화 분석은 협죽도과의 산업적인 잠재력을 향상시키는 데에도 기여할 것이다. 본 연구에서는 데이터베이스에 공개된 염기서열과 대한민국에 자생하는 식물에서 확보한 염기서열을 포함하여 총 224 개 협죽도과 식물의 엽록체 유전체를 분석하였으며 각 유전체에서 추출한 coding sequence를 기반으로 계통학적 분석 및 분화 시기 추정을 수행하였다. 그 결과 협죽도과의 3 개 주요 아과 중 첫 분화 시기는 약 7150 만 년 전으로 계산되었으며, 이는 백악기-팔레오기 대량절멸 사건이 일어난 시기와 근접하다. 이 시기 이후에 분화한 아과는 초본인 종의 비율 및 배수체가 아닌 종의 비율이 더 높은 것으로 조사되었는데, 이는 해당 대량절멸 사건이 협죽도과의 진화에 영향을 가했음을 시사한다. 또한 특정 아과의 분화와 연관성이 있을 것으로 추정되는 엽록체 유전체의 고유한 구조적 변이와 유전자 변형을 발견하였다. 미토콘드리아 서열로부터 삽입된 4.6 kb의 DNA 서열 조각과 PsaI 단백질의 3개 아미노산 삭제는 초본이 우세한 박주가리아과만이 보이는 특징이다.

핵심어: 협죽도과, 백악기-팔레오기 대량절멸, 엽록체 유전체, 진화 분석, 계통학적 분석, 분화 시기 추정, *psaI*.

ACKNOWLEDGEMENT

이 지면을 빌려 석사과정 2년, 혹은 그 이상의 시간 동안 제가 입은 은혜에 대하여 감사함을 표하고자 합니다.

먼저 스스로의 가치를 깨닫지 못하고 있던 저에게 망설임 없이 손을 내밀고 아낌없는 지원과 조건 없는 사랑을 베풀어주신 지도교수 양태진 교수님께 진심으로 감사의 말씀을 드립니다. 부족하던 저를 무한히 신뢰하시고, 허물은 덮어주시고 장점을 부각시키면서, 날카로운 통찰력으로 끊임없는 가르침을 반복하시어, 저는 자신감을 찾고 재능있는 분야를 찾아 원없이 몰두한 끝에 제 2의 인생을 시작할 수 있었습니다.

시간이 흘러도 한결같은 마음으로 저를 지도해주시고 연구자로써의 기초소양과 더불어 어른으로써의 기본 마음가짐을 지도해주신 작물생명과학전공의 이변우 교수님, 고희종 교수님, 이석하 교수님, 김도순 교수님, 백남천 교수님, 서학수 교수님, 김광수 교수님께도 감사의 인사를 드립니다.

2년 동안 저의 집이었던 기능성식물연구실의 구성원 분들께도 감사의 말씀을 드립니다. 제가 연구에 몰두할 수 있도록 물심양면으로 지원을 아끼지 않으셨던 박지영 박사님, 비가 오나 눈이 오나 수원을 지키시는 유홍섭 박사님, 항상 젊은 마인드로 친절한 형님이 되어주셨던 박현승 박사님, 겉으로 드러나지 않게 후배를 배려하고 지원해주신 장우종 박사님, 진심으로 사람을 대하고 저의 첫 외국인 친구가 되어주신 파푸 박사님, 먼 타지에 와서 타국의 사람들을 도와주신 공 박사님, 자신의 몫

까지 내어주며 지도해주신 강종수 박사님, 언제나 품위 있게 후배를 지원해주신 조호준 선배님, 모르는 것을 물어보면 언제나 친절하고 자세히 알려주신 린 선배님, 연구실 구성원들의 또 하나의 일터인 서버를 책임감 있게 관리하신 구현진 선배님, 몇 날 몇 일 밤을 같이 지새워주던 저의 사수 이세현 선배님, 뛰어난 능력으로 대학원생의 모범을 보여주신 심현아 선배님, 유일한 동갑으로 저를 많이 신경써주신 박영상 선배님, 신중하고 사려 깊은 일처리를 알려주신 조우현 선배님, 2년간 같은 방에 살며 친구 이상의 사이가 된 김진태 선배님, 힘든 일도 기쁜 일도 처음부터 함께한 저의 동기 은비와 연정이, 부족한 저를 높여주고 항상 많은 도움을 준 용혁, 선희, 민녕이, 거친일을 마다하지 않으시며 농장에 진심어린 애정을 쏟으시는 박세원 선생님과 최화춘 선생님. 한 분 한 분께 모두 진심을 담아 감사의 말씀을 전합니다.

마지막으로, 적지 않은 나이에 학위 과정에 도전한 저를 믿고, 일생을 바쳐 저를 지원해주신 사랑하는 부모님께 감사의 말씀을 전합니다.

저는 저에게 새로운 인생을 베풀어준 농업생명과학대학의 졸업생임을 잊지 않고 사회에 진출해서도 항상 그 은혜를 갚기 위해 노력하겠습니다. 감사합니다.