Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening

Peitao Lü[1,11], Sheng Yu[1,11], Ning Zhu[1,11], Yun-Ru Chen[1,11], Biyan Zhou[2,11], Yu Pan[3], David Tzeng[1], Joao Paulo Fabi[4], Jason Argyris[5], Jordi Garcia-Mas[5], Nenghui Ye[6], Jianhua Zhang[6], Donald Grierson[7,8], Jenny Xiang[9], Zhangjun Fei[10], James Giovannoni[10] and Silin Zhong[1*]

[1]State Key Laboratory of Agrobiotechnology, School of Life Sciences, Chinese University of Hong Kong, Hong Kong, China.
[2]College of Horticulture, South China Agricultural University, Guangzhou, China.
[3]College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China.
[4]Department of Food Science and Experimental Nutrition, FCF, University of Sao Paulo, Sao Paulo, Brazil.
[5]IRTA, Centre for Research in Agricultural Genomics, Barcelona, Spain.
[6]Department of Biology, Hong Kong Baptist University, Hong Kong, China.
[7]School of Crop Sciences, University of Nottingham, Nottingham, UK.
[8]Zhejiang Provincial Key Laboratory of Horticultural Plant Integrative Biology, Zhejiang University, Hangzhou, China.
[9]Weill Medical College, Cornell University, New York, NY, USA.
[10]US Department of Agriculture—Agricultural Research Service and Boyce Thompson Institute for Plant Research, Cornell University, Ithaca, NY, USA.

[11]These authors contributed equally.
[*]e-mail: silin.zhong@cuhk.edu.hk

Abstract

Fleshy fruits using ethylene to regulate ripening have developed multiple times in the history of angiosperms, presenting a clear case of convergent evolution whose molecular basis remains largely unknown. Analysis of the fruitENCODE data consistint of 361 transcriptome, 71 accessible chromatin, 147 histone and 45 DNA methylation profiles reveals three types of transcriptional feedback circuits controlling ethylene-dependent fruit ripening. These circuits are evolved from senescence orfloral organ pathways in the ancestral angiosperms either by neofunctionalisation or repurposing pre-existing genes. The epigenome, H3K27me3 in particular, has played a conserved role in restricting ripening genes and their orthologues in dry and ethylene-independent fleshy fruits. Our findings suggest that evolution of ripening is constrained by limited hormone molecules and genetic and epigenetic materials, and whole-genome duplications have provided opportunities for plants to successfully circumvent these limitations.

Keywords: ENCODE, fleshy fruit; ripening, ethylene, convergent evolution, genome duplication, senescence.

Introduction

Angiosperms are the largest and most diverse group of land plants. Unlike gymnosperms, seeds of angiosperms are enclosed and protected by a structure called a fruit, which is differentiated from the ovary or its surrounding floral tissues. Fruits can be classified as dry or fleshy, and the more ancient dry fruits or their seeds are adapted for dispersal by mechanical expulsion, wind and by attaching to the fur of animals[1]. The development of fleshy fruits enabled angiosperms to interact with coevolving animals (frugivores), which consumed the fruits and dispersed the seeds to different locations, thus enhancing distribution, minimizing parental competition and increasing plant reproductive success[2].

Many fleshy fruits, such as apples, bananas and tomatoes are climacteric, where a respiratory burst occurs at the commencement of ripening as a prelude to the molecular changes that alter fruit colour, flavour, texture, aroma and nutritional properties. Despite having evolved independently, climacteric fruits use the same plant hormone ethylene as a ripening signal[3,4]. These climacteric fruits are often harvested unripe, stored and treated with ethylene to complete maturation. Too much ethylene, on the other hand, leads to rapid deterioration of the fruit. Hence, controlling ethylene synthesis or signalling is of great practical importance during post-harvest storage, shipping and for maintaining shelf-life and quality.

Climacteric fruits also evolved a mechanism to synthesize ethylene in an autocatalytic manner, which is historically referred to as system II ethylene to distinguish it from the self-inhibitory system I ethylene in other tissues, such as immature fruit and leaves[3,4]. Its autocatalytic nature suggests a positive feedback loop controlling ethylene synthesis during ripening. Although extensive studies have identified isoforms of ethylene biosynthesis genes that are specifically required for system I and II ethylene production, their regulation remains largely unknown[5,6].

In tomato, the most studied fruit model, a series of transcription factors such as COLOURLESS NON-RIPENING (CNR), NONRIPENING (NOR) and RIPENING INHIBITOR (RIN) are required for ripening and autocatalytic ethylene synthesis[7,8]. In addition, genome-wide DNA hypomethylation is associated with tomato fruit development, and silencing the DNA demethylase *DML2* could delay ripening[9,10]. These

results suggest that the epigenome acts as a developmental switch to restrict the activities of ripening regulators before seed maturation. Although tomato is still the predominant model for fruit research, many fleshy fruit genomes have now been sequenced, raising the questions whether and to what degree the tomato model is universal. In addition, tomato has experienced whole-genome duplication (WGD) ~ 71 Myr, and key ripening regulators including RIN and ethyene biosynthesis genes are paralog members of duplicated gene families[11]. Hence, plants without WGD or demethylase expressed during ripening might have evolved diferent regulatory systems.

It is difficult to resolve complex convergent traits, such as ripening, in diverse taxa by sequencing and comparing genomes if the convergence occurred through the evolution of different genes and pathways, or if the genes are the same but the *cis*-regulatory elements or points of epigenetic regulation are different. To address these questions, we used an ENCODE-style functional genòmic approach to systematically characterize the molecular circuits controlling ripening in multiple plant species. We found three major types of transcriptional circuits controlling climacteric fleshy fruit ripening (Fig. 1). Eudicots with recent WGD utilized their duplicated MADS transcription factors to form the ripening circuits, while those without WGD used carpel senescence-related NAC transcription factors. The monocot plant banana also experienced recent WGD and uses both MADS and NAC genes to form two interconnected circuits. We also found that the ripening genes, as well as their epigenetic marks restricting their expression, are conserved in their orthologues in non-climacteric fruits and even dry fruits, suggesting that these independently evolved ripening mechanisms are originated from pre-existing pathways that served different functions in the ancestral angiosperms.

Results

The fruitENCODE data. The fruitENCODE project aims to generate a comprehensive annotation of functional elements in seven climacteric fruit species (apple, banana, melon, papaya, peach, pear and tomato) with sequenced reference genomes. Four non-climacteric fleshy fruit species (cucumber, grape, strawberry and watermelon) and two dry fruit plants (Arabidopsis and rice) were also included for comparative analysis. To construct a multidimensional dataset for fleshy fruit functional genomics, we have used wholegenome bisulfite sequencing (WGBS), ChIP-Seq, DNaseI-Seq and RNA-Seq to

profile their tissue-specific DNA methylation, histone modifications, accessible chromatin and transcriptome profiles, respectively (Supplementary Tables 1–34).

We have also included a large collection of mutants with altered ripening phenotypes. For other non-model species such as melon, which lack large mutant collections, we have included four varieties with different ripening characteristics for comparative analysis. The current dataset encompassed 361 transcriptome, 71 accessible chromatin, 147 histone modification and 45 methylome profiles. All processed datasets can be accessed from the fruitENCODE data base (www.epigenome.cuhk.edu.hk/encode.html).

Using the fruitENCODE data, we sought to clarify the regulatory circuits controlling climacteric fruit ripening. We first identified transcription factors and ethylene biosynthetic genes expressed during ripening. The accessible chromatin dataset enabled us to identify their cis-regulatory elements and candidate transcription factors. We could then validate these regulatory interactions by performing transcription factor ChIP-Seq and promoter activation assays. In view of the difficulty in carrying out genetic assays in crops, particularly fruit trees, we developed a heterologous tobacco system involving ectopic expression of gene components under their native promoters from all seven climacteric species to recreate the autocatalytic ethylene symptomatic of climacteric ripening.

MADS-type positive feedback loop controlling ripening in tomato, apple and pear. We first reconstructed the ripening regulatory circuit for the model tomato fruit, which has three known components: ethylene, transcription factor RIN and DNA methylation. From the DNaseI hypersensitive sites (DHS) dataset, we found an EIN3 binding motif in the promoter of RIN, the functional significance of which was confirmed by our EIN3 ChIP-Seq (Fig. 2a and Supplementary Table 32). RIN is a MADS-box transcription factor, which functions in a multimeric complex with TAGL1. TAGL1 is expressed during both early and late fruit development[12]. We performed ChIP-Seq for both RIN and TAGL1, and found that they can directly target the ripening ethylene biosynthesis genes ACC SYNTHASE2 (ACS2) and ACC OXIDASE1 (ACO1) (Fig. 2a).

Our findings suggest that ethylene transcription factor EIN3 and MADS-box transcription factors RIN–TAGL1 could form a positive feedback loop to synthesize autocatalytic

system II ethylene, while the downstream ripening genes are coupled to the loop through RIN–TAGL1 (Fig. 1a). Our ChIP-Seq data confirmed that the RIN–TAGL1 complex targets well-known ripening genes that are involved in fruit softening, colour change, aroma production and sugar metabolism (Supplementary Tables 30 and 31). Given the central role of the MADS genes in this ripening model, we named it
the MADS positive feedback loop.

To test this feedback loop, we attempted to recreate the autocatalytic ethylene synthesis in tobacco leaf by expressing the core tomato loop genes. It is well known that the ethylene synthesis and signalling pathways are conserved in plants[4,13]. Leaf has endogenous EIN3 and ACO, and lacks ACS activity, which is the rate-limiting enzyme In its ethylene biosynthesis under normal growth conditions. When we expressed the tomato RIN and ACS2 using their native promoters, and TAGL1 was supplied under a constitutive CaMV35S promoter, spontaneous ethylene synthesis was observed (Fig. 3a and Supplementary Fig. 1). We could then mutate the EIN3 binding motif in the RIN promoter and the RIN binding motif in the ACS2 promoter, both of which disrupted the spontaneous ethylene synthesis (Fig. 3a and Supplementary Fig. 1c). These results confirmed that the exogenous tomato ACS2 and MADS genes are responsable for the observed ethylene burst. We also treated the tobacco leaf expressing the loop with ethylene inhibitor 1-MCP, which causes degradation of the EIN3 protein[14]. We found that 1-MCP blocked the ethylene synthesis, suggesting that the ethylene generated by the tomato MADS loop in tobacco leaf is indeed autocatalytic, a key characteristic of the system II ethylene produced by ripening climacteric fruits (Fig. 3a and Supplementary Fig. 1b).

However, ethylene is a stress hormone and such an autocatalytic feedback loop involving a diffusible signal molecule poses a major threat to the plant itself and its neighbours because any leakiness could cause developmental perturbations, including senescence and tissue death (Supplementary Fig. 1f). Our DHS and epigenome datasets showed that the EIN3 binding site in the RIN promoter is demethylated and becomes accessible only in ripening fruit tissues, while RIN and ACS2 are associated with the repressive histone mark H3K27me3 in leaf and immature fruit (Fig. 2a).

We examined the fruitENCODE data from other species and found that only apple and pear have similar MADS-type positive feedback loop. Their MADS promoters contain EIN3 binding motifs, while their ethylene biosynthesis genes have MADS binding motifs (Supplementary Figs. 2 and 3). We also used the tobacco system to confirm that the apple and pear loops were capable of generating autocatalytic ethylene. Interestingly, apple and pear shared a recent WGD[15], and their ripening MADS genes targeted by EIN3 are orthologues. Both of them are members of the duplicated MIKCc MADS transcription factor family and are homologues of the tomato RIN (Supplementary Fig. 4). These suggest that similar ripening circuits have idependently evolved through neofunctionalization of the duplicated MADS genes originally controlling floral organ identity in the ancestral angiosperms.

Tomato is the only reported species to activate DNA demethylase expression during fruit ripening[10]. Our methylome datasets confirmed that apple and pear lack a tomato-like whole-genome demethylation during fruit development (Supplementary Fig. 19).
However, from the histone modification datasets, we found that tomato, apple and pear have the same tissue-specific H3K27me3 on their ethylene biosynthesis and MADS gene loci, whereas they are absent in the ripe fruit tissues (Fig. 2a, Supplementary Fig. 4). We also found that the ethylene-independent pear cultivar Dangshanshuli contains hyper-H3K27me3 in its ethylene biosynthesis gene loci compared to the ethylene-dependent cultivar Williams (Supplementary Fig. 3). The non-ripening tomato mutant nor also contains hyper-H3K27me3 in the ACS2 and RIN loci (Supplementary Fig. 5). This suggests that instead of using DNA methylation, H3K27me3 could play a conserved role from preventing the MADS positive feedback loop from generating autocatalytic ethylene.

Peach, papaya and melon operate a NAC positive feedback loop. Climacteric fruits, such as peach, papaya and some climacteric melon cultivars, can also produce and require autocatalytic ethylene for ripening but, unlike the MADS-type fruits, did not undergo recent WGD[15,16]. To reconstruct their ripening circuits, we first examined what transcription factors could regulate their ACS and ACO genes during ripening. Interestingly, in their promoter DHS, we identified NAC instead of MADS transcription factor binding motifs. In addition, they all have NAC genes with ripening-specific expression pattern (Supplementary Tables 17, 23–25).

NAC is one of the largest plant-specific transcription factor families, with members involved in many developmental processes such as senescence, stress, cell wall formation and embryo development. The peach, melon and papaya ripening-specific NAC genes we found are orthologues of the Arabidopsis carpel senescence-related transcription factors NARS1/2 and are distantly related to the leaf senescence-related AtNAP[17,18] (Supplementary Fig. 6). Examination of their NAC gene promoter DHS revealed EIN3 binding sites, suggesting that instead of neofunctionalization of the duplicated MADS genes, plants without WGD might have repurposed their carpel senescence NAC to generate a positive feedback loop with ethylene to regulate ripening (Fig. 1b).

To test this, we performed ChIP-Seq in ripening peach fruit tissues using an antibody against the NAC protein (ppa007577m), and found that it can bind to the ACS and ACO promoter (Fig. 2b). Next, we used the tobacco system to test whether they could form a positive feedback loop. We ectopically expressed the peach NAC and ACS (ppa004774m) genes under their native promoters and found that they were capable of generating ethylene spontaneously (Fig. 3b). We also performed EIN3 motif deletion to confirm that EIN3 binding to the *NAC* promoter is required. The ethylene synthesis could be blocked by treatment with 1-MCP, suggesting that the ethylene generated by the NAC positive feedback loop is autocatalytic.

In tomato, the downstream ripening genes are directly coupledto the MADS positive feedback loop through the RIN–TAGL1 transcription factors. Our ChIP-Seq data showed that the NAC transcription factor also binds to the promoter of key fruit ripening genes, such as those involved in pigment accumulation, volatile secondary metabolite production, cell wall softening and sugar accumulation (Supplementary Table 33).

Genes involved in the MADS-type loop found in tomato, apple and pear are associated with conserved H3K27me3 marks (Fig. 2a). Our epigenome data revealed similar tissue-specific H3K27me3 patterns in the peach NAC and ACS loci (Fig. 2b). In papaya and the climacteric melon cultivar Védrantais, we also found this NAC-type positive feedback loop with key genes associated with H3K27me3 in non-ripening tissues (Supplementary

Figs. 7 and 8). In the nonclimacteric melon cultivar Piel de Sapo, we found that NAC is downregulated and is associated with increased H3K27me3 level (Supplementary Fig. 8). These results suggest that H3K27me3 plays a conserved and perhaps central role in regulating both the MADS and NAC-type positive feedback loops that generate ripening ethylene in different plant species, despite having evolved independently.

Monocot banana operates a dual-loop system. Banana is also a climacteric fruit that requires autocatalytic ethylene to ripen and it has experienced three recent WGD[15,19]. The autocatalytic ethylene production in other climacteric fruits such as tomato can be interrupted by the ethylene-action inhibitor 1-MCP, a scenario that we reproduced in the heterologous tobacco system. However, a unique ripening feature in banana is that inhibitor treatment is unable to interrupt its ethylene production after ripening has been initiated, indicating a transition from autocatalytic to ethylene-independent ripening[20]. Examination of the banana data showed that it has two positive feedback loops, and the second one is able to maintain the ethylene synthesis when the first ethylene-dependent loop is blocked (Fig. 1c).

The first banana loop is similar to the NAC-type positive feedback loop in eudicots without WGD (Fig. 1c). The banana ACS (Ma04_t35640.1) and ACO (Ma07_t19730.1) have NAC motifs in their promoter DHS, while the NAC gene (Ma06_t33980.1) contains an EIN3 binding motif. To test the loop, we ectopically expressed the banana NAC and ACS genes under their native promoters in tobacco and found that they are sufficient to generate ethylene in an autocatalytic manner. Ethylene inhibitor 1-MCP, as well as EIN3 motif deletion, could block ethylene production in the absence of the second loop, suggesting that loop I is a functional NAC-type positive feedback loop (Fig. 3c).

It should be noted that this banana ripening NAC is an orthologue of the rice leaf senescence transcription factor OsNAP[21] and is distantly related to the carpel senescence-related NACs utilized by the eudicots climacteric fruits (Supplementary Fig. 6). We also profiled gene expression and histone modifications in the young, matured and aged banana leaves and found that this NAC is expressed during leaf senescence (Fig. 4). Consistently, the banana NAC and ACS genes are associated with tissue-specific H3K27me3 as those in the eudicots, except that the NAC locus lost H3K27me3

in the aged leaves (Fig. 4).

Both the eudicots MADS- and NAC-type positive feedback loops are directly coupled to the downstream ripening genes, which we confirmed by ChIP-Seq using tomato MADS and peach NAC as examples. In the absence of a suitable antibody against the banana NAC, we used the dual luciferase assay to show that it is capable of activating known ripening gene promoters (Supplementary Fig. 9), suggesting that banana ripening genes are likely to be coupled to the positive feedback loop, as in their eudicots counterparts.

It should be noted that the first loop alone could not explain how banana bypasses the ethylene dependence after ripening initiation[20]. It has been shown that three MADS transcription factors (MADS1/2/5) are expressed in banana fruit, where MADS1/2 have been further functionally characterized via transgenic repression resulting in delayed ripening[22,23]. We found a NAC motif in the banana MADS1 gene promoter and a MADS motif in its NAC gene promoter, suggesting that the MADS and NAC genes could form a second positive feedback loop to bypass the first loop (Fig. 1c).

To validate the second loop, we first used the tobacco system to show that the first NAC-type loop could be blocked by inhibidor 1-MCP treatment (Fig. 3c). Co-expressing the three MADS genes from the second loop with MADS1 driven by its native promoter and MADS2/5 with the constitutive 35S promoter enabled the
tobacco leaves to synthesize more ethylene than expressing loop 1 alone (54.27%, $P = 2.91 \times 10{-}5$). Most importantly, 1-MCP was unable to block the ethylene production when the second loop was present, mimicking the behaviour of the ripening banana fruit (Fig. 3c). To confirm that the second loop is dependent on the interaction of the NAC and MADS genes, we deleted the NAC motif in the MADS1 promoter, as well as the MADS motif in the NAC promoter. We found that the tobacco leaf expressing the loop 2 without these motifs could no longer bypass the ethylene inhibitor treatment
(Fig. 3c). Taken together, our results showed that banana fruit ripening is controlled by a dual-loop circuit that consists of both leaf senescence NAC and floral organ identity MADS genes. The banana NAC and MADS genes are also associated with tissue-specific H3K27me3 marks (Fig. 4), suggesting that their epigenetic regulation
is conserved in both eudicots and monocot.

Climacteric fruit ripening gene orthologues in non-climacteric and dry fruit species. The fruitENCODE project included four non-climacteric species (cucumber, grape, strawberry and watermelon), none of which have undergone recent WGD[15]. They have orthologues of the carpel senescence NAC with tissue-specific H3K27me3 and a ripening-specific gene expression pattern similar to those in the NAC-type climacteric fruits without WGD (Fig. 5a and Supplementary Figs. 6,10–13). However, they often lack EIN3 motif in their NAC gene promoters or the NAC motif in their ethylene biosynthesis genes promoters, both of which would preclude participation in an ethylene positive feedback loop. The dry fruit-bearing plant Arabidopsis also has orthologues of the climacteric fruit ripening NAC and MADS genes. We examined their H3K27me3 levels in its leaf, carpel and senescence silique, which is the equivalent tissue of a ripening fleshy fruit (Fig. 5b). We found that its NAC and MADS gene have similar tissue-specific expression pattern and H3K27me3.

Banana is a monocot that has diverged from eudicots over 100 Mya. The dual-loop system in banana utilizes an orthologue of the rice leaf senescence OsNAP gene[21]. We have examined the gene expression and histone modification in young and senescence rice leaves, as well as its carpel tissues. We found that OsNAP is expressed in the aged leaves and is also associated with the repressive H3K27me3 mark in the young leaf and carpel tissues (Fig. 5c). Taken together, we showed that dry fruit and non-climacteric fleshy fruit species have orthologues of the climacteric fruit ripening genes. They are involved in leaf senescence, carpel senescence or floral development, and associated with tissue-specific H3K27me3 marks[17,18,21]. This suggests that the three ripening circuits in climacteric fruits were evolved from pre-existing pathways that served diferent functions in the ancestral angiosperms.

Discussion

The innovative fertilization and seed dispersal mechanisms are key to the evolutionary success of the flowering plants[23,24]. The fruitENCODE project has identified three major routes for angiosperms to evolve the ethylene-dependent climacteric fruit ripening process, and the core genetic elements and epigenetic mechanisms for these are present in non-climacteric and even dry-fruited species. Eudicots with recent WGD like tomato, utilized their duplicated MADS to form a positive feedback loop with ethylene to regulate

ripening, while eudicots without WGD repurposed the carpel senescence NAC (Fig. 6). Banana has two positive feedback loops using both leaf senescence NAC and duplicated MADS, which enables it to bypass ethylene inhibitor treatment after ripening initiation[20].

The tomato and banana ripening models we proposed are consistent with the observations that silencing key genes in the positive feedback loop could delay or abolish fruit ripening[22,23,25]. Mutation in the core NAC gene in climacteric melon cultivar Charentais Mono could also delay ripening[26], while both the peach and melon NAC loci are located in quantitative trait loci that are associated with late ripening phenotypes[16,27]. However, it should be noted that our proposed models only define the core transcriptional regulatory mechanisms centred on the ripening ethylene and do not preclude discovery of additional transcription factors or regulatory mechanisms such as post-translational regulations, which also contribute to ripening.

It is common for different species to evolve similar features when exposed to the same selection pressure. However, the probability of complex traits like ripening originating multiple times through similar trajectories would be expected to be very small, unless there is strong constraint. This constraint could be the limited set of suitable signalling molecules like the ethylene gas that can easily diffuse from cell to cell. For plants without WGD, another constraint could be the limited transcription factors available in the carpel tissues, hence leading to the repurposing of the senescence NAC to form the ripening circuit. In addition to evolving key regulators, plants also need to gain ethylene responsive cis-regulatory elements, such as EIN3 binding motifs, in the promoters of the NAC or MADS genes, as well as the corresponding motifs in their ACS genes to complete the core positive feedback loop. They also need to gain hundreds or even thousands of cis-regulatory elements in downstream ripening gene promoters in order to couple them to the loop.

DNA demethylation is required for the tomato fruit ripening[6]. Although local DNA methylation changes during fruit development are widespread in all species we examined (Supplementary Fig. 23), direct genetic evidence to link DNA methylation change to a regulatory role in ripening was only found in tomato[9,10]. Our study

revealed a surprisingly conserved role of H3K27me3 in regulating the core ripening genes and their orthologues (Figs. 1 and 5), while DNA methylation dynamics were often associated with promoter chromatin accessibility changes (Supplementary Fig. 21).

H3K27me3 is associated with silencing of key developmental genes in both animals and plants[28]. In animals, it is catalysed and bound by the polycomb repressive complexes, which condense chromatin and silence gene expression. In plants, H3K27me3 is best known for silencing the flowering regulator FLOWERING LOCUS C and floral homeotic gene AGAMOUS, both of which are MADS-box transcription factors[29]. For fleshy fruit species, it is of significant evolutionary advantage to use a stable epigenetic mark like H3K27me3 to keep the autocatalytic ripening loop under strict developmental control. We found that H3K27me3 targets key ethylene biosynthesis genes, as well as the MADS and NAC transcription factors in the loops. Their orthologues in four non-climacteric and two dry fruit species also have similar tissue-specific H3K27me3 dynamics, suggesting that the climacteric fruits have not just hijacked the genetic pathways in the ancestral angiosperms, but also utilized their epigenetic marks to regulate ripening.

However, the trigger for epigenome reprograming, including the tomato whole-genome cytosine demethylation and theremoval of H3K27me3 in the NAC and MADS loci, remains largely unknown. Our tomato dataset included two mutants, cnr and nor, the fruits of which do not synthesize ethylene or ripen if ethylene is supplied externally. Cnr is an SBP-box transcription factor epimutant[7]. Although it does not disrupt the expression of RIN, our data showed that the ripening gene promoters targeted by RIN became hypermethylated in Cnr[6,9]. The nor mutant fruit on the other hand is unable to express both RIN and ACS2, and these loci are associated with hyper-H3K27me3 when compared to wild-type (Supplementary Fig. 5). The nor fruit contains a missense mutation in a *NAC* gene orthologous to the ripening NAC used by other climacteric fruits (Supplementary Fig. 6), suggesting that the carpel senescence pathway is involved in tomato ripening initiation by controlling the H3K27me3.

The fruitENCODE project has generated a comprehensive funcional genomic resource for 11 fleshy fruit species, which opens the door for addressing some important problems in agricultural practices. For example, post-harvest loss is a major concern for horticulture

produce worldwide, especially in developing countries, but is also prevalent in modern food supply chains. Control of ethylene and ripening is critical because deterioration and rotting is an inevitable consequence of unhindered ripening. However, improvement in shelf-life through manipulation of ethylene often leads to reduced quality and nutritional value, which is to be expected because most of the downstream ripening genes are tightly coupled to the autocatalytic ethylene loop. With a comprehensive annotation of the cis-regulatory elements, and much improved understanding of their regulators, it is now possible to design strategies to engineer promoter cis-regulatory element to manipulate candidate gene expression to alter specific ripening attributes to improve nutritional quality, consumer appeal and shelf-life without affecting the general ripening process.

References

1. Coombe, B. G. The development of fleshy fruits. *Annu. Rev. Plant. Physiol.* **27**, 207–228 (1976).
2. Tiffney, B. H. Vertebrate dispersal of seed plants through time. *Annu. Rev. Ecol. Evol. Syst.* **35**, 1–29 (2004).
3. McMurchie, E. J., McGlasson, W. B., & Eaks, I. L. Treatment of fruit with propylene gives information about the biogenesis of ethylene. *Nature* **237**, 235–236 (1972).
4. McManus, M. T. *The plant hormone ethylene.* (Wiley–Blackwell, Hoboken, NJ, 2012).
5. Barry, C. S., Llop-Tous, M. I., & Grierson, D. The regulation of 1-aminocyclopropane-1-carboxylic acid synthase gene expression during the transition from system-1 to system-2 ethylene synthesis in tomato. *Plant Physiol.* **123**, 976–986 (2000).
6. Giovannoni, J., Nguyen, C., Ampofo, B., Zhong, S., & Fei, Z. The epigenome and transcriptional dynamics of fruit ripening. *Annu. Rev. Plant. Biol.* **68**, 61–84 (2017).
7. Manning, K. et al. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* **38**, 948–952 (2006).
8. Vrebalov, J. et al. A MADS-box gene necessary for fruit ripening at the

450 tomato RIPENING INHIBITOR (rin) locus. *Science* **296**, 343–346 (2002).
451 9. Zhong, S. et al. Single-base resolution methylomes of tomato fruit
452 development reveal epigenome modifications associated with ripening.
453 *Nat. Biotechnol.* **31**, 154–159 (2013).
454 10. Liu, R. et al. A DEMETER-like DNA demethylase governs tomato fruit
455 ripening. *Proc. Natl Acad. Sci. USA* **112**, 10804–10809 (2015).
456 11. Sato, S. et al. The tomato genome sequence provides insights into fleshy fruit
457 evolution. *Nature* **485**, 635–641 (2012).
458 12. Vrebalov, J. et al. Fleshy fruit expansion and ripening are regulated by the
459 tomato SHATTERPROOF gene TAGL1. *Plant Cell* **21**, 3041–3062 (2009).
460 13. Lin, Z., Zhong, S., & Grierson, D. Recent advances in ethylene research.
461 *J. Exp. Bot.* **60**, 3311–3336 (2009).
462 14. Guo, H. & Ecker, J. R. Plant responses to ethylene gas are mediated by
463 SCF(EBF1/EBF2)-dependent proteolysis of EIN3 transcription factor.
464 *Cell* **115**, 667–677 (2003).
465 15. Lee, T. H., Tang, H., Wang, X. & Paterson, A. H. PGDD: A database of gene
466 and genome duplication in plants. *Nucleic Acids Res.* **41**, 1152–1158 (2013).
467 16. Périn, C. et al. Molecular and genetic characterization of a non-climacteric
468 phenotype in melon reveals two loci conferring altered ethylene response in
469 fruit. *Plant Physiol.* **129**, 300–309 (2002).
470 17. Kunieda, T. et al. NAC family proteins NARS1/NAC2 and NARS2/NAM in
471 the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell Online*
472 **20**, 2631–2642 (2008).
473 18. Guo, Y., & Gan, S. AtNAP, a NAC family transcription factor, has an
474 important role in leaf senescence. *Plant J.* **46**, 601–612 (2006).
475 19. D'Hont, A. et al. The banana (*Musa acuminata*) genome and the evolution of
476 monocotyledonous plants. *Nature* **488**, 213–217 (2012).
477 20. Golding, J., Shearer, D., Wyllie, S., & McGlasson, W. Application of 1-MCP
478 and propylene to identify ethylene-dependent ripening processes in mature
479 banana fruit. *Postharvest. Biol. Technol.* **14**, 87–98 (1998).
480 21. Liang, C. et al. OsNAP connects abscisic acid and leaf senescence by
481 fine-tuning abscisic acid biosynthesis and directly targeting senescenceassociated
482 genes in rice. *Proc. Natl Acad. Sci. USA* **111**, 10013–10018 (2014).
483 22. Roy Choudhury, S., Roy, S., Nag, A., Singh, S. K. & Sengupta, D. N.

Characterization of an AGAMOUS-like MADS-box protein, a probable constituent of flowering and fruit ripening regulatory system in banana. *PLoS One* **7**, e44361 (2012).
23. Elitzur, T. et al. Banana *MaMADS* transcription factors are necessary for fruit ripening and molecular tools to promote shelf-life and food security. *Plant Physiol.* **171**, 380–391 (2016).
24. Primack, R. B. Relationships Among Flowers, Fruits, and Seeds. *Annu. Rev. Ecol. Syst.* **18**, 409–430 (1987).
25. Tieman, D. M., Ciardi, J. A., Taylor, M. G. & Klee, H. J. Members of the tomato LeEIL (EIN3-like) gene family are functionally redundant and regulate ethylene responses throughout plant development. *Plant J.* **26**, 47–58 (2001).
26. Ríos, P. et al. ETHQV6.3 is involved in melon climacteric fruit ripening and is encoded by a NAC domain transcription factor. *Plant J.* **91**, 671–683 (2017).
27. Pirona, R. et al. Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. *BMC Plant Biol.* **13**, 166 (2013).
28. Schwartz, Y. B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nat. Rev. Genet.* **8**, 9–22 (2007).
29. He, Y. Chromatin regulation of flowering. *Trends. Plant. Sci.* **17**, 556–562 (2012).

Acknowledgements: This work is supported by Hong Kong UGC GRF-14119814/14104515 and Area of Excellence Scheme AoE/M-403/16 to S.Z., Shenzhen Peacock-KQTD201101 to J.Z., Spanish Ministry of Economy and Competitively grant AGL2015-64625-C2-1-R, Centro de Excelencia Severo Ochoa 2016-2020 and the CERCA Programme/Generalitat de Catalunya to J.G.-M., and National Science Foundation IOS-1339287 to Z.F. and J.G.

Author contributions: S.Z. designed the research; P.L., N.Z., Y.C., B.Z., Y.P., J.F., J.A., N.Y. and J.Z. performed theexperiments. S.Y., D.T., S.Z. and J.X. analysed the data. D.G., J.G.-M., Z.F., J.G. and S.Z. wrote the paper.

Competing interests: The authors declare no competing interests.

Supplementary information

Supplementary figures S1-S27

Supplementary tables S1-S34

Fig. 1 Three types of transcriptional feedback circuits controlling climacteric fruit ripening. a, Model for tomato fruit ripening regulation. Ethylene transcription factor EIN3 activates the MADS transcription factor RIN. RIN forms a complex with TOMATO AGAMOUS-LIKE1 (TAGL1), and activates the ethylene biosynthesis genes, forming a positive feedback circuit that generates autocatalytic ethylene during ripening. Downstream ripening genes are directly coupled to the loop through the MADS transcription factors. In leaf and immature fruits, the loop is repressed with key genes associated with promoter DNA hypermethylation and repressive histone mark H3K27me3 in the gene body. b, Model for peach fruit ripening regulation, which utilizes a NAC instead of a MADS transcription factor. c, Model for banana fruit ripening regulation. An additional loop between the NAC and MADS enables the banana fruit to synthesize ethylene in the presence of ethylene inhibitor 1-methylcyclopropene (MCP) after ripening initiation.

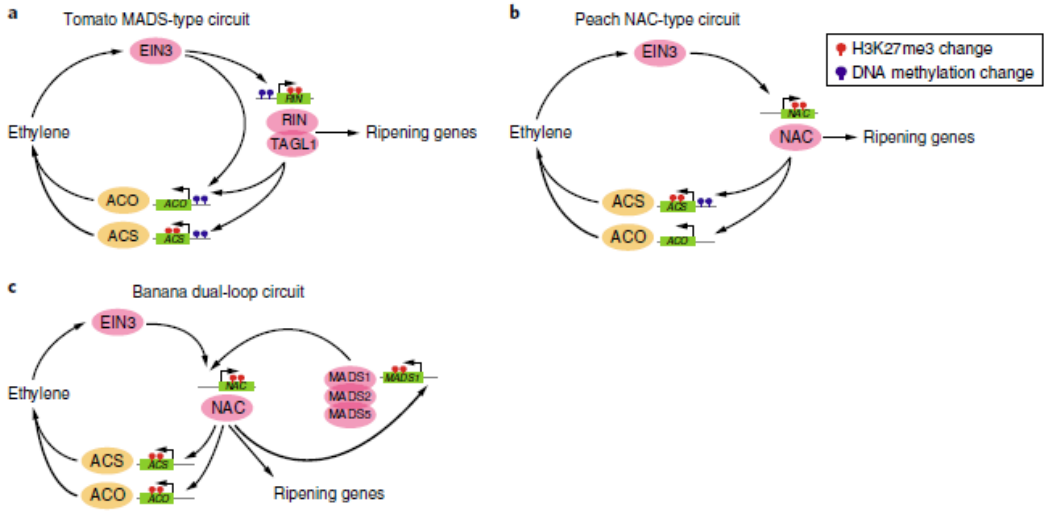

Fig. 2 Tomato and peach ripening genes are associated with tissue-specific epigenetic marks. a, Examples of dynamic chromatin accessibility, histone modification, DNA methylation and transcription in ripening gene loci in leaf, 17 days post-anthesis immature fruit and fully ripened tomato fruit tissues. Browser track shows normalized coverage of different features using merged data from multiple biological replicates. Detailed information is shown in Supplementary Table 2, 3, 5, 9 and 13. Individual data can be accessed on the fruitENCODE website. b, Peach ripening genes chromatin dynamics in leaf, 21 days post-anthesis immature fruit and fully ripened fruit tissues.

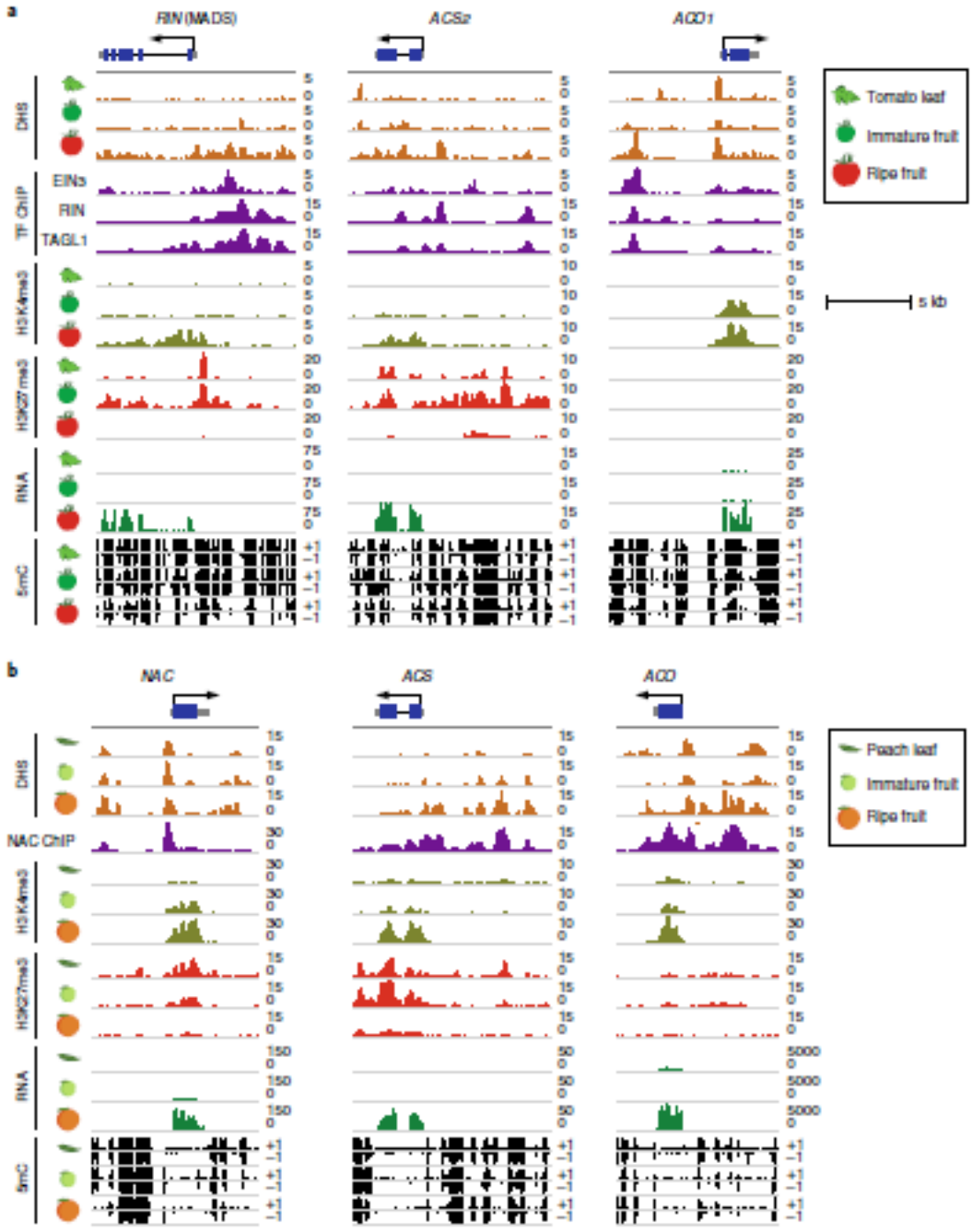

Fig. 3 Recreation of the positive feedback circuits for autocatalytic ethylene synthesis in tobacco. a, Expression of tomato MADS-loop components *RIN* and *ACS2* under their native promoter and *TAGL1* under the 35 S promoters is sufficient to generate autocatalytic ethylene. Mutation of the EIN3 motif in the *RIN* promoter or ethylene inhibitor 1-MCP treatment can disrupt the autocatalytic ethylene production. b, Ectopic expression of the peach *NAC* and *ACS* genes under native promoter generated autocatalytic ethylene. c, Ectopic expression of the *NAC* and *ACS* genes from the banana loop 1 generated autocatalytic ethylene, which could be disrupted by inhibitor treatment or EIN3 motif deletion. When loop 1 is co-expressed with the three *MADS* from loop 2, the autocatalytic ethylene could not be blocked by 1-MCP. Deletion of the NAC motif in the *MADS1* promoter and the MADS motif in the *NAC* promoter could disrupt the second loop. Sample sizes are shown in the figure legends. Individual values and their mean value are shown as dots and bars, respectively. Error bars represent ± s.e.m. *P* values were calculated using two side Student's *t*-test. Tobacco leaves infiltrated with empty vector were used as a mock control.

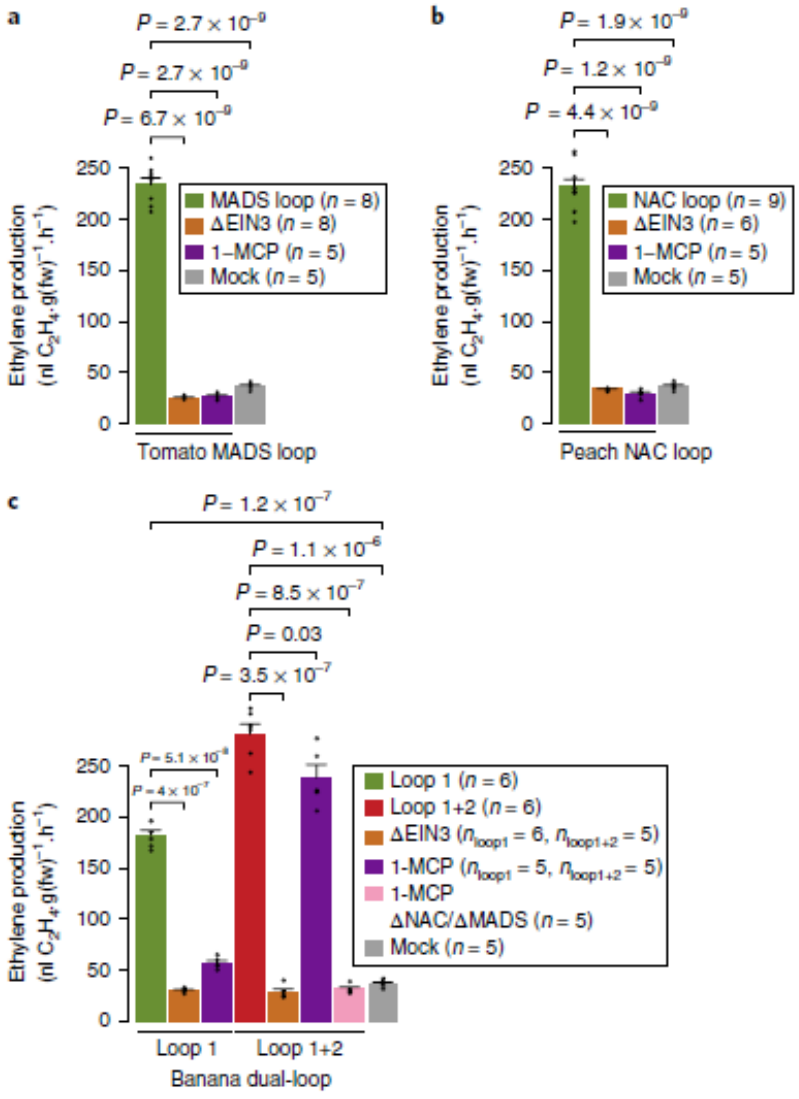

Fig. 4 Chromatin and epigenome features of the banana ripening genes. Key banana fruit ripening genes are associated with H3K27me3 in non-ripening tissues. The banana *NAC* is an orthologue of the monocot rice leaf senescence *OsNAP* and is associated with reduced H3K27me3 level in aged leaf and ripening fruit tissues. Browser track shows normalized coverage of different chromatin features using merged data from multiple biological replicates. Detailed information is shown in Supplementary Tables 2, 5, 9 and 13. Individual data can be accessed on the fruitENCODE website.

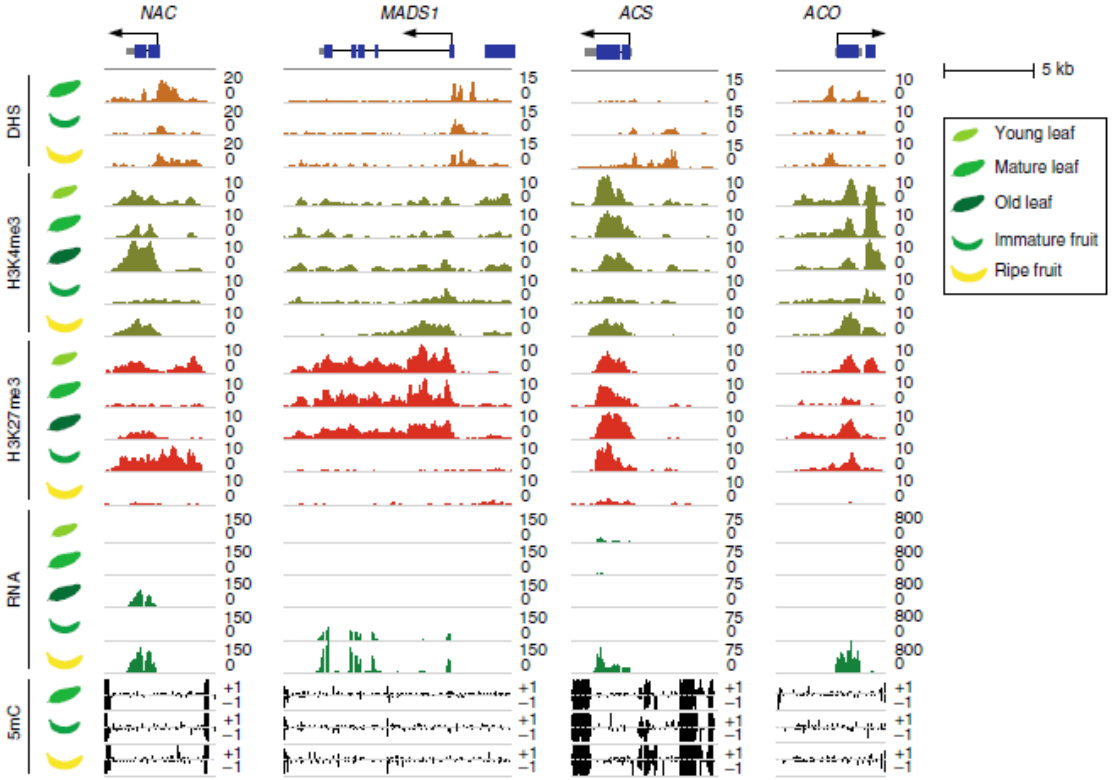

Fig. 5 Fruit ripening gene orthologues in non-climacteric and dry fruit species are associated with tissue-specific H3K27me3. **a**, Watermelon is a nonclimacteric species that does not require ethylene for fruit ripening. Its *NAC* and *MADS* orthologues are associated with H3K27me3 in leaf and immature fruit tissues. b, *NARS1* controlling *Arabidopsis* carpel senescence is an orthologue of the NAC transcription factors involved in eudicots climacteric fruit ripening. *SEP4* is the orthologue of the tomato *RIN*. c, Banana ripening gene orthologues in the monocot rice. Browser track shows normalized coverage of different chromatin features using merged data from multiple biological replicates. Detailed information is shown in Supplementary Table 2, 9 and 13. Individual data can be accessed on the fruitENCODE website.

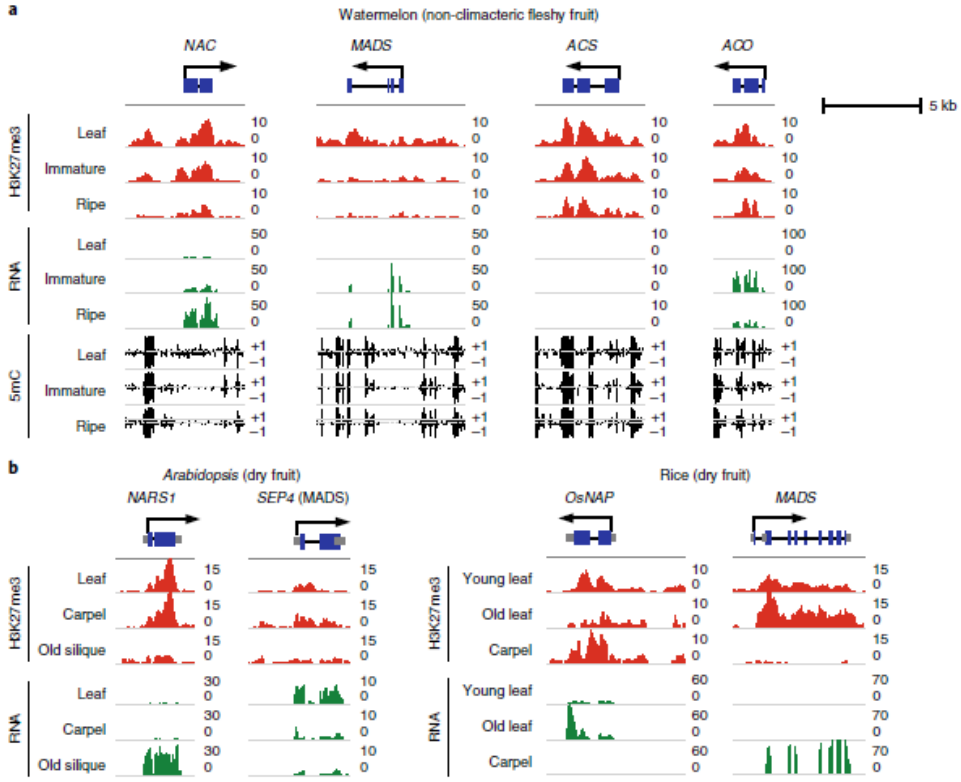

Fig. 6 Speciation, fruit ripening types and polyploidization in diferent angiosperms lineages. Plant species bearing dry, fleshy climacteric and fleshy non-climacteric fruits are indicated in black, red and green, respectively. The basal angiosperm *Schisandra chinesis* is highlighted with a question mark because it bears fleshy fruit with uncharacterized ripening behaviour. Confirmed whole-genome duplications and triplications are shown with red and green circles. The three types of ethylene-dependent regulatory circuit are shown in parenthesis.