International Journal of Molecular Evolution and Biodiversity, 2025, Vol.15, No.1, 40-50
http://ecoevopublisher.com/index.php/ijmeb

Research Report **Open Access**

Phylogenetic and Evolutionary Trajectories of Ananas: Evidence from Whole-Genome Data

Zhonggang Li [1] Yeping Han [2]
1 Tropical Specialty Crops Research Center, Hainan Institute of Tropical Agricultural Resources, Sanya, 572026, Hainan, China
2 Institute of Life Science, Jiyang College of Zhejiang A&F University, Zhuji, 311800, Zhejiang, China
Corresponding author: yeping.han@jicat.org

International Journal of Molecular Evolution and Biodiversity, 2025, Vol.15, No.1 doi: 10.5376/ijmeb.2025.15.0004
Received: 27 Dec., 2024
Accepted: 03 Feb.,2025
Published: 16 Feb., 2025

Preferred citation for this article:
Li Z.G., and Han Y.P., 2025, Phylogenetic and evolutionary trajectories of ananas: evidence from whole-genome data, International Journal of Molecular Evolution and Biodiversity, 15(1): 40-50 (doi: 10.5376/ijmeb.2025.15.0004)

Abstract Pineapple (*Ananas comosus*) is a tropical plant. It belongs to the Poales group and split off early in its evolution. Thanks to whole-genome sequencing, scientists can now study the pineapple and related plants in more detail. Research using both mitochondrial and nuclear DNA shows that all Ananas species come from a single common ancestor. It also helps place pineapple clearly within the Bromeliaceae family. This study looks at how comparing genomes helps us understand how important traits in pineapple evolved. The results show that the pineapple genome has not had a recent whole-genome duplication. But other changes, like sudden increases in transposons and repeated copies of certain genes, have changed many gene families. These discoveries show that having full genome data has made pineapple a useful plant for studying how monocots have evolved over time.

Keywords Pineapple; Phylogenomics; CAM photosynthesis; Gene family evolution; Genome sequencing

1 Introduction

The genus Ananas, which includes the common pineapple (*Ananas comosus*), is made up of small bromeliad plants. These plants come from the Neotropical region. Pineapple is the second most produced and traded tropical fruit in the world. People value it for its sweet taste and many uses. Besides its farming importance, pineapple has some special traits. It uses water very efficiently because it performs crassulacean acid metabolism (CAM). Most pineapple types also grow without seeds and are propagated asexually (Chen et al., 2019). Pineapple has a small genome-about 563 Mb per haploid-which makes it useful for research. It also sits at a basic position in the plant order Poales, so it is a good plant for studying how monocots evolved and adapted.

Ananas belongs to the family Bromeliaceae, an early-diverging lineage within the Poales, providing a phylogenetic reference for comparative evolutionary studies with major cereal crops (e.g., Poaceae) and other bromeliads (Ouyang et al., 2022; Yow et al., 2022). All species in the genus Ananas are native to the Americas and include only a few taxa, such as *Ananas comosus* and its wild relatives *Ananas bracteatus*, *Ananas ananassoides*, etc. Prior to the genomic era, phylogenetic relationships within the genus were primarily inferred from morphological traits and a limited number of DNA markers, leaving species boundaries and hybridization events unresolved.

This study presents the currently available pineapple genome data and focuses on the evolutionary processes of several important traits in pineapple, such as CAM photosynthesis, reproductive strategies, fruit development, and its ability to cope with stress. The study emphasizes how genomic changes-such as alterations in chromosome regions, duplicated segments, or active transposable elements-lead to distinct differences among plants. By integrating data from population genomics and trait mapping, this study identifies genes related to domestication and breeding strategies. It also discusses future research priorities for pineapple, highlighting how pineapple serves as a good model for studying monocotyledons and how its genomic data can support the improvement of vegetatively propagated crops.

2 Genomic Resources and Data Availability in Ananas Genus

2.1 Assembly of pineapple reference genomes

The first pineapple reference genome was made using the F153 variety. But because this species has many repeated sequences and is highly heterozygous, the genome wasn't very complete. Later, scientists used long-read sequencing and Hi-C technology to assemble a better version for the MD2 variety, which is popular in the market (Yow et al., 2022). This improved genome has a contig N50 of over 2.9 Mb. It also includes about 161 Mb of new or previously incomplete sequences and over 3 000 new genes. Pineapple has 25 chromosomes (2n = 50), and its haploid genome size is about 563 Mb. Compared to grasses, pineapple has experienced fewer ancient whole-genome duplications, keeping an early chromosome structure from the Poales group.

2.2 Resequencing and genetic diversity data

One major study resequenced 89 genomes from the Ananas genus. This included both wild and cultivated types. The study found over 7.2 million SNPs and around 920 000 indels (Xu et al., 2018). The cultivated *A. comosus* samples grouped into known types such as 'Smooth Cayenne', 'Queen', and 'Spanish'. Wild types like *A. bracteatus* and *A. ananassoides* formed their own separate groups (Feng et al., 2022). Some types, like 'Smooth Cayenne' and 'Queen', show signs of mixing from older lineages. But 'Singapore Spanish' seems to come from a single domestication without much mixing (Chen et al., 2019; Feng et al., 2022). This is likely due to clonal propagation. However, occasional sexual reproduction still brings new variations into some genetic backgrounds (Chen et al., 2019).

2.3 Functional genomics and organelle genomes

Scientists have sequenced the full chloroplast genomes of pineapple and its wild relatives. These genomes are about 162 kb in size and help in phylogenetic studies. One wild type, *A. comosus* var. *erectifolius*, is genetically very close to the cultivated pineapple, which agrees with how they are traditionally grouped (Liu et al., 2022). In terms of gene expression, several RNA-seq datasets are available. These cover different tissues like leaves, flowers, and fruits, as well as responses to stress. Some RNA-seq experiments done under day-night cycles found CAM photosynthesis genes controlled by the internal biological clock. In cold-treated pineapples of two different genotypes, many genes related to cold stress were found. These include genes for wax formation and hormone signaling (Yow et al., 2023). Xu et al. (2024) found 21 *HSP70* genes in pineapple. Many of these genes respond to heat and drought stress.

2.4 Data repositories and genomic databases

The pineapple genomics database (PGD) was created as a public platform for researchers. It includes genome sequences, gene information, genetic markers, and expression data for *A. comosus* (Xu et al., 2018). PGD provides online tools like genome browsers, search bars, and co-expression networks. Users can check the locations of about 46 000 pineapple genes, search for specific genes, or download SNP marker files from resequencing studies. Pineapple genome data can also be found in public databases like NCBI and Ensembl Plants. Research groups have connected genomic data to important traits like fruit quality and disease resistance, which helps in breeding programs.

3 Phylogenetic Relationships within the Genus Ananas

3.1 Species diversity and classification of ananas

The genus Ananas includes only a few species. In the past, many scientists considered it as a single species, *A. comosus*, with several botanical varieties or closely related species. These commonly include *A. comosus* (both farmed and wild types), *A. bracteatus*, *A. ananassoides*, *A. parguazensis*, *A. lucidus*, *A. monstrosus*, and *A. microstachys*. However, classification can differ depending on the source. New genetic research supports the idea that Ananas is a single group clearly different from other genera in the Bromeliaceae family. For example, Pseudananas is now often grouped into Ananas as *A. sagenaria*. All seven or so types in the genus show only small genetic differences, suggesting they split apart not long ago in evolutionary terms. Wild pineapples mostly grow in South America, especially in Brazil, Paraguay, and nearby places. Scientists believe this region is where both the genus and the cultivated pineapple first appeared (Figure 1) (Chen et al., 2019). Human movement spread the cultivated form, *A. comosus* var. *comosus*, to other tropical areas. Still, all farmed pineapples belong to this same group.

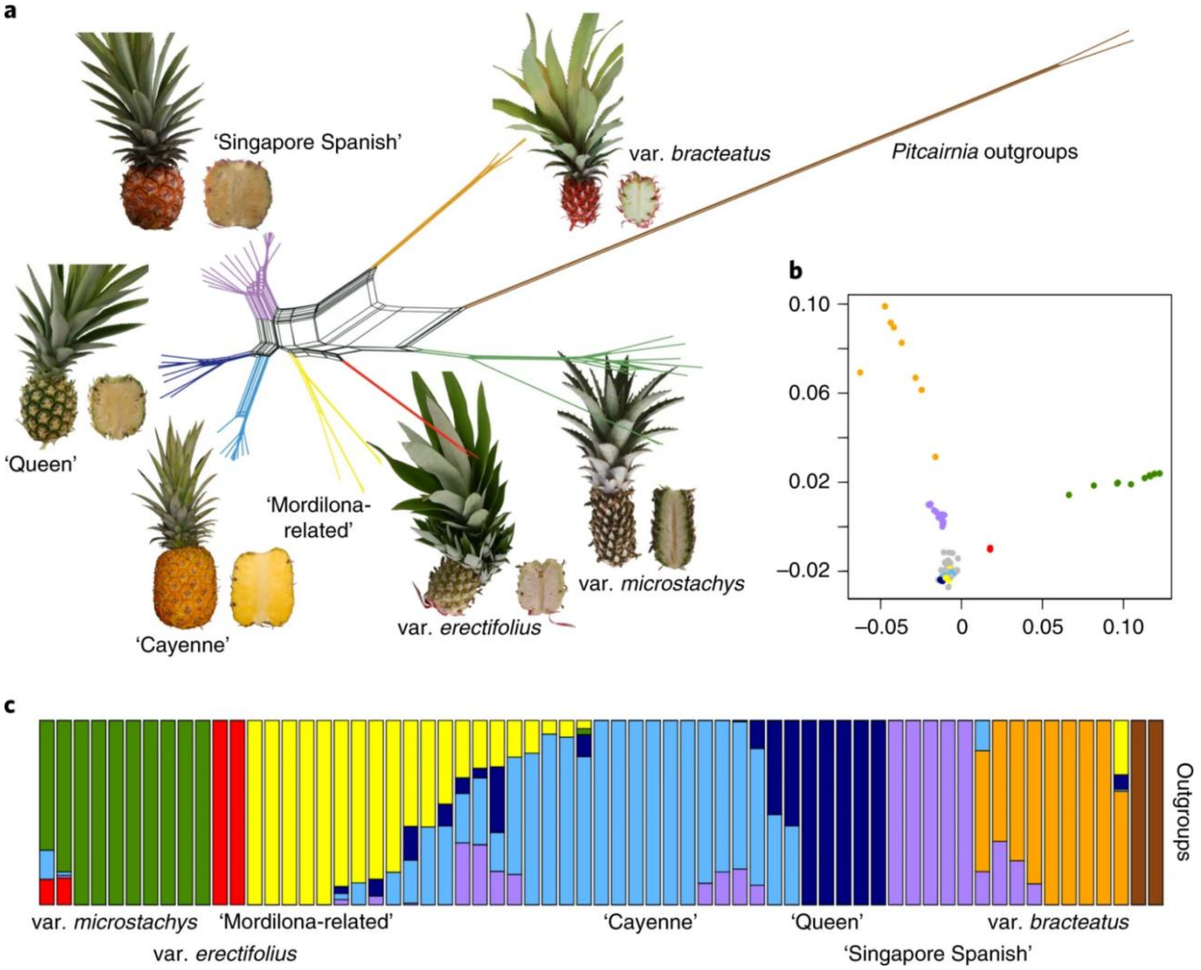

Figure 1 Pineapple population structure and admixture (Adopted from Chen et al., 2019)

Image caption: a: SplitsTree network for *Ananas* accessions excluding admixed samples. Green, variety *microstachys*; red, variety *erectifolius*; orange, variety *bracteatus*; yellow, variety *comosus*/Mordilona-related cultivars Cambray/Monte Lirio; purple, variety *comosus*/cultivar Singapore Spanish; light blue, variety *comosus*/cultivar Smooth Cayenne; dark blue, variety *comosus*/cultivar Queen; brown, genus *Pitcairnia* outgroups. A network of admixed samples is shown in Supplementary; b: MDS graphs of the studied *Ananas* accessions, with horizontal and vertical axes explaining 33.0% and 20.6% of the variance, respectively; color code follows that in a and admixed *A. comosus* genotypes are indicated in gray; c: Ancestry results from ADMIXTURE under the K = 8 model supported by an examination of cross-validation errors (Adopted from Chen et al., 2019)

3.2 Molecular phylogenetic studies based on organelle and nuclear genomes

Earlier DNA studies using plastid DNA or AFLP markers could not clearly show the family tree. But now, full organelle genomes and many SNP markers give better data. For example, *A. comosus* var. *erectifolius* is very close to *A. comosus* in plastid genome trees (Liu et al., 2022). Studies using nuclear genome SNPs also show only small differences among Ananas types. A tree made from about 7.2 million SNPs shows clear groups within *A. comosus*, like the "Spanish" group, the "Queen/Cayenne" group, and different lines for wild *A. bracteatus* and A. microstachys (Feng et al., 2022). All Ananas types fall into a single group, and the species once called *Pseudananas sagenarius* is actually part of Ananas (Ouyang et al., 2022).

3.3 Phylogenetic placement of ananas within bromeliaceae and poales

Inside the subfamily Bromelioideae, Ananas stands out. Most members of this group live on trees (epiphytes), but Ananas grows on the ground and uses CAM photosynthesis strongly. Studies have confirmed that Bromelioideae is one group and that Ananas is closely related to Bromelia. These two likely share a common ancestor. Nuclear and chloroplast genes have helped show deeper family links within Bromeliaceae (Ouyang et al., 2022). At the level of the plant order Poales, full genome data shows that Bromeliaceae (and Ananas) split from the rest of the group before the grass family (Poaceae) did (Feng et al., 2022).

3.4 Reticulate evolution and hybridization in ananas

Although species within the genus Ananas are generally intercrossable, natural hybridization is limited due to geographic isolation. Cultivars such as ‘Smooth Cayenne’ and ‘Queen’ carry genome segments derived from multiple wild lineages, consistent with hybridization among distinct wild populations during domestication. The ‘Singapore Spanish’ cultivar, by contrast, is almost entirely derived from a single ancestral population. Chen et al. (2019) proposed a mixed domestication model: some cultivated groups were domesticated mainly through clonal propagation from a single origin (a one-step process), while others experienced a phase of sexual reproduction, during which genetic material from different lineages was introduced. Further evidence comes from the genome of the wild species *A. bracteatus* CB5, whose comparative analysis reveals a mosaic of genomes from three different Ananas species (Feng et al., 2022).

4 Evolutionary Trajectories of Key Traits in Pineapple

4.1 CAM photosynthesis and metabolic adaptation

Pineapple developed CAM photosynthesis in a way that is different from C4 grasses. In grasses, new enzymes came from gene duplications. In pineapple, existing genes changed how they are used. These changes happened through gene regulation. The pineapple genome has not gone through a recent whole-genome duplication. Important CAM enzymes like phosphoenolpyruvate carboxylase and malate dehydrogenase are mostly single-copy or belong to small gene groups. Many CAM genes in pineapple leaves turn on and off in a daily rhythm. This allows the plant to take in and fix CO_2 at night. Also, photosynthetic tissues (green leaves) and non-photosynthetic ones (white leaves) show different levels of CHH methylation, which is linked to CAM gene expression (Yow et al., 2023).

Because there is no recent genome duplication, the pineapple genome is simpler. It uses time-based gene activity to support CAM function. This helps the plant save water but leads to slower growth compared to C3 plants. CAM in pineapple evolved alongside other traits, like thick leaves with big vacuoles (for storing malate) and stomata that open at night.

4.2 Reproductive biology: flowering and compatibility

Pineapple has some unique reproductive traits. It often produces fruit without pollination (parthenocarpy). Wild types usually have flowers that can't self-pollinate, so they outcross. For example, *A. comosus* var. *bracteatus* is self-incompatible, but some cultivated types can self-pollinate. Chen et al. (2019) found four possible self-incompatibility (SI) genes in the wild genome F153. These genes are missing or not working in the cultivated variety CB5.

Pineapple does not flower at a steady time. Flowering can be triggered by cold or drought stress (Yow et al., 2023). Studies show that cold affects plant hormone signals, like ethylene and ABA. Genes related to vernalization and flowering (like FT-like and VIN3-like) act differently in various genotypes.

The B-box transcription factor gene AcBBX5 has been shown to promote flowering in pineapple (Ouyang et al., 2022). The AcBBX gene family in pineapple (19 members) displays some degree of diversification, though gene duplications are limited. Flowering time regulation can be achieved through expression changes in a few key regulators such as AcBBX5.

4.3 Fruit development and quality trait evolution

Pineapple fruit is made up of many flowers joined together. During domestication, fruit traits changed to meet human taste. Cultivated pineapples have more sugar than wild ones. Genes that manage sugar metabolism and transport were affected by selection. For example, areas in the genome with sucrose synthase and sugar transporter genes show signs of selection. These genes may help the fruit store more sugar (Chen et al., 2019).

In pineapple hybrids, researchers found QTLs linked to shell (rind) color. The gene AcCCD4, which breaks down carotenoids, is a likely candidate (Nashima et al., 2023). Some cultivars have special AcCCD4 alleles that slow down this breakdown. Another gene, AcMYB266, helps make the fruit peel red. Overexpression of this gene turns the peel red (Zhang et al., 2024).

4.4 Stress adaptation and defense mechanisms

CAM photosynthesis itself is a water-saving adaptation, but pineapple also possesses various defensive traits: tough, spiny leaves (to deter herbivores), high tissue concentrations of bromelain (a proteolytic enzyme), and metabolic responses to abiotic stress. *Cysteine proteases* such as bromelain and their inhibitors may help defend against pests and pathogens. In cultivated pineapple, a genomic region containing tandemly repeated bromelain inhibitor genes has undergone strong selective sweeps (Chen et al., 2019).

The WRKY gene family is important for plant stress responses. Pineapple has at least 54 WRKY genes (Wai et al., 2024). When AcWRKY28 is overexpressed, the plant tolerates salt stress better. But overexpressing AcWRKY31 makes the plant less tolerant to salt and drought. Still, it improves resistance to pineapple mealybug (Dysmicoccus brevipes).

5 Genomic Duplication and Structural Variation in Pineapple

5.1 Ancestral whole-genome duplication and karyotype conservation

Pineapple did not go through recent whole-genome duplication (WGD), unlike many other flowering plants. It comes from a monocot ancestor that had seven original chromosomes. Grasses went through more WGD events, but pineapple did not. As a result, its genome stayed more stable and shows less duplication. Pineapple split from grasses before the “rho” WGD event. This helped it avoid extra genome doubling and kept its chromosome number steady at 25 for millions of years. Because pineapple did not become polyploid in recent times, its genome changed more slowly. It also avoided problems like subgenome dominance or losing too many genes. The fact that it is clearly diploid makes it easier to build genetic maps and find genes linked to traits. Fewer duplicate genes also mean it is easier to see which genes cause which traits.

5.2 Structural variations and genome structural evolution

High-resolution comparisons between different Ananas genomes (cultivated and wild species) have revealed differences such as insertions, deletions, inversions, and copy number variations (CNVs). A comparison between the genome of *A. bracteatus* CB5 and *A. comosus* F153 showed high overall genome synteny, but also identified several large inversions and translocations (Feng et al., 2022). In a comparison between the ornamental red pineapple (*A. bracteatus* tricolor variety GL1) and the wild type CB5, researchers found two genomic fragments present in CB5 but absent in GL1, one of which (~1.1 kb) contains a photosynthetic protein gene.

The MD2 pineapple genome was improved using haplotype phasing. This revealed many structural variations even within this single cultivated type (Yow et al., 2022). Most of these were small insertions, deletions, or small CNVs. These differences affect gene copies and may lead to trait diversity. One example is the duplication of the AcTI gene, which codes for a trypsin inhibitor. Cultivated pineapple has more copies of this gene than wild types (Chen et al., 2019).

5.3 Transposable elements and genome dynamics

In the MD2 pineapple genome, about 45% of the genome is made up of long terminal repeat (LTR) retrotransposons (Feng et al., 2022). In *A. bracteatus*, the Gypsy-type LTR retrotransposons are the most common, making up around 44.8%. Pineapple had a burst of transposon activity in the late Pleistocene. Many intact LTR elements in CB5 and GL1 appeared about 1.7 to 2.1 million years ago.

The genome of cultivated *A. comosus* (F153) has fewer recent LTRs. This could be because repeat regions were not fully assembled or because domestication reduced transposon diversity. In some pineapple types, a DNA transposon from the CACTA family inserted before the *AcFT* gene, which may affect when the plant flowers. Pineapple uses DNA methylation to keep transposons silent. This is especially important because most pineapples are not grown from seeds, so harmful insertions are not easily removed. Studies on pineapple leaves found high CHH methylation levels, which shows active transposon silencing. Both green and white tissues have this pattern, though some differences were found (Shi et al., 2021).

5.4 Gene families and copy number variations

Pineapple has 43 *GRAS* genes, which help control growth and signals. These genes are spread across 19 of the 25 chromosomes (Lin et al., 2024). Some of them formed through tandem duplication on the same chromosome. When comparing with other plants, scientists saw that some GRAS subfamilies, like SHR and PAT1, had their own duplication patterns in pineapple.

Some gene families are reduced in copy number in pineapple-for example, the photoreceptor family is smaller compared to C3 grasses, aligning with pineapple's adaptation to sunlit, open environments (which reduces the need for shade responses).

On chromosome 3, a gene related to fruit protease has different copy numbers in different pineapple lines. Some have one copy, and others have two, which could affect how much protease the fruit makes. More genome sequencing has found many other CNVs, including in genes tied to disease resistance (like NBS-LRR), secondary metabolism, and development. Because pineapples are mostly grown from shoots and not seeds, rare gene duplications can stay in the same line for a long time.

6 Molecular Evolution of Key Gene Families

6.1 Photosynthesis-related gene families and the evolution of CAM photosynthesis

CAM photosynthesis in pineapple evolved in a different way compared to some C4 plants. In C4 plants, genes like PEPC or *Rubisco activase* often expand a lot. But in pineapple, this did not happen. The PEPC gene family, which helps fix CO_2 at night, is small. Only a few PEPC versions exist, and there's no sign that new ones evolved recently. Pineapple uses just one PEPC gene for CAM, and its activity is tightly controlled by the biological clock. The PEPC-activating kinase (PPCK) is also encoded by a single gene.

Another important group is aquaporins. These proteins help move CO_2 in leaf tissues. Pineapple has a moderate number of aquaporin genes in the PIP1 and PIP2 groups, and some work mainly at night. Genes that control stomata, like OST1 kinase and ABA receptors, did not expand much, but their activity patterns fit well with CAM.

The bHLH transcription factor family is large in pineapple, with about 121 genes (Hu et al., 2021). Some of these might help control stomata at night. In pineapple, CAM likely formed through small changes in when common photosynthesis genes are active. This was helped by promoter changes and new roles for transcription factors from the bHLH and NAC families.

6.2 Gene families related to flowering regulation and development

Pineapple flowers in a unique way. It needs a long juvenile phase and often responds to ethylene to flower. Flowering in pineapple is controlled by genes seen in other plants, but it also has its own features. Main regulators include MADS-box genes, FT-like genes, and hormone-related genes. Scientists found 43 MADS-box genes in pineapple-32 type II (MIKC-type) and 11 type I. These genes match all major groups seen in flowering plants. But unlike grasses or orchids, pineapple doesn't have many extra *MADS* genes.

Pineapple has only one SEP gene. It works during both flower and fruit development. This shows it helps with organ fusion and ripening. Genes like AcAP1 and AcAG follow the "ABC model" and are active in sepals, petals, stamens, and carpels. Pineapple flowers are small and packed together. The thick flower stem that forms the fruit may be due to continued action of genes like AcFUL and AcAGL11 during development.

CONSTANS (CO) and FT genes help plants track day length. Pineapple has at least two FT-like genes. One reacts to drought and stress, which may cause flowering during tough times (Yow et al., 2023). Pineapple also has 19 B-box (BBX) transcription factor genes (Ouyang et al., 2022). Some are like CONSTANS (COL) and help respond to light. AcBBX5 is one such gene. If you increase its expression, the plant flowers faster; if you reduce it, flowering is delayed.

6.3 Stress resistance gene families (NBS-LRR, WRKY, NAC)

Pineapple has about 80~90 *NBS-LRR* genes, which help fight disease. This number is lower than in many seed-grown crops, which often have hundreds. This may be because pineapple mostly reproduces asexually, so it doesn't need as many defense genes. There's no clear sign that these genes expanded a lot recently. Usually, there is one gene per lineage.

Some pineapple varieties show different numbers of these genes. For example, one gene on chromosome group LG12 appears several times in resistant types but only once in sensitive ones. WRKY transcription factors also play a role in stress response. Pineapple has around 54 *WRKY* genes (Wai et al., 2024), spread across groups I, II (a-e), and III. These genes have different roles. For example, AcWRKY31 (Group III) helps defend against pests, while AcWRKY28 (Group II) helps with heat or drought. The NAC gene family has about 73 members. These genes help control stress response, plant growth, and aging.

6.4 Metabolic pathway genes and secondary metabolites

Pineapple makes proteases, which are enzymes that break down proteins. These come from a small group of C1A peptidase genes. They fall into two types: fruit-type and stem-type. The fruit-type (e.g., AcCP1) is from a single gene. The stem-type likely came from gene duplications and shows different traits, like isoelectric points.

Pineapple also has protease inhibitors like cystatins. These are encoded by a few genes, though some duplicated over time. This shows an evolutionary balance between proteases and inhibitors (Chen et al., 2019). Pineapple's smell comes from esters, terpenes, and aldehydes. These are made by gene families like acetyltransferases, terpene synthases, and alcohol dehydrogenases. The BAHD acetyltransferase family is active in ripe fruit. One member, AcAAT, helps make key esters like ethyl butyrate. Pineapple has about a dozen TPS (terpene synthase) genes. Some of them evolved under positive selection, which may increase aroma diversity.

Pineapple has key genes for anthocyanin pigment production (e.g., DFR, ANS, UFGT), usually one copy each. But regulators like AcMYB266 can change skin color from green to red by changing gene activity or mutating (Zhang et al., 2024). Enzymes like ACC synthase and ACC oxidase, which make ethylene, have evolved under pressure to meet flowering and ripening needs in pineapple.

7 Case Study: Domestication and Genetic Improvement of Pineapple

7.1 Origin and domestication of pineapple

Pineapple was domesticated by the indigenous peoples of tropical America before the time of Columbus, and its domestication process appears to have been a "single-step" event (Chen et al., 2019). Wild populations of Ananas in the Paraguay-Paraná river basin are considered the ancestors of cultivated pineapple (*A. comosus*). Cultivated pineapple exhibits characteristics of reduced genetic diversity, consistent with a model of a single domestication origin followed by clonal propagation. Most cultivated varieties share a common genomic foundation, particularly those in the "Spanish group," which are almost entirely derived from a single wild lineage (supporting the single-step domestication hypothesis). In fact, some varietal groups, such as "Smooth Cayenne" and "Queen," carry genomic segments from other wild pineapple populations.

Traits selected during domestication include: increased fruit size, high sugar content, reduced flesh fiber, development of seedless fruit, and the loss of sexual reproduction capability. The domestication process achieved a transition from sexual reproduction to complete reliance on clonal propagation through suckers and slips.

7.2 Genomic features selected in cultivated pineapple

In cultivated pineapple, genetic diversity is lower than in wild types. Some parts of the genome are very different from wild alleles. Chen et al. (2019) found 25 possible domestication sweep regions in the pineapple genome. These regions range in size from 150 kb to 1.2 Mb. The strongest sweep is on chromosome 5. It includes several genes that control fruit growth and fiber traits. Another important region is on chromosome 10. It has a gene for sugar transport and another gene that helps control how the fruit grows (Figure 2).

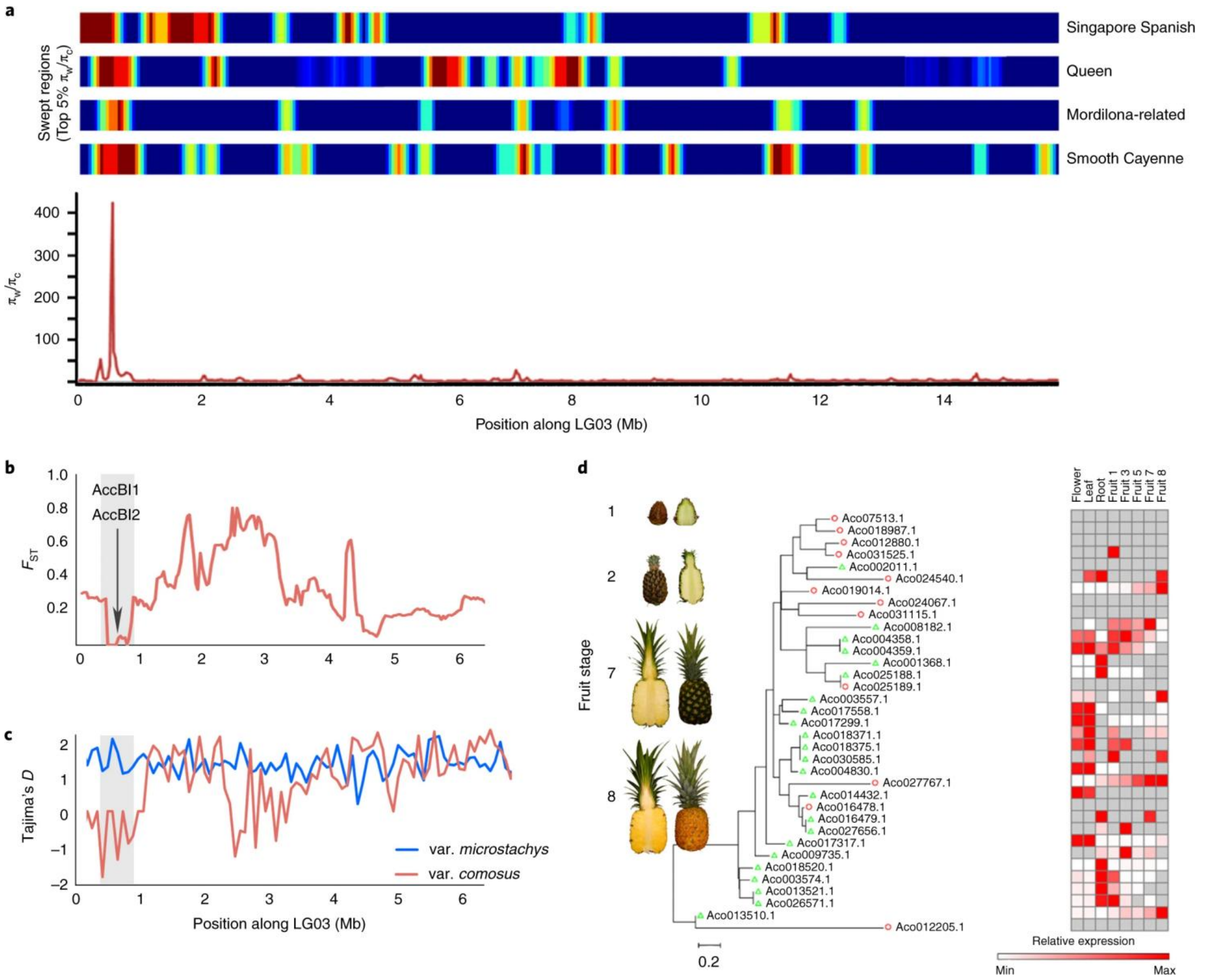

Figure 2 Putative domestication sweep around a bromelain inhibitor gene that helps control fruit ripening (Adopted from Chen et al., 2019)

Image caption: a: top: heat maps showing the distribution of domestication sweeps (top 5% π_w/π_c) for the four cultivars; Bottom: a putative swept region at the end of LG03 containing *AccBI1* and *AccBI2*. π_w/π_c across all cultivars is plotted using a sliding window of 0.5 Mb with 0.1 Mb shift; b: genetic distance (F_{ST}) between the smooth cayenne, queen, singapore spanish and mordilona-related clusters for the 6.5 Mb of LG03. Mean F_{ST} values are plotted in sliding windows of 50 kb with 25 kb step size; c: Tajima's D values for the four combined cultivar clusters (var. *comosus*) and wild (var. *microstachys*); Mean Tajima's D values are plotted in sliding windows of 50 kb with a step size of 25 kb; d: left: pineapple fruit at select stages from a fruit ripening series (stages 1, 2, 7 and 8); right: maximum likelihood phylogeny of bromelain genes with log2 transformed RPKMs of expression in fruit, flower, leaf and root tissue (Adopted from Chen et al., 2019)

Cultivated pineapples have accumulated numerous nonsynonymous mutations in genes related to disease resistance, consistent with the loss of trait-specific adaptive selection pressures. A typical feature of asexually propagated crop genomes is the accumulation of deleterious mutations (genetic load), due to the absence of purifying effects of sexual recombination in each generation. In pineapple, a large number of mildly deleterious variants are retained in a heterozygous state (Yow et al., 2022). The domestication process by humans may have driven these mutations to drift toward higher frequencies. Currently, more than 48% of the genes in the pineapple genome are observed to carry putatively heterozygous deleterious variants.

7.3 Breeding and genetic improvement of modern pineapple

In the past, pineapple breeding mainly relied on picking parent plants, doing hand pollination, collecting seeds, and then testing the new plants made from cuttings. This whole process usually takes many years. Now, with the pineapple genome and many SNP markers available, scientists can use genetic tools to speed up the work. Researchers have used an F1 mapping population to find important regions (QTLs) in the genome linked to traits like fruit weight, sugar level (measured by Brix), sourness, and flesh color (Nashima et al., 2023). One major QTL that controls peel color was found in a gene called *AcCCD4*. This gene helps break down pigments, so it affects whether the peel looks yellow or orange. Another QTL related to vitamin C was found on chromosome 7. It includes a gene that helps make vitamin C, called a galacturonate reductase. A separate GWAS study looked at 89 pineapple types, including wild and farmed ones. This study found DNA markers linked to traits like plant height and when the plant flowers (Nashima et al., 2024).

The company Del Monte used genetic engineering to make a pink-fleshed pineapple. They did this by turning off the *AcCCD4* gene, which lets lycopene (a pink pigment) build up in the fruit. Another gene, *AcMYB266*, was found to control red peel color. Breeders can use this gene to cross red-skinned ornamental types with common yellow-fleshed ones. Then, they can use DNA markers to pick out new plants that have both red skin and good fruit traits (Zhang et al., 2024).

8 Implications for Future Phylogenomics and Evolutionary Research

8.1 Pineapple as a model for phylogenomics of monocotyledonous plants

Pineapple occupies a key phylogenetic position as an outgroup to commelinid monocots such as grasses and palms (Xu et al., 2018). Its genome serves as a valuable reference point for comparative analyses aimed at inferring the ancestral genome content and chromosomal structure of monocots. The Ananas genome has already been used to identify conserved non-coding sequences across monocots.

Due to its relatively small genome and extensive functional annotations (e.g., the PGD database), pineapple is gradually becoming an important model for evolutionary genomics of monocots, following rice and banana. Future studies may include sequencing more Bromeliaceae genomes (some projects are already underway, such as the *Aechmea fasciata* genome), which will uncover the genomic basis of epiphytism, CAM (Crassulacean Acid Metabolism) evolution, and other lineage-specific innovations in Bromeliaceae (Li et al., 2022).

8.2 Genetic diversity and conservation of wild ananas species

Cultivated pineapple includes only a small part of the genetic diversity found in the Ananas genus (Chen et al., 2019). Some wild species like *A. macrodontes* (previously called Pseudananas) grow in tough environments such as dry savannas. These plants may carry useful genes that help them survive heat and drought. Another wild species, *A. bracteatus*, is known for its strong leaf fibers and its resistance to certain soil diseases. These traits could be helpful in breeding pineapples that produce both fruit and fiber, or in improving root strength.

By sequencing and genotyping wild species, scientists can find special gene variants and structural differences. For example, if a wild plant has a gene that naturally protects it from pineapple mealybug wilt-associated virus, breeders can try to introduce that gene into new pineapple varieties. Conservation genomics can also help explain how wild Ananas species are related and how their genes move between populations. Feng et al. (2022) found that *A. bracteatus* CB5 has genes from different sources, so it’s important to protect more than one group of this species.

Gene editing can be used to improve wild pineapples that resist disease but have small fruit, such as A. ernestii. By changing just a few key genes, scientists can make these wild plants produce sweeter or larger fruits, creating new pineapple-like crops.

8.3 Integrated omics and systems biology studies in pineapple

Combining DNA methylation maps with gene expression data can help scientists understand how genes control complex traits like CAM metabolism and flowering. Researchers have already used chromatin accessibility methods, such as DNase-seq or ATAC-seq, on pineapple leaves. These tests give clues about which genes are turned on or off by the plant's internal clock. In the future, the same methods can be used to study fruit growth, stress response, and other traits (Sharma et al., 2017). Some important traits, like sugar buildup or the start of flowering, are controlled by transcription factors and DNA patterns. Looking at these elements together will help us understand them better.

Pineapple fruit development can be viewed as a network output involving interactions among genes (transcription factors, enzymes) and metabolites (sugars, phytohormones, secondary metabolites). With the aid of computational modeling, one can predict how changes in specific parts of the system (e.g., upregulation of sugar transporter expression) influence the whole system (sugar levels, osmotic balance, and even cross-regulation of ethylene signaling) (Chomthong and Griffiths, 2023).

Acknowledgments

Sincerely thank the reviewers for their constructive criticisms and suggestions during the review process.

Conflict of Interest Disclosure

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

References

Chen L.Y., VanBuren R., Paris M., Zhou H.Y., Zhang X.T., Wai C.M., Yan H.S., Chen S., Alonge M., Ramakrishnan S., Liao Z.Y., Liu J., Lin J.S., Yue J.J., Fatima M., Lin Z.C., Zhang J.S., Huang L.X., Wang H., Hwa T.Y., Kao S.M., Choi J.Y., Sharma A., Song J., Wang L.L., Yim W.C., Cushman J.C., Paull R.E., Matsumoto T., Qin Y., Wu Q.S., Wang J.P., Yu Q.Y., Wu J., Zhang S.L., Boches P., Tung C.W., Wang M.L., d'Eeckenbrugge G.C., Sanewski G.M., Purugganan M.D., Schatz M.C., Bennetzen J.L., Lexer C., and Ming R., 2019, The bracteatus pineapple genome and domestication of clonally propagated crops, Nature Genetics, 51(10): 1549-1558.
https://doi.org/10.1038/s41588-019-0506-8
PMid:31570895

Chomthong M., and Griffiths H., 2023, Prospects and perspectives: inferring physiological and regulatory targets for CAM from molecular and modelling approaches, Annals of Botany, 132(4): 583-596.
https://doi.org/10.1093/aob/mcad142
PMid:37742290 PMCid:PMC10799989

Feng L.J., Wang J.T., Mao M.Q., Yang W., Adje M.O., Xue Y.B., Zhou X.Z.X., Zhang H.L., Luo J.H., Tang R.M., Tan L., Lin D.P., Zhang X.P., Zang Y.Q., He Y.H., Chen C.M., Luan A.P., Lin W.Q., Xu W.T., Li X., Sun L.X., Jiang F.X., and Ma J., 2022, The highly continuous reference genome of a leaf-chimeric red pineapple (*Ananas comosus* var. *bracteatus*) provides insights into elaboration of leaf color, G3: Genes, Genomes, Genetics, 12(2): jkab452.
https://doi.org/10.1093/g3journal/jkab452
PMid:35100332 PMCid:PMC8824783

Hu J., Chang X., Zhang Y., Yu X., Qin Y., Sun Y., and Zhang L., 2021, The pineapple MADS-box gene family and the evolution of early monocot flower, Scientific Reports, 11(1): 849.
https://doi.org/10.1038/s41598-020-79163-8
PMid:33441609 PMCid:PMC7806820

Li Z., Wang J., Zhang X., Zhu G., Fu Y., Jing Y., Huang B., Wang X., Meng C., Yang Q., and Xu L., 2022, The genome of *Aechmea fasciata* provides insights into the evolution of tank epiphytic habits and ethylene-induced flowering, Communications Biology, 5(1): 920.
https://doi.org/10.1038/s42003-022-03918-4
PMid:36071139 PMCid:PMC9452560

Lin J.T., Wu J.H., Zhang D., Cai X.K., Du L.M., Lu L., Liu C.J., Chen S.C., Yao Q.L., Xie S.Y., Xu X.W., Wang X.M., Liu R.Y., Qin Y., and Zheng P., 2024, The GRAS gene family and its roles in pineapple (*Ananas comosus* L.) developmental regulation and cold tolerance, BMC Plant Biology, 24(1): 1204.
https://doi.org/10.1186/s12870-024-05913-9
PMid:39701971 PMCid:PMC11657692

Liu C., Zhang W., and He Y., 2022, The complete chloroplast genome of *Ananas comosus* var. *erectifolius* (LB Smith) Coppens & Leal, Mitochondrial Dna Part B, 7(3): 431-433.
https://doi.org/10.1080/23802359.2022.2039081
PMid:35252578 PMCid:PMC8896190

Nashima K., Omine Y., Shirasawa K., Sato T., Yamada M., Shoda M., and Takeuchi M., 2024, Genome-wide association study of pineapple breeding population, Scientia Horticulturae, 338: 113757.
https://doi.org/10.1016/j.scienta.2024.113757

Nashima K., Takeuchi M., Moromizato C., Omine Y., Shoda M., Urasaki N., Tarora K., Irei A., Shirasawa K., Yamada M., Kunihisa M., Nishitani C., and Yamamoto T., 2023, Identification of quantitative trait loci of fruit quality and color in pineapples, The Horticulture Journal, 92(4): 375-383.
https://doi.org/10.2503/hortj.QH-063

Ouyang Y.W., Pan X.L., Wei Y.Z., Wang J., Xu X.X., He Y.K., Zhang X.H., Li Z.Q., and Zhang H.G., 2022, Genome-wide identification and characterization of the BBX gene family in pineapple reveals that candidate genes are involved in floral induction and flowering, Genomics, 114(4): 110397.
https://doi.org/10.1016/j.ygeno.2022.110397
PMid:35675877

Sharma A., Wai C.M., Ming R., and Yu Q., 2017, Diurnal cycling transcription factors of pineapple revealed by genome-wide annotation and global transcriptomic analysis, Genome Biology and Evolution, 9(9): 2170-2190.
https://doi.org/10.1093/gbe/evx161
PMid:28922793 PMCid:PMC5737478

Shi Y., Zhang X., Chang X., Yan M., Zhao H., Qin Y., and Wang H., 2021, Integrated analysis of DNA methylome and transcriptome reveals epigenetic regulation of CAM photosynthesis in pineapple, BMC Plant Biology, 21: 1-14.
https://doi.org/10.1186/s12870-020-02814-5
PMid:33407144 PMCid:PMC7789485

Wai M.H., Luo T., Priyadarshani S.V.G.N., Zhou Q., Mohammadi M.A., Cheng H., Aslam M., Liu C., Chai G.F., Huang D.P., Liu Y.H., Cai H.Y., Wang X.M., Qin Y., and Wang L.L., 2024, Overexpression of AcWRKY31 increases sensitivity to salt and drought and improves tolerance to mealybugs in pineapple, Plants, 13(13): 1850.
https://doi.org/10.3390/plants13131850
PMid:38999690 PMCid:PMC11243833

Xu H.M., Yu Q.Y., Shi Y., Hua X.T., Tang H.B., Yang L., Ming R., and Zhang J., 2018, PGD: pineapple genomics database, Horticulture Research, 5: 66.
https://doi.org/10.1038/s41438-018-0078-2
PMid:30245835 PMCid:PMC6139296

Xu R., Wei F., Chen Y., Khan F.S., Wei Y., and Zhang H., 2024, Identification, phylogeny, and expression profiling of pineapple heat shock proteins (HSP70) under various abiotic stresses, International Journal of Molecular Sciences, 25(24): 13407.
https://doi.org/10.3390/ijms252413407
PMid:39769172 PMCid:PMC11678451

Yow A.G., Bostan H., Castanera R., Ruggieri V., Mengist M.F., Curaba J., Young R., Gillitt N., and Iorizzo M., 2022, Improved high-quality genome assembly and annotation of pineapple (*Ananas comosus*) cultivar MD2 revealed extensive haplotype diversity and diversified FRS/FRF gene family, Genes, 13: 52.
https://doi.org/10.3390/genes13010052
PMid:35052394 PMCid:PMC8774480

Yow A.G., Laosuntisuk K., Young R.A., Doherty C.J., Gillitt N., Perkins-Veazie P., Xiang Q.Y.J., and Massimo Iorizzo M., 2023, Comparative transcriptome analysis reveals candidate genes for cold stress response and early flowering in pineapple (*Ananas comosus*), Scientific Reports, 13: 18890.
https://doi.org/10.1038/s41598-023-45722-y
PMid:37919298 PMCid:PMC10622448

Zhang W., Wu J., He J.H., Liu C.Y., Yi W., Xie J.Y., Wu Y., Xie T., Zhong Z.Q., Yang M.Z., Chen C.J., Luan A.P., and He Y.H., 2024, AcMYB266, a key regulator of the red coloration in pineapple peel: a case of subfunctionalization in tandem duplicated genes, Horticulture Research, 11(6): uhae116.
https://doi.org/10.1093/hr/uhae116
PMid:38919552 PMCid:PMC11197299