Review and Progress Open Access

Phylogenetic Position of Durian and Comparative Genomics with Close Relatives in Malvaceae

Wang May [1,2]
1 Hainan Provincial Institute of Biological Engineering, Haikou, 570206, Hainan, China
2 Hainan Institute of Tropical Agricultural Resources, Sanya, 572025, Hainan, China
Corresponding author: haimei.wang@hibio.org

Genomics and Applied Biology, 2025, Vol.16, No.1 doi: 10.5376/gab.2025.16.0004
Received: 20 Dec., 2024
Accepted: 26 Jan., 2025
Published: 21 Feb., 2025

Preferred citation for this article:
Wang M., 2025, Phylogenetic position of durian and comparative genomics with close relatives in malvaceae, Genomics and Applied Biology, 16(1): 34-45 (doi: 10.5376/gab.2025.16.0004)

Abstract *Durio zibethinus Murr*. is a famous tropical fruit tree of the Mallow family. It is known as the "king of fruits" for its unique and strong odor and high economic value. This study reviews the phylogenetic location of durian in the Malvaceae family, and a comparative analysis of related species of durian and Malvaceae family based on the latest genomics research. Through comparative genomic analysis, it was found that the genomic collinearity of durian and Malvaceae remained conservative in genomic collinearity, but there were also chromosomal rearrangements and gene families. We also compared the evolution of key functional genes of durian and its relative species: the ethylene signal and secondary metabolic pathway genes related to durian fruit maturation showed unique expression and evolutionary patterns, and the disease-resistant and stress-responsive genes also had copy number variation and selection pressure differences among different species. Case analysis shows that the genomes of durian and cocoa (Theobroma cacao) are similar in overall structure but different in scale, and durian and kapok (*Bombax Ceiba*) are phylogenetically closer, sharing some conserved gene clusters. The end of this study looks forward to the application prospects of durian genome research in variety improvement. Based on the research progress in the past five years, we have deepened our understanding of the genome evolution of durian and its related species, which will provide a scientific basis for the genetic improvement of durian and the development and utilization of plant resources in the Malvaceae family.

Keywords Durian; Malvaceae; Comparative genomics; Whole genome replication; Secondary metabolism

1 Introduction

Durian belongs to the durian plant of the Mallowe family. It is a famous fruit with great economic value in Southeast Asia and is known as the "king of fruits" for its sweet and rich flesh and unique and strong odor. Durian fruits are nutritious and expensive. They are widely cultivated in Thailand, Malaysia, Indonesia and other countries, and are an important local cash crop. In recent years, with the increase in consumer demand and agricultural cooperation under the background of the "Belt and Road", durian has also attracted attention in new regions, including China, and research on variety improvement and introduction and cultivation has gradually emerged. The durian plant is an evergreen tree, with both flowers and a thick and prickly shell and a edible pulp with fragrance. The unique smell of durian comes from the complex combination of volatile compounds in its flesh, including sulfur-containing compounds, esters, etc. This unique trait makes durian valuable both economically and scientifically research.

The Malvaceae family belonging to durian is one of the major families among angiosperms, also known as the Hibiscus family. It contains about 244 genera and 4 225 plants, distributed in tropical, subtropical and temperate regions around the world. It has both herbs and trees, including cotton, cocoa, hibiscus, kapok and many other species with economic or ornamental value. The genus *Durio* has traditionally been classified in the family Bombacaceae and is closely related to genera such as *Bombax* and *Ceiba* in the modern classification.

The Malvaceae can be divided into evolutionary branches composed of two main branches, and the other branch includes the subfamily Bombacoideae, the Malvaceae, and the other seven subfamily. Molecular systems studies show that a branch of the evolutionary tree of the subfamily Bombacoideae and Malvaceae. Therefore, durian, a member of the subfamily Bombacoideae, occupies a relatively derived position in the Malvaceae family, and its relative species include plants such as the Kapok, and is slightly evolutionarily far from Cocoa (formerly classified in the family Sterculiaceae, now placed in the genus *Grewia* of the family Malvaceae.).

Although durian is of great economic value, genetic improvement has been slow for a long time due to its tall tree, long growth cycle and special after-ripening behavior of fruits (rapid quickly soft ripen and deteriorate). In the past decade, with the development of genomic technologies such as high-throughput sequencing, major breakthroughs have been made in durian genomic research. This study aims to systematically review the phylogenetic location of durian and its comparative research progress with the genome of related species of the Malvaceae family. The content covers phylogenetic reconstruction methods, durian genome characteristics, gene family and genome replication, key functional gene evolution, as well as the application prospects of these studies in durian variety improvement and Malvaceae crop breeding, in order to provide reference for the in-depth research and utilization of durian and Malvaceae plants in the future.

2 Phylogenetic Analysis of Durian

2.1 Molecular markers and methods for phylogenetic reconstruction

With the development of molecular biology technology, DNA sequence data has become the main basis for reconstructing plant phylogenetic relationships. Early studies based on combined analyses of chloroplast *rbcL* and *atpB* gene sequences supported the merger of the formerly separate families Bombacaceae, Sterculiaceae, and others into the family Malvaceae. Subsequent studies have gradually increased the types of molecular markers to obtain a more reliable phylogenetic tree. In recent years, with the popularization of high-throughput sequencing, the whole genome of chloroplasts of multiple genera of the Malvaceae have been compared and the phylogenetic tree has been constructed to parse to a more refined level for the whole genome of the Malvaceae plants (Liu et al., 2023). Maximum likelihood (ML) and Bayesian inference (BI) are commonly used phylogenetic tree construction methods, and differentiation time can be estimated in combination with molecular clock models. It is worth noting that the evolutionary history of Malvaceae is relatively complex, and genome-wide replication and hybridization may lead to the phenomenon that trees are inconsistent with species trees. Therefore, in recent years, higher-resolution data sources such as genome ubiquitous single-copy sequences and metagenomic analysis have also been used to reconstruct the phylogenetic of Malvaceae (Cheng et al., 2020).

2.2 The evolutionary relationship of durian in the malvaceae

Based on the latest phylogenetic research, durian belongs to the Bombacoideae branch of the subfamily Bombacoideae (Bombacoideae) on the Malvaceae system tree, and is closely related to the genus *Bombax* and *Ceiba*. Phylogenetic analysis of the Malvaceae shows that the subfamily Bombacoideae is a relatively derived branch of the Malvaceae and often forms a sister group with the Malvaceae. In the phylogenetic tree constructed by Huy et al. (2023) using chloroplast fragments, the subfamily Bombacoideae and the Malvaceae together form the "Malvatheca" branch. In a larger analysis, the subfamily Bombacoideae, including durian, belongs to the "Malvadendrina" branch of the Malvadendrina family, together with the Sterculioideae, which contains cocoa (Sterculioideae), etc. (Huy et al., 2023). This is consistent with the expectations of morphological classification: durian and kapok belong to the same subfamily of Kapok, both of which are tall trees and have large broad fruits (Figure 1). In addition, durian has a relatively distant relationship with cocoa (Startaceae subfamily), and cocoa was separated from the main trunk of the Mallowe family earlier. The latest chloroplast genome analysis further supports this: the winged subfamily tends to form an independent branch with the Grewioideae subfamily, while the subfamily Bombacoideae where durian is located is clustered into another large branch with the Mallowe subfamily (Cvetković et al., 2021).

2.3 Estimation of species differentiation time and its evolution events

The molecular clock model can be used to infer the differentiation age and evolutionary history of durian and its relative species. Based on the correction of the molecular phylogeny trees of the Malvaceae, some studies speculate that the radiation differentiation of the major branches of the Malvaceae occurred between the end of the Cretaceous and the early Paleogene (about 60 to 70 million years ago), which may be related to continental division and climate change at that time (Cvetković et al., 2021). The subfamily Bombacoideae where durian is located is believed to have begun to differentiate during the Paleogene period, and its common ancestor with the Mallowe subfamily may have experienced a genome-wide replication event. At the genus level, molecular systems analysis estimated that Durio and *Bombax* diverged in the early Cenozoic, and the preliminary time was estimated to be about 30 to 40 million years ago. This period coincided with the movement of plates in Southeast Asia and the warmth of climate, which provided an opportunity for the evolution and spread of tropical tree species such as durian. Although the direct fossil evidence of durian in the fossil record is limited, the Malvaceae family has generally seen a diverse pollen fossils appear in the Eocene-Oligocene, supporting that plants in this family have flourished in the Paleogene period. However, the lineage containing durian did not undergo such extensive whole-genome duplication events, and thus its genome structure has remained relatively ancestral.

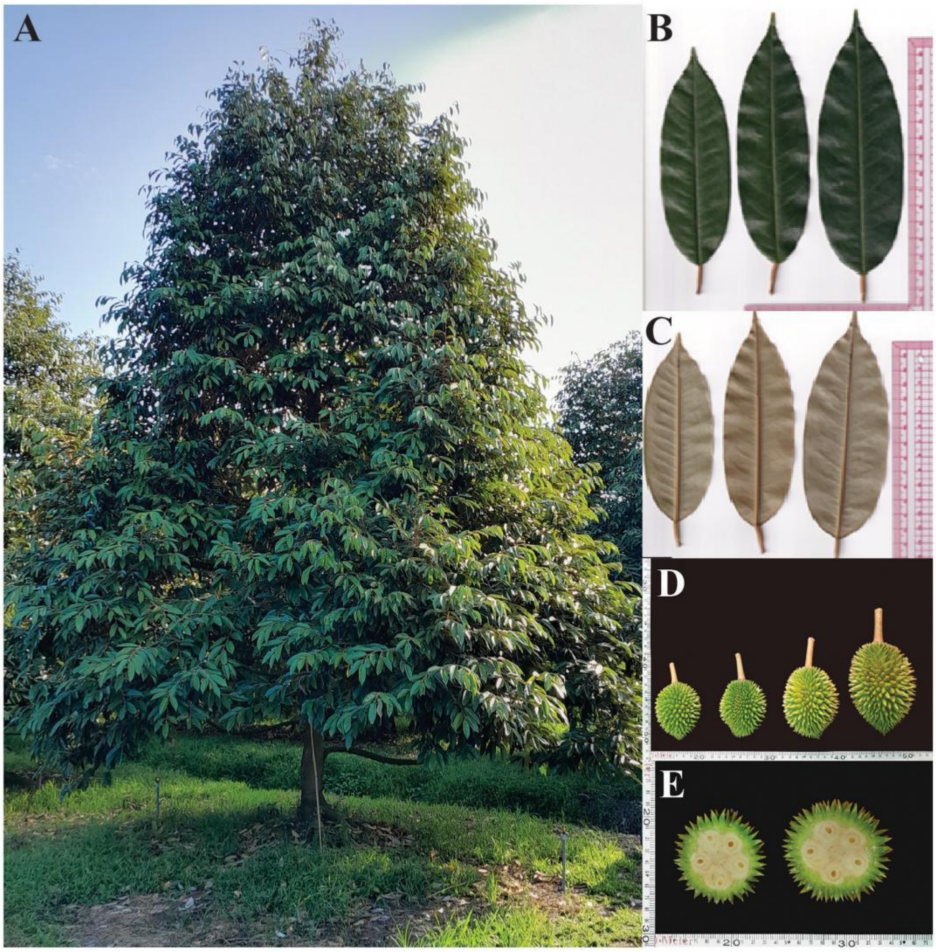

Figure 1 Phylogenetic development of durian (Adapted from Huy et al., 2023)

3 Genome Characteristics of Durian

3.1 Assembly and annotation of the durian genome

The deciphering of the durian genome is a milestone achievement in the genomics of the Mallowe family. In 2017, an international research team took the lead in reporting the genome sketch of the durian commercial breed "Maoshanwang". The genome is assembled by third-generation PacBio sequencing combined with second-generation Illumina sequencing. Its size is about 735 Mb, and the number of predicted encoding genes exceeds 45 000. The number of this gene is almost twice that of Arabidopsis and nearly twice that of human genes. In recent years, with the advancement of sequencing and assembly technologies, the quality of the durian genome has been further improved. Scientists from Thailand and other countries have genome resequencing and assembled other cultivars of durian, and obtained assembly sequences of about 762~833 Mb for high-quality genome assembly of varieties such as gold pillows and long handles (Nawae et al., 2023). Using the auxiliary framework constructed by Hi-C technology, researchers have improved the assembly of durian genome to chromosome level (2n=56), which can clearly identify 28 pairs of durian chromosomes. By constructing the durian pan-genome, the study found that the core genome covers about 92% to 96% of the reference gene set, while there is a certain number of deletion/existence variant (PAV) genes between different varieties. Some of these variant genes are related to flowering and disease resistance, which may explain the differences in traits of durian varieties.

3.2 Structural and functional characteristics of the durian genome

The durian genome has the general characteristics of the Malvaceae family and also shows some of its own uniqueness. The durian genome is about 750~830 Mb in size, and it belongs to the medium and large range in the Malvaceae family. In contrast, another tropical fruit tree representative of the Mallowe family, the genome of the cocoa tree is only about 430 Mb in size; while the genome of the cultivated cotton (ali-tetraploid) is as large as about 2.2 Gb. Therefore, the durian genome size is between cocoa and cotton. The chromosome number of durian is consistent with 2n. High-quality assembly revealed that the durian genome contains 28 pairs of chromosomes, and the gene arrangement on each chromosome follows a certain collinearity, but some translocations or inversions of chromosome fragments have also occurred. In terms of gene density, the durian genome is relatively dense, with an average of one gene for every about 13 to 15 kb, which is the same as the cocoa genome gene density of about one gene for every 10 kb. Functionally, through gene family analysis, it was found that a considerable proportion of genes in the durian genome belonged to the multi-copy family. It is particularly worth mentioning that the methionine gamma-lyase (MGL) gene responsible for the generation of volatile sulfur-containing compounds is encoded by a small family of genes in the durian genome, while in some related species, the gene has only a single copy (Pinsorn et al., 2023). This gene expansion may be closely related to the traits of durian flesh that release strong sulfur odors (Teh et al., 2017). In addition to odor-related genes, some other gene families of durian also show unique characteristics. Studies have shown that compared with the cotton genome, most protein families in the durian genome have a slower evolution rate and a higher degree of conservatism, while the protein families involved in transcriptional regulation and signaling evolve slightly faster in durian, which may reflect their directed evolution in adapting to a specific ecological environment.

3.3 Transposons and their role in genomic evolution

Like many higher plants, a considerable proportion of the durian genome consists of repeat sequences and transposons. According to preliminary statistics, about half of the sequences in the entire genome of durian belong to various repeating elements, among which long-term repeating retroretrosposons (LTR retrotransposons) occupy an important position. The complexity of the durian genome is largely due to these repeats. Transposons are believed to play the dual role of "genome builder" and "mutation maker" in the evolution of durian genome. On the one hand, the insertion and amplification of transposons increase genome size and lead to changes in chromosomal structure through mechanisms such as asymmetric cross-lines. On the other hand, transposon activity may provide new functions for the genome. Some transposon sequences were found in the durian genome to adjacent genes related to fruit development, suggesting that they may be involved in regulating the expression of these genes (Lin et al., 2019).

Transposon evolution dynamics are not the same in different species of the Malvaceae. Some studies compared the LTR retroposons in the durian and cotton genome, and found that the LTR copy accumulation rate in the durian genome is slower than that in cotton. Cotton has experienced a recent LTR surge in evolution, resulting in a significant expansion of its genome, while durian appears to have less large-scale LTR expansion in recent times. This may explain why the genome size of durian is significantly smaller than that of cotton, although durians have also experienced ancient polyploidization events. Transposons may also carry gene fragments through a copy-paste mechanism, forming new gene fusion or regulatory elements.

4 Comparative Genomics Study of Related Species of Durian and Malvaceae

4.1 Analysis of genome collinearity and evolutionary conservatism

An important part of comparative genomics is the genomic collinearity analysis between different species to reveal the conservatism of chromosomal arrangement and gene order. Collinear comparison between durian and multiple species within the Malvaceae family shows that on a larger evolutionary scale, their genome and relative species have maintained a certain degree of conservatism. The chromosomes of durian and cocoa can be traced back to about 7 original chromosomes of the ancestors of the Malvaceae family, and most homologous genes still maintain one-to-one correspondence between the two (Teh et al., 2017). During species-specific evolution, different recombination and mutations have also taken place in their respective genomes. A comparative study of durian and cotton found that although the two share some large collinear segments, the gene sequence of doubling and translocation has been relatively rearranged, while the durian genome is still relatively close to the configuration of the ancestral ancestors of the Malvaceae family (Alves et al., 2023). To quantify collinearity and conservatism, researchers often calculate the coverage ratio of homologous fragments between genomes. The comparison between lotus and cocoa or durian and cotton is consistent with the proximity of phylogenetic relationships. Additionally, comparative analysis with other sequenced genomes in the Malvaceae family—such as ornamental Hibiscus species like *Hibiscus syriacus* and *Hibiscus rosa-sinensis*—revealed that certain regions of the durian genome are highly conserved and form a "syntenic backbone.". These regions are rich in basic metabolic and developmental key genes and maintain synonyms in almost all Malvaceae species.

4.2 Gene family expansion and contraction modes

Expansion or contraction of gene families is one of the main sources of genome differences in different species. The study pointed out that the MGL gene family of durian undergoes germline-specific expansion, thereby increasing the accumulation of sulfur metabolites and providing a strong odor to the flesh. Similarly, the ethylene signal pathway genes that regulate fruit ripening also show patterns in durians that differ from other species. A comparative analysis revealed that the number of genes encoding ethylene biosynthetic enzymes and ethylene signaling components in the durian genome is comparable to that in cacao, but these genes in durian tend to have multiple copies of homologous genes and are expressed in large quantities during maturity, while the expression levels of related genes in odorless cocoa fruits are lower (Thongkum et al., 2018). There are also copy number differences between genes such as polygalacturonidase and cellulase related to fruit softening: durian may have expanded some cell wall degradation enzyme genes, so that their fruits can quickly soften and fall off after ripening. Comparative results show that the number of *R* gene family members encoding NBS-LRR disease-resistant proteins in the durian genome reaches more than 2 500, slightly higher than species such as cocoa. Cortaga et al. (2022) identified 2586 RGA genes in the durian genome in detail and found that some of these subfamilies (such as the receptor kinase RLK family) are particularly abundant in durians, while in some of the close relatives, are relatively small. This may be related to the pathogenic pressure in the long-term wild growth environment of durian, which allows it to retain more disease-resistant genes.

4.3 Genome-wide replication events and their evolutionary significance

Genome-wide replication (WGD) is an important recurring event in the history of angiosperm evolution and has a profound impact on species diversity and gene function innovation. The Mallow family is no exception, with WGD occurring multiple times in its lineage. Comparative genomics evidence shows that there is an ancient WGD event in the Malvaceae family, which may be a γ-multiplier event shared by all chrysanthemums (core vermicelli) that occurred around the mid-Cretaceous period (Cvetković et al., 2021). As a member of the true dicot, durian retains the remains of this ancient WGD: some paired homologous segments can be detected within the genome, as well as many paired homologous genes (i.e., ancient duplicate genes). Some of these ancient repeat genes have differentiated into new functions, while others have gradually become inactivated by functional redundancy. Conover et al. (2019) conducted collinear analysis of the transcriptome and genomes of multiple genera of the Malvaceae family.

They found that cotton has a 5-6-fold genome doubling relative to cocoa (including ancient gamma events and cotton itself doubling), while subfamily Bombacoideae species including durian only exhibit ancient gamma events and no additional new WGD. This shows that durian and its close relatives, Kapok, Jibe, etc., may have evolved in diploid form after the Pliocene, without large-scale polyploidization. On the contrary, cotton has experienced multiple doublings, and has obtained rich genes related to fiber development, which has led to the emergence of a new phenotype of long velvet fiber.

5 Evolution of Key Genes of Durian and Its Relative Species

5.1 Evolutionary analysis of genes related to fruit development and maturation

The fruit is the most prominent organ of durian, and its development and maturity process is precisely regulated by a series of genes. Durian is a respiratory jump-transformed fruit that produces ethylene peaks during maturity and undergoes rapid softening and shedding. The relative cocoa fruit is non-jumping, lacks obvious ethylene peaks, and has fewer copies of ACS and ACO genes in its genome (1 each) and has a stable expression level. Although the durian ethylene receptor (ETR1/ETR2) gene did not expand significantly, its coding sequence undergoes some positively selected site changes, which may affect the sensitivity to ethylene concentration. Key transcriptional regulators such as the NAC and MADS-box families that regulate fruit maturity, and unique members also exist in the durian genome.

DzNAC, a Durian-specific NAC transcription factor, was identified in durian, which was strongly upregulated when the fruit was ripe and was supposed to be a "ripening switch" similar to the tomato RIN gene (Thongkum et al., 2018). This gene has not seen significant homologous expansion in other Malvaceae species and may be a novel functional gene unique to durian. Transcriptome data show that during the stage where durian fruits are from unripe (green hard) to ripe (soft) , a large number of cell wall hydrolase genes (such as polygalacturonidase PG, cellulase Cel, pectin methyl esterase PME, etc.) are significantly upregulated (Palapol et al., 2015). Evolutionary analysis of these genes found that some PG genes in durian have longer promoters than the homologous genes of kapok, and ethylene response elements are enriched on them, indicating that durian may have enhanced the regulation of ethylene on these mature enzyme genes through promoter evolution. Although the basal metabolic pathways are generally similar in fruit development of the three, durian is particularly enriched for upregulation of vitamin metabolism and amino acid metabolism-related genes when ripened.

5.2 Comparative analysis of biosynthesis pathways of secondary metabolites

Durian is known for its special aromatic scent, which comes from the rich and unique secondary metabolites in the fruit. Comparative genome and metabolomic studies revealed that sulfur-containing compounds and esters dominate the volatile components of durian fruits, which is closely related to the expansion and high expression of durian in the metabolic pathway gene. In terms of sulfur-containing volatiles biosynthesis, methionine is a precursor substance, which degrades and metabolizes to produce odorous molecules such as methylene mercaptan. There are significantly more genes encoding methionine conversion-related enzymes (such as MGL and ACC decarboxylase) in the durian genome than those of related species such as kapok and cocoa, and the mRNA levels of these genes increase significantly when the fruit matures. Experiments have shown that treating durian fruits with 1-MCP (an ethylene action inhibitor) can significantly inhibit the formation of thiol volatiles (such as diethylene disulfide and trisulfide cyclopentane), while decreasing MGL activity and expression were detected. This shows that the synthesis of volatile sulfides of durian is regulated by ethylene, which promotes sulfur metabolic flow by upregulating genes such as MGL. However, in fruits such as cocoa without strong sulfur odor, no similar gene expansion and expression surge were observed. In addition to sulfur-containing compounds, durian fruits are also rich in esters, giving them a sweet fragrance.

The fruits of Kapok are floc-shaped seeds with no obvious odor. Although the corresponding pathway gene exists homologously, it may become a pseudogene or have extremely low expression due to loss of function. In other secondary metabolism, durian fruits are rich in polyphenols such as flavonoids and anthocyanins, which give them certain antioxidant properties. Genome comparisons show that durian and cocoa are relatively conservative in the polyphenol synthesis pathway genes, and both have complete copies of key genes such as *PAL*, *CHS*, and *DFR*, indicating that this path is strongly retained in the Mallowaceae fruit trees. However, the color of the durian flesh is pale yellow, and the accumulation of anthocyanins is not significant. The corresponding UDP-glucose:anthocyanin glucose transferase gene is weakly regulated in durian, which is related to the species difference in the expression of the pulp color gene (Zheng et al., 2024). The secondary metabolism of other parts of the durian tree is also unique, and the bark and leaves contain some flavonols and terpenes, which are considered to have medicinal value.

5.3 Evolutionary characteristics of disease resistance and stress response genes

Durians face challenges in wild growth environments with pathogenic infection and abiotic stress (such as drought, salting, etc.), so genes related to disease resistance and stress resistance in their genome have undergone interesting changes in long-term evolution. Comparative genomic studies show that the disease-resistant genes of Malvaceae (mainly NBS-LRR-like *R* genes) vary greatly between different species, and the *R* gene family of durian is particularly large. Cortaga et al. (2022) identified 2 586 *RGA* candidate genes in the durian genome, including multiple types of disease-resistant gene subclasses. The number of *R* genes reported in the cocoa genome is about 1,100, significantly less than that of durian. The number of *R* genes in the Kapok genome is between durian and cocoa, which may reflect the different pathogenic pressures of each species in their respective ecological environments. Further phylogenetic analysis showed that in the R gene family of durian, several clusters are dilated clusters unique to durians. These expanded *R* gene clusters often undergo positive selection, with higher variation in key amino acid sites encoding proteins suggesting that durian maintains a diverse library of resistance in the arms race against pathogens. In addition to the *R* gene, there are a large number of genes in the durian genome that encode receptor-like kinases (RLKs) and receptor-like proteins (RLPs), which are involved in pathogen recognition and signal transduction (Khan et al., 2016). In terms of abiotic stress, durian is native to a wet and hot tropical rainforest, but it also needs to deal with periodic drought stress and soil poverty. Comparative analysis found that durian and kapok with strong drought resistance have differences in some stress-resistant genes (Sathapondecha et al., 2024). When durian is under drought stress, these genes respond quickly, improving the ability to close stomatals, while kapok may be less dependent on these genes due to its deep roots. Nawae et al. (2023) found that by comparing the gene evolution rates of durian and cotton, etc., that protein families related to biological and abiotic stress responses (such as transcriptional regulators, protein kinases, etc.) evolved faster in durian. This shows that durian may have experienced great environmental pressure selection in recent days, which has rapidly improved and optimized these stress-resistant genes.

6 Case Analysis: Comparison of Genomes Between Durian and Specific Related Species

6.1 The similarities and differences between durian and cocoa

Cocoa and durian are both famous tropical fruit crops of the Malvaceae family, but the two belong to different subfamilies in phylogenetic conditions, and are completely different in terms of fruit shape and flavor. Comparing the genomes of the two, it was found that there were both conservative commonalities and significant differences. In terms of genome size and gene number, the cocoa genome is approximately 430 Mb, encoding about 28 798 genes, significantly smaller than the ~750 Mb and 46 000 genes of durian. This difference is mainly due to the different proportion of repeats and transposons: about 35% of the cocoa genome is repeated, while the durian is close to 50%. Cocoa has evolved in a relatively stable under-forest environment for a long time, with a low degree of genome expansion and has not experienced recent large-scale transposon bursts and doubling events. In contrast, although no new WGD occurred in durian, some transposon amplification is more active and the genome "expansion" is more obvious). In terms of chromosomes and collinearity, the number of chromosomes of cocoa is 2n=20, which is quite different from that of durian 2n=56. However, through comparative studies, it was found that 10 pairs of cocoa can correspond to some chromosome combinations of durian, which implies that the ancestral chromosomes of the two have undergone several fusion/schizoosomes (Argout et al., 2011). In terms of functional genes, both have some common core genes, such as genes related to basal metabolism and growth development, which are highly conserved (collinear and almost no sequence changes) in both genomes. But genes associated with fruit traits show significantly different evolutionary trajectories. Cocoa beans are rich in cocoa alkali and flavanols, giving chocolate a bitter flavor and excitement properties. Genes responsible for synthesis of these secondary metabolites (such as the caffeine synthesis gene *XAN* and polyphenol oxidase PPO, etc.) are particularly developed in the cocoa genome, whereas most of these genes in the durian genome have only basic copies or are not present at all. Durian has its own unique volatile thiol synthesis pathway, while the cocoa genome has no corresponding expansion gene. In addition, cocoa seeds are rich in fat (cocoa butter), while durian seeds are mainly starch. Comparatively, the gene expression level of oil and fat synthesis pathways (fatty acid synthesizers, oleate dehydrogenase, etc.) in the cocoa genome is much higher than that of durian seeds, which may be caused by regulatory evolution rather than differences in gene copy number.

6.2 Phylogenetic relationship and genome comparison between durian and kapok (*Bombax Ceiba*)

Kapok and durian have similarities in tree morphology and ecological habits (both tropical trees), but they are very different in reproduction and fruit characteristics: Kapok fruit is a capsule, which cracks and releases fiber-like flocculent hairs when ripening, and spreads according to the wind; while durian fruit is a thick-skinned berries that rely on animals to spread seeds. Therefore, comparing the genome of durian and kapok can help resolve the genetic differences in different transmission strategies species within the subfamily Bombacoideae. In terms of phylogenetic, the latest genomic data supports that durian and kapok are indeed very close: the genus DNA sequence of kapok has a high homology with durian, and the whole genome comparison found that about 70% of the sequences of the two can be compared in comparison, and the differences are mainly concentrated in durian regions and species-specific genes. Collinear analysis shows that durian and kapok show a one-to-one correspondence on most chromosomes, and only a small-scale inversion occurred at the ends of individual chromosomes. This high degree of collinearity shows that the two have not experienced major chromosomal recombination events since their common ancestor differentiation, and it also once again confirms the closeness of their systemic relationship (Shao et al., 2024). In terms of gene content, the number of genes encoding durian and kapok genomes is similar. Kapok predicts that there are about 38,000 genes, slightly less than the 46 000 of durian, but considering the differences in annotation strategies, this gap may not be significant. Compared with durian, a unique content worthy of attention in the kapok genome is the collection of genes related to fiber development. Preliminary analysis found that some cytoskeleton regulation and cellulose synthase gene families in the kapok genome were in an expanded state. For example, the gene encoding cellulose synthase CesA has more copies of the cellulose synthase than durian in kapok, and was found to be highly expressed during capsule development (Figure 2). Durians do not produce fibrous structures, so these genes may not be so strongly selected to expand or express in durians (Ding et al., 2023). In addition, differences in reproductive strategies between kapok and durian may be genetically cues. Durians rely on animals (such as civets, primates) to spread seeds, which requires conspicuous fruits and nutritious pseudo-seed coats to attract animals. Kapok is spread by wind, and its seeds do not require a large amount of nutrients to be wrapped, but instead require reduced mass and increased air delivery distance. This distinction may be reflected in gene regulation related to seed development and metabolism. Sequencing data show that the starch and oil content of kapok seeds is lower than that of durian (*Bombax* seeds are dry and contain a small amount of oil and fat), while durian seed development has higher oil and fat metabolism gene activity.

6.3 Analysis of the genetic basis of unique traits of durian through comparative genomics

The unique biological traits of durian have always been a question of interest to researchers. Comparative genomics provides an effective way to compare durians with other species to identify gene or sequence characteristics that may confer uniqueness to durians. By comparing with relative species without strong odor (such as kapok and cocoa), a batch of candidate genes closely related to odor formation was determined. For example, transporters and sulfur metabolic enzymes that control the release of sulfur compounds are specifically expressed in durian flesh, which are usually silent or missing in odorless fruits. Analysis of these genes on phylogenetic trees revealed that most of them belong to new or rapidly evolving branches in the durian lineage, suggesting that the genetic basis of durian odor is its relatively recent innovation in evolution. Verifying the function of these candidate genes through gene editing will be an important direction in the future.

If the durian *MGL* gene is knocked out, whether the volatile thiol content of the fruit will decrease, the contribution of the gene to the odor will be directly tested. Secondly, the shape of durian fruit is also unique: the shell is thick, hard and densely covered with spikes, which is conducive to preventing predation when immature, but after ripening, the shell automatically cracks the seeds to expose and facilitate animal feeding and spreading. The genetic mechanisms associated with this can be found in the comparison. The fruit forms of durian and kapok are very different. Researchers compared the two genes related to peel development and found that in the durian genome, several genes that control the formation of the choke layer have significantly expanded and are highly expressed in the peel, while these genes in kapok only have a single copy (Pinsorn et al., 2024). This may explain the durian peels are thick and hard and resin-rich. On the other hand, enzymes that control peel cracking (such as endoglucanase that hydrolyzes the gum) are upregulated during the ripening period of durian, causing regular cracks to form fruit shells, while kapok fruits crack through different destratigraphic mechanisms. These different gene regulatory networks can be identified by comparing the transcriptome, thereby locking in key genes that regulate durian husk cracking. Durian seeds have a thick and juicy pseudo-seed coat (easy-eating flesh), a structure that attracts animal transmission. However, relatives such as kapok have no similar structure, and their seeds directly attach fibers. It was found that when durian develops, it initiates a new gene network similar to that formed by cotton seeds, except that the action is different: cotton activates fibroblast elongation genes, and durian activates beads thickened by cells and stored nutrient genes. This suggests that durian may have “borrowed” the seed development regulation modules already in ancestral plants to create its unique pseudo-seed coat through changes in expression patterns.

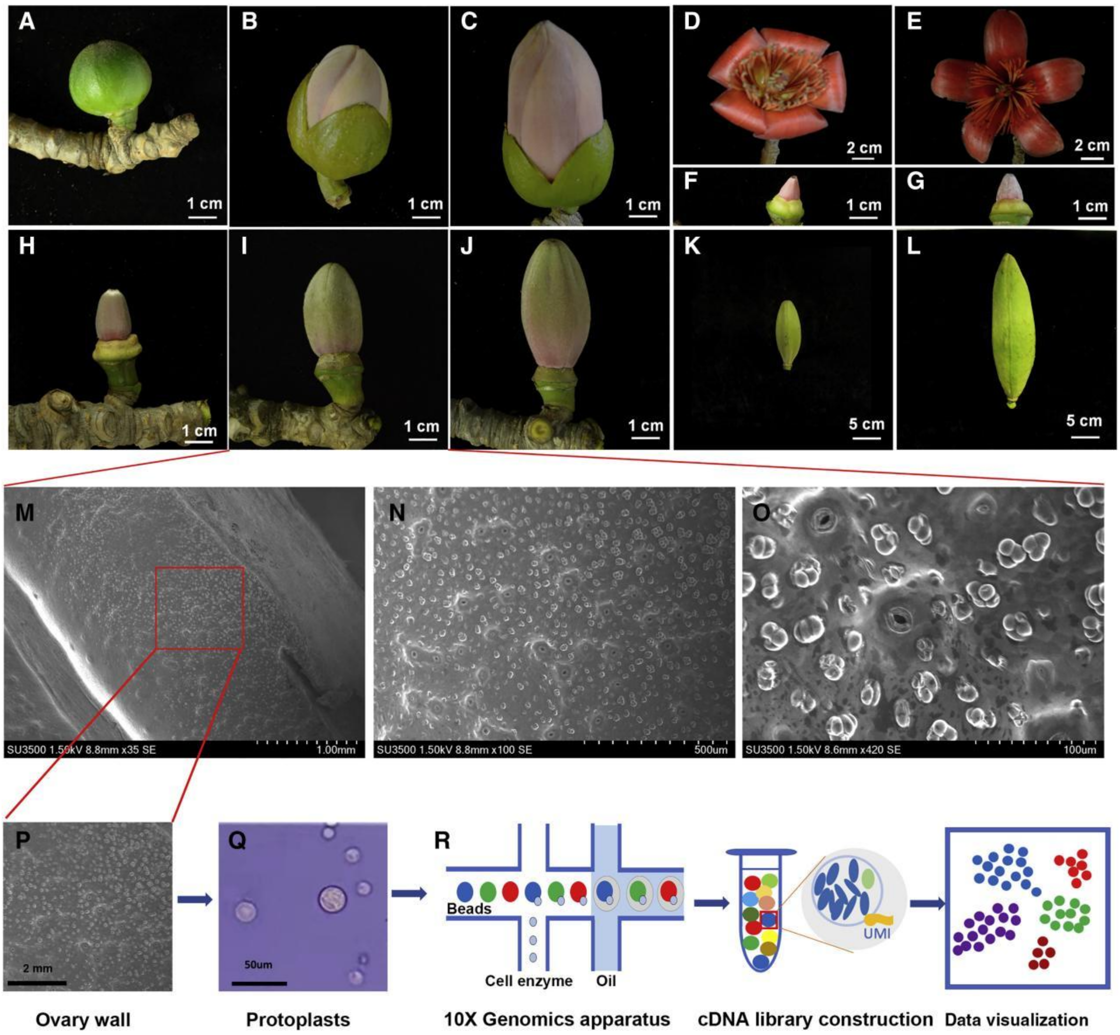

Figure 2 Identification and application of scRNA-seq in fibre initiation of *B. ceiba* (Adopted from Ding et al., 2023)

7 Application and Future Prospects of Durian Genome Research

7.1 Potential Application of Durian Genomics in Variety Improvement

The deciphering of the durian genome brings unprecedented opportunities for variety improvement. Traditional durian breeding mainly relies on real seed selection and hybridization. Due to the long generation cycle and unclear genetic basis, breeding progress is slow. The massive genetic information and molecular markers provided by genomics can greatly accelerate this process. Genome data allow the construction of high-density genetic linkage maps and genome-wide association analysis (GWAS) to localize the main-effect gene or quantitative trait loci (QTL) that control important agronomic traits. Research in Thailand has identified some dominant alleles unique to varieties such as "long-handle", including candidate sites that affect the mellowness and aroma of the flesh. The genetic diversity pattern of durian revealed by genomic sequences can be used for the protection and rational use of variety resources. There are hundreds of durian cultivars in Malaysia and Thailand. Genomic analysis has found that the genetic distance between different varieties is far and near, and some varieties are highly similar in actual genotypes. Genomics can also help discover excellent genes of wild relatives of durian. For example, wild species such as *Durio graveolens* may have genes that are resistant to disease or early fruit. These differential genes can be found by comparing the genome and introduced into cultivated durians through interspecies hybridization (Mursyidin et al., 2024). The application of genomics in the durian industry also includes the use of gene molecular markers for variety authenticity identification and germplasm identity identification. Since morphological differences between durian varieties are sometimes not obvious, DNA-based identification techniques can improve accuracy. For example, by developing varieties-specific SNP markers or barcode sequences, it is possible to quickly detect whether durian fruits or seedlings come from designated varieties, which is helpful for trade and planting (Kulkarni et al., 2024).

7.2 Optimize the fruit quality and stress resistance of durian through gene editing technology

The rise of gene editing technologies such as CRISPR/Cas9 provides a revolutionary tool for genetic improvement in crops. In durian breeding, the potential application value of gene editing is huge, especially in optimizing fruit quality and stress resistance. Through gene editing, it is expected to cultivate new varieties that "low-odor durian" or change the proportion of aroma components. For example, targeting knockout of 1-2 *MGL* gene copies in durian can be considered to reduce the production of volatile thiols. Because the MGL family has multiple copies of the presence in durian, partial knockout may not be fatal, but may reduce odor concentration. For example, the *acetyltransferase* gene that regulates the ester synthesis pathway can affect the composition of the sweet smell of durian and make the aroma more pleasant. Properly downregulating the expression of some softening enzyme genes by gene editing or knocking out regulators that promote premature softening may delay the fruit softening speed, thereby extending shelf life without significantly affecting the flavor. Similar strategies have been successfully implemented on crops such as tomatoes, improving shelf life without only having a slight impact on texture. Durians are susceptible to some fungal diseases (such as tree trunk rot) and root diseases, and the selection of disease-resistant varieties is limited. Gene editing can be used to create new disease-resistant varieties. One idea is to knock out the gene (i.e., *S* gene) that encodes susceptibility factors in the durian genome, thereby improving disease resistance. The case of wheat obtaining broad-spectrum anti-powder mildew by knocking out the Mlo gene is inspired by screening genes in durian that play a negative role in the pathogen infection process and knocking it out, making durian immune to the pathogen (Panpetch and Sirikantaramas, 2021). Although the pathology research of durian is relatively lagging, these candidate *S* genes can be gradually discovered as genomic and transcriptomic data accumulate. Another idea is to edit durian's own disease-resistant gene to enhance its function. For example, mutate a disease-resistant *R* gene promoter through gene editing, causing it to be expressed stronger or constitutively when it encounters a pathogen, thereby providing a faster and stronger defense response. In terms of stress resistance, such as drought and flood tolerance, you can consider knocking out some negative regulatory genes that control stomatal closure, so that plants are more likely to retain water under drought stress; or regulate root development genes to enhance root tolerance to waterlogging. These technical ideas have been verified in model plants and are expected to be transplanted on durian. In the future, we may be able to see new durian products optimized by "genetic surgery", such as durian with controllable fragrance, durian that is resistant to storage and transportation after harvest, durian that is immune to specific diseases, etc., which provide strong support for industrial development.

7.3 Application of comparative genomics in Mallowaceae crop breeding and biodiversity conservation

The Mallow family includes many important crops and economic plants, such as cotton, cocoa, hibiscus, jute, etc. Comparing durian with the genomes of these species is not only valuable in scientific research, but also can learn from each other in practice to promote the breeding and diversity protection of the entire Malvaceae family. Comparative genomics can help us discover common beneficial genes or excellent alleles among different Mallowaceae crops, thus achieving cross-species breeding inspiration. Comparative genomics can reveal the genetic basis of important agronomic traits of various species in the Mallowe family and promote the formulation of breeding strategies. For cocoa and durian, it can be found that the genes that affect seed flavor, size, etc. are different, so different genes should be focused on when breeding each. Many species of the Mallowe family face habitat loss and population decline in biodiversity conservation. Genome comparisons can help determine which genes or fragments of genetic diversity require focused protection. Genomics technology can also be used to construct population genetic databases of endangered plants in the Mallowaceae family, conduct DNA testing and genetic monitoring of wild populations to assist in the formulation of protective measures. Data and methods for comparing genomes can play a role in the domestication of Malvaceae crops. By comparing the genome differences between cultivars and their wild relatives, key domesticated genes can be identified under human selection.

Acknowledgments

The author appreciates the modification suggestions from Professors Rudi Mai and Qixue Liang on the manuscript of this study.

Funding

This study was funded by the Hainan Tropical Agricultural Resources Research Foundation Project (Project No. H2025-03).

Conflict of Interest Disclosure

The author affirms that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Alves R., Abreu V.A.C., Oliveira R.P., Almeida J.V.A., Oliveira M.M., Silva S.R., Paschoal A.R., de Almeida S.S., de Souza P.A.F., Ferro J.A., Miranda V.F.O., Figueira A., Domingues D.S., and Varani A.M., 2023, Genomic decoding of Theobroma grandiflorum (cupuassu) at chromosomal scale: evolutionary insights for horticultural innovation, bioRxiv, 13: giae027.
https://doi.org/10.1101/2023.12.23.573069

Argout X., Salse J., Aury J.M., Guiltinan M. J., Droc G., Gouzy J., Allegre M., Chaparro C., Legavre T., Maximova S.N., Abrouk M., Murat F., Fouet O., Poulain J., Ruiz M., Roguet Y., Rodier-Goud M., Barbosa-Neto J.F., Sabot F., Kudrna D., Ammiraju J.S.S., Schuster S.C., Carlson J.E., Sallet E., Schiex T., Dievart A., Kramer M., Gelley L., Shi Z., Bérard A., Viot C., Boccara M., Risterucci A.M., Guignon V., Sabau X., Axtell M.J., Ma Z., Zhang Y., Brown S., Bourge M., Golser W., Song X., Clement D., Rivallan R., Tahi M., Akaza J.M., Pitollat B., Gramacho K., D'Hont A., Brunel D., Infante D., Kebe I., Costet P., Wing R., McCombie W.R., Guiderdoni E., Quetier F., Panaud O., Wincker P., Bocs S., and Lanaud C., 2011, The genome of Theobroma cacao. Nature Genetics, 43(2): 101-108.
https://doi.org/10.1038/ng.736

Cheng Y., Zhang L., Qi J., and Zhang L., 2020, Complete chloroplast genome sequence of Hibiscus cannabinus and comparative analysis of the Malvaceae family, Frontiers in Genetics, 11: 227.
https://doi.org/10.3389/fgene.2020.00227

Conover J.L., Karimi N., Stenz N.W.M., Ané C., Grover C., Skema C., Tate J.A., Wolff K., Logan S.A., Wendel J.F., and Baum D.A., 2019, A Malvaceae mystery: A mallow maelstrom of genome multiplications and maybe misleading methods? Journal of Integrative Plant Biology, 61(1): 12-31.
https://doi.org/10.1111/jipb.12746

Cvetković T., Areces-Berazain F., Hinsinger D., Thomas D., Wieringa J., Ganesan S., and Strijk J.S., 2021, Phylogenomics resolves deep subfamilial relationships in Malvaceae s.l. G3: Genes|Genomes|Genetics, 11(7): jkab136.
https://doi.org/10.1093/g3journal/jkab136

Ding Y., Gao W. J., Qin Y., Li X., Zhang Z., Lai W., Yang Y., Guo K., Li P., Zhou S., and Hu H., 2023, Single-cell RNA landscape of the special fiber initiation process in *Bombax Ceiba*. Plant Communications, 4(5): 100554.
https://doi.org/10.1016/j.xplc.2023.100554

Huy T.G., Thi N. P.A., Do H.D.K., and Khang D.T., 2024, The complete chloroplast genome of *Durio zibethinus* L. cultivar Ri6 (Helicteroideae, Malvaceae). Mitochondrial DNA Part B, 9: 625-630.
https://doi.org/10.1080/23802359.2024.2350619

Khan A.M., Khan A., Azhar M., Amrao L., and Cheema H., 2016, Comparative analysis of resistance gene analogues encoding NBS-LRR domains in cotton, Journal of the Science of Food and Agriculture, 96(2): 530-538.
https://doi.org/10.1002/jsfa.7120

Kulkarni K.P., Appiah R.K., Reddy U.K., and Melmaiee K., 2024, Cleaved amplified polymorphic sequence markers in horticultural crops: current status and future perspectives, Agronomy, 14(11): 2598.
https://doi.org/10.3390/agronomy14112598

Liu R., Xie L., and Wang D., 2023, The complete chloroplast genome of Malvaviscus penduliflorus (Malvaceae), Mitochondrial DNA Part B, 8: 886-889.
https://doi.org/10.1080/23802359.2023.2246670

Nawae W., Naktang C., Charoensri S., U-thoomporn S., Narong N., Chusri O., Tangphatsornruang S., and Pootakham, W., 2023, Resequencing of durian genomes reveals large genetic variations among different cultivars. Frontiers in Plant Science, 14: 1137077.
https://doi.org/10.3389/fpls.2023.1137077

Teh B.T., Lim K., Yong C.H., Ng C.C.Y., Rao S.R., Rajasegaran V., Lim W.K., Ong C.K., Chan K., Cheng V.K.Y., Soh P.S., Swarup S., Rozen S.G., Nagarajan N., and Tan P., 2017, The draft genome of tropical fruit durian (*Durio zibethinus*). Nature Genetics, 49(11): 1633-1641.
https://doi.org/10.1038/ng.3972

Pinsorn P., Sanachai K., Rungrotmongkol T., Hoefgen R., Watanabe M., Brueckner F., Nakabayashi R., Mori T., Oikawa A., Sasaki R., Sucharitakul J., Saito K., and Sirikantaramas S., 2023, Sulfur metabolism in durian pulps: Factors contributing to the production of volatile sulfur compounds during fruit ripening, Postharvest Biology and Technology, 206: 112533.
https://doi.org/10.1016/j.postharvbio.2023.112533

Lin J., Cai Y., Huang G., Yang Y., Li Y., Wang K., and Wu Z., 2019, Analysis of the chromatin binding affinity of retrotransposases reveals novel roles in diploid and tetraploid cotton, Journal of Integrative Plant Biology, 61(1): 32-44.
https://doi.org/10.1111/jipb.12740

Palapol Y., Kunyamee S., Thongkhum M., Ketsa S., and van Doorn W.G., 2015, Expression of expansin genes in the pulp and the dehiscence zone of ripening durian (*Durio zibethinus*) fruit, Journal of Plant Physiology, 182: 33-39.
https://doi.org/10.1016/j.jplph.2015.04.005

Panpetch P., and Sirikantaramas S., 2021, Fruit ripening-associated leucylaminopeptidase with cysteinylglycine dipeptidase activity from durian suggests its involvement in glutathione recycling, BMC Plant Biology, 21(1): 69.
https://doi.org/10.1186/s12870-021-02845-6

Pinsorn P., Hoefgen R., and Sirikantaramas S., 2024, The novel durian (*Durio zibethinus* L.) homeodomain-leucine zipper (DzHD-ZIP) transcription factor 1.8 regulates the production of volatile sulfur compounds during fruit ripening, Postharvest Biology and Technology, 210: 112776.
https://doi.org/10.21203/rs.3.rs-2611573/v1

Mursyidin D.H., Nazari Y., and Firnanda M.R., 2024, Assessment of durian diversity and its wild relatives (Durio spp.) based on leaf morphology and molecular marker. Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi, 34(3): 393-405.
https://doi.org/10.29133/yyutbd.1423236

Sathapondecha P., Suksri P., Nuanpirom J., Nakkanong K., Nualsri C., and Whankaew S., 2024, Development of gene-based InDel markers on putative drought stress-responsive genes and genetic diversity of durian (*Durio zibethinus*), Biochemical Genetics, 62(6): 4396-4407..
https://doi.org/10.1007/s10528-023-10638-9

Shao L., Jin S., Chen J., Yang G., Fan R., Zhang Z., Deng Q., Han J., Ma X., Dong Z., Lu H., Hu W., Wang K., Hu L., Shen Z., Huang S., Zhao T., Guan X., Hu Y., Zhang T., and Fang L., 2024, High-quality genomes of *Bombax Ceiba* and *Ceiba* pentandra provide insights into the evolution of Malvaceae species and differences in their natural fiber development, Plant Communications, 5(5): 100832.
https://doi.org/10.1016/j.xplc.2024.100832

Thongkum M., McAtee P., Schaffer R., Allan A., and Ketsa S., 2018, Characterization and differential expression of ethylene receptor genes during fruit development and dehiscence of durian (*Durio zibethinus*), Scientia Horticulturae, 240: 623-630.
https://doi.org/10.1016/j.scienta.2018.06.052

Zheng L., Zhang J., He H., Meng Z., Wang Y., Guo S., and Liang C., 2024, Anthocyanin gene enrichment in the distal region of cotton chromosome A07: mechanisms of reproductive organ coloration, Frontiers in Plant Science, 15: 1381071.
https://doi.org/10.3389/fpls.2024.1381071