



*This is a post-peer-review, pre-copyedit version of an article published in Journal of Applied Phycology. The final authenticated version is available online at:*

<https://doi.org/10.1007/s10811-018-1456-8>

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

3

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

8

9 <sup>1</sup>State Key Laboratory of Agrobiotechnology, School of Life Sciences, Chinese  
10 University of Hong Kong, Hong Kong, China.

11 <sup>2</sup>College of Horticulture, South China Agricultural University, Guangzhou, China.

12 <sup>3</sup>College of Horticulture and Landscape Architecture, Southwest University, Chongqing,  
13 China.

14 <sup>4</sup>Department of Food Science and Experimental Nutrition, FCF, University of Sao Paulo,  
15 Sao Paulo, Brazil.

16 <sup>5</sup>IRTA, Centre for Research in Agricultural Genomics, Barcelona, Spain.

17 <sup>6</sup>Department of Biology, Hong Kong Baptist University, Hong Kong, China.

18 <sup>7</sup>School of Crop Sciences, University of Nottingham, Nottingham, UK.

19 <sup>8</sup>Zhejiang Provincial Key Laboratory of Horticultural Plant Integrative Biology, Zhejiang  
20 University, Hangzhou, China.

21 <sup>9</sup>Weill Medical College, Cornell University, New York, NY, USA.

22 <sup>10</sup>US Department of Agriculture—Agricultural Research Service and Boyce Thompson  
23 Institute for Plant Research, Cornell University, Ithaca, NY, USA.

24

25 <sup>11</sup>These authors contributed equally.

26 \*e-mail: [silin.zhong@cuhk.edu.hk](mailto:silin.zhong@cuhk.edu.hk)

27

28 **Abstract**

29

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

42

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

## 45 **Introduction**

46

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

55

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

65

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

73

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

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

87

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

103

## 104 **Results**

105

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

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

115

116 We have also included a large collection of mutants with altered ripening phenotypes. For  
117 other non-model species such as melon, which lack large mutant collections, we have  
118 included four varieties with different ripening characteristics for comparative analysis.

119 The current dataset encompassed 361 transcriptome, 71 accessible chromatin,  
120 147 histone modification and 45 methylome profiles. All processed datasets can be  
121 accessed from the fruitENCODE data base ([www.epigenome.cuhk.edu.hk/encode.html](http://www.epigenome.cuhk.edu.hk/encode.html)).

122

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

133

### 134 **MADS-type positive feedback loop controlling ripening in tomato, apple and pear.**

135 We first reconstructed the ripening regulatory circuit for the model tomato fruit, which  
136 has three known components: ethylene, transcription factor RIN and DNA methylation.

137 From the DNaseI hypersensitive sites (DHS) dataset, we found an EIN3 binding motif in  
138 the promoter of RIN, the functional significance of which was confirmed by our EIN3  
139 ChIP-Seq (Fig. 2a and Supplementary Table 32). RIN is a MADS-box transcription  
140 factor, which functions in a multimeric complex with TAGL1. TAGL1 is

141 expressed during both early and late fruit development<sup>12</sup>. We performed ChIP-Seq for  
142 both RIN and TAGL1, and found that they can directly target the ripening ethylene  
143 biosynthesis genes ACC SYNTHASE2 (ACS2) and ACC OXIDASE1 (ACO1) (Fig. 2a).

144

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

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

153

154 To test this feedback loop, we attempted to recreate the autocatalytic ethylene synthesis  
155 in tobacco leaf by expressing the core tomato loop genes. It is well known that the  
156 ethylene synthesis and signalling pathways are conserved in plants<sup>4,13</sup>. Leaf has  
157 endogenous EIN3 and ACO, and lacks ACS activity, which is the rate-limiting enzyme  
158 in its ethylene biosynthesis under normal growth conditions. When we expressed the  
159 tomato RIN and ACS2 using their native promoters, and TAGL1 was supplied under a  
160 constitutive CaMV35S promoter, spontaneous ethylene synthesis was observed (Fig. 3a  
161 and Supplementary Fig. 1). We could then mutate the EIN3 binding motif in the RIN  
162 promoter and the RIN binding motif in the ACS2 promoter, both of which disrupted the  
163 spontaneous ethylene synthesis (Fig. 3a and Supplementary Fig. 1c). These results  
164 confirmed that the exogenous tomato ACS2 and MADS genes are responsible for the  
165 observed ethylene burst. We also treated the tobacco leaf expressing the loop with  
166 ethylene inhibitor 1-MCP, which causes degradation of the EIN3 protein<sup>14</sup>. We found that  
167 1-MCP blocked the ethylene synthesis, suggesting that the ethylene generated by the  
168 tomato MADS loop in tobacco leaf is indeed autocatalytic, a key characteristic of the  
169 system II ethylene produced by ripening climacteric fruits (Fig. 3a and Supplementary  
170 Fig. 1b).

171

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

179

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

191

192 Tomato is the only reported species to activate DNA demethylase expression during fruit  
193 ripening<sup>10</sup>. Our methylome datasets confirmed that apple and pear lack a tomato-like  
194 whole-genome demethylation during fruit development (Supplementary Fig. 19).

195 However, from the histone modification datasets, we found that tomato, apple and pear  
196 have the same tissue-specific H3K27me3 on their ethylene biosynthesis and MADS gene  
197 loci, whereas they are absent in the ripe fruit tissues (Fig. 2a, Supplementary Fig. 4). We  
198 also found that the ethylene-independent pear cultivar Dangshanshuli contains hyper-  
199 H3K27me3 in its ethylene biosynthesis gene loci compared to the ethylene-dependent  
200 cultivar Williams (Supplementary Fig. 3). The non-ripening tomato mutant nor also  
201 contains hyper-H3K27me3 in the ACS2 and RIN loci (Supplementary Fig. 5). This  
202 suggests that instead of using DNA methylation, H3K27me3 could play a conserved role  
203 from preventing the MADS positive feedback loop from generating  
204 autocatalytic ethylene.

205

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



214

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

225

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

235

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

241

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

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

253

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

264

265 The first banana loop is similar to the NAC-type positive feedback loop in eudicots  
266 without WGD (Fig. 1c). The banana ACS (Ma04\_t35640.1) and ACO (Ma07\_t19730.1)  
267 have NAC motifs in their promoter DHS, while the NAC gene (Ma06\_t33980.1) contains  
268 an EIN3 binding motif. To test the loop, we ectopically expressed the banana NAC and  
269 ACS genes under their native promoters in tobacco and found that they are sufficient to  
270 generate ethylene in an autocatalytic manner. Ethylene inhibitor 1-MCP, as well as EIN3  
271 motif deletion, could block ethylene production in the absence of the second loop,  
272 suggesting that loop I is a functional NAC-type positive feedback loop (Fig. 3c).

273

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

281 in the aged leaves (Fig. 4).

282

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

289

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

297

298 To validate the second loop, we first used the tobacco system to show that the first NAC-  
299 type loop could be blocked by inhibitor 1-MCP treatment (Fig. 3c). Co-expressing the  
300 three MADS genes from the second loop with MADS1 driven by its native promoter and  
301 MADS2/5 with the constitutive 35S promoter enabled the  
302 tobacco leaves to synthesize more ethylene than expressing loop 1 alone (54.27%,  $P =$   
303  $2.91 \times 10^{-5}$ ). Most importantly, 1-MCP was unable to block the ethylene production  
304 when the second loop was present, mimicking the behaviour of the ripening banana fruit  
305 (Fig. 3c). To confirm that the second loop is dependent on the interaction of the NAC and  
306 MADS genes, we deleted the NAC motif in the MADS1 promoter, as well as the MADS  
307 motif in the NAC promoter. We found that the tobacco leaf expressing the loop 2 without  
308 these motifs could no longer bypass the ethylene inhibitor treatment  
309 (Fig. 3c). Taken together, our results showed that banana fruit ripening is controlled by a  
310 dual-loop circuit that consists of both leaf senescence NAC and floral organ identity  
311 MADS genes. The banana NAC and MADS genes are also associated with tissue-specific  
312 H3K27me3 marks (Fig. 4), suggesting that their epigenetic regulation  
313 is conserved in both eudicots and monocot.

314

315 **Climacteric fruit ripening gene orthologues in non-climacteric and dry fruit species.**

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

328

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

340

341 **Discussion**

342

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

349 ripening, while eudicots without WGD repurposed the carpel senescence NAC (Fig. 6).  
350 Banana has two positive feedback loops using both leaf senescence NAC and duplicated  
351 MADS, which enables it to bypass ethylene inhibitor treatment after  
352 ripening initiation<sup>20</sup>.

353

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

363

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

377

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

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

385

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

399

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

412

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

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

426

427

## 428 **References**

429

430 1. Coombe, B. G. The development of fleshy fruits. *Annu. Rev. Plant. Physiol.* **27**,  
431 207–228 (1976).

432 2. Tiffney, B. H. Vertebrate dispersal of seed plants through time. *Annu. Rev.*  
433 *Ecol. Evol. Syst.* **35**, 1–29 (2004).

434 3. McMurchie, E. J., McGlasson, W. B., & Eaks, I. L. Treatment of fruit with  
435 propylene gives information about the biogenesis of ethylene. *Nature* **237**,  
436 235–236 (1972).

437 4. McManus, M. T. *The plant hormone ethylene*. (Wiley–Blackwell, Hoboken,  
438 NJ, 2012).

439 5. Barry, C. S., Llop-Tous, M. I., & Grierson, D. The regulation of  
440 1-aminocyclopropane-1-carboxylic acid synthase gene expression during  
441 the transition from system-1 to system-2 ethylene synthesis in tomato.  
442 *Plant Physiol.* **123**, 976–986 (2000).

443 6. Giovannoni, J., Nguyen, C., Ampofo, B., Zhong, S., & Fei, Z. The epigenome  
444 and transcriptional dynamics of fruit ripening. *Annu. Rev. Plant. Biol.* **68**,  
445 61–84 (2017).

446 7. Manning, K. et al. A naturally occurring epigenetic mutation in a gene  
447 encoding an SBP-box transcription factor inhibits tomato fruit ripening.  
448 *Nat. Genet.* **38**, 948–952 (2006).

449 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 senescence-associated  
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.



- 484 Characterization of an AGAMOUS-like MADS-box protein, a probable  
485 constituent of flowering and fruit ripening regulatory system in banana.  
486 *PLoS One* **7**, e44361 (2012).
- 487 23. Elitzur, T. et al. Banana *MaMADS* transcription factors are necessary for  
488 fruit ripening and molecular tools to promote shelf-life and food security.  
489 *Plant Physiol.* **171**, 380–391 (2016).
- 490 24. Primack, R. B. Relationships Among Flowers, Fruits, and Seeds. *Annu. Rev.*  
491 *Ecol. Syst.* **18**, 409–430 (1987).
- 492 25. Tieman, D. M., Ciardi, J. A., Taylor, M. G. & Klee, H. J. Members of the  
493 tomato LeEIL (EIN3-like) gene family are functionally redundant and  
494 regulate ethylene responses throughout plant development. *Plant J.* **26**,  
495 47–58 (2001).
- 496 26. Ríos, P. et al. ETHQV6.3 is involved in melon climacteric fruit ripening  
497 and is encoded by a NAC domain transcription factor. *Plant J.* **91**,  
498 671–683 (2017).
- 499 27. Pirona, R. et al. Fine mapping and identification of a candidate gene  
500 for a major locus controlling maturity date in peach. *BMC Plant Biol.* **13**,  
501 166 (2013).
- 502 28. Schwartz, Y. B. & Pirrotta, V. Polycomb silencing mechanisms and the  
503 management of genomic programmes. *Nat. Rev. Genet.* **8**, 9–22 (2007).
- 504 29. He, Y. Chromatin regulation of flowering. *Trends. Plant. Sci.* **17**,  
505 556–562 (2012).

506

507 **Acknowledgements:** This work is supported by Hong Kong UGC GRF-  
508 14119814/14104515 and Area of Excellence Scheme AoE/M-403/16 to S.Z., Shenzhen  
509 Peacock-KQTD201101 to J.Z., Spanish Ministry of Economy and Competitively grant  
510 AGL2015-64625-C2-1-R, Centro de Excelencia Severo Ochoa 2016-2020 and the  
511 CERCA Programme/Generalitat de Catalunya to J.G.-M., and National Science  
512 Foundation IOS-1339287 to Z.F. and J.G.

513

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

517

518 **Competing interests:** The authors declare no competing interests.

519

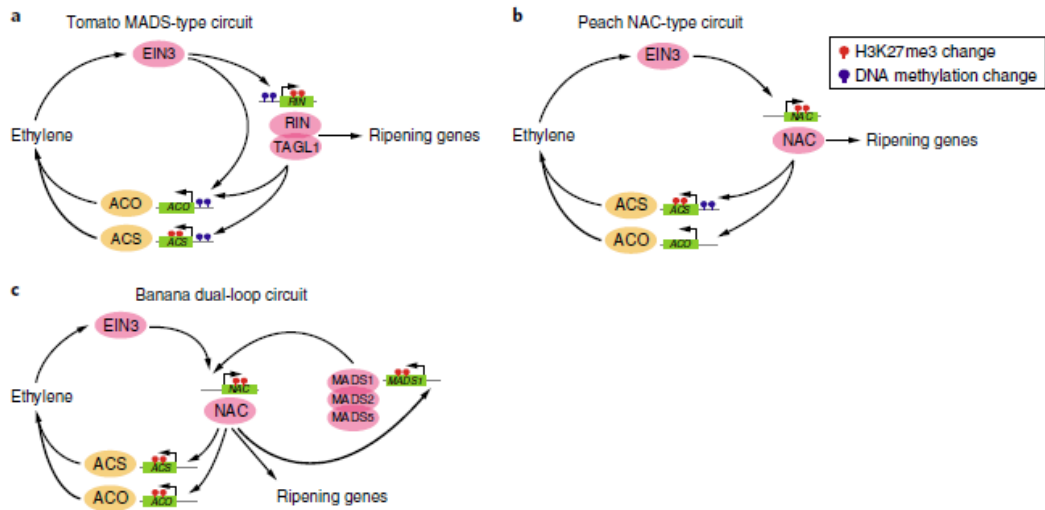
520 **Supplementary information**

521 **Supplementary figures S1-S27**

522 **Supplementary tables S1-S34**

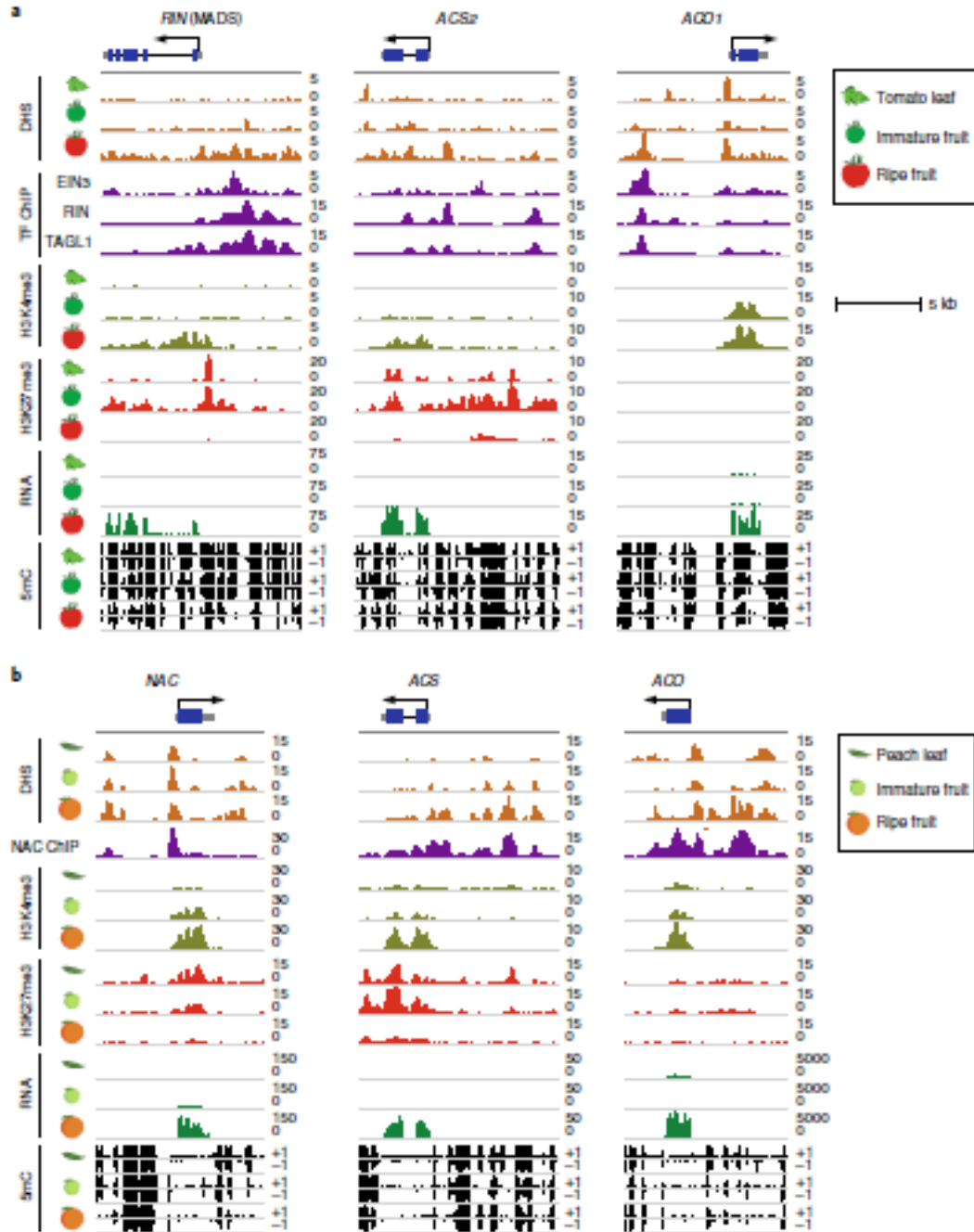
523

524 **Fig. 1** Three types of transcriptional feedback circuits controlling climacteric fruit ripening. a, Model for  
 525 tomato fruit ripening regulation. Ethylene transcription factor EIN3 activates the MADS transcription  
 526 factor RIN. RIN forms a complex with TOMATO AGAMOUS-LIKE1 (TAGL1), and activates the  
 527 ethylene biosynthesis genes, forming a positive feedback circuit that generates autocatalytic ethylene during  
 528 ripening. Downstream ripening genes are directly coupled to the loop through the MADS transcription  
 529 factors. In leaf and immature fruits, the loop is repressed with key genes associated with  
 530 promoter DNA hypermethylation and repressive histone mark H3K27me3 in the gene body. b, Model for  
 531 peach fruit ripening regulation, which utilizes a NAC instead of a MADS transcription factor. c, Model for  
 532 banana fruit ripening regulation. An additional loop between the NAC and MADS enables the  
 533 banana fruit to synthesize ethylene in the presence of ethylene inhibitor 1-methylcyclopropene (MCP) after  
 534 ripening initiation.  
 535



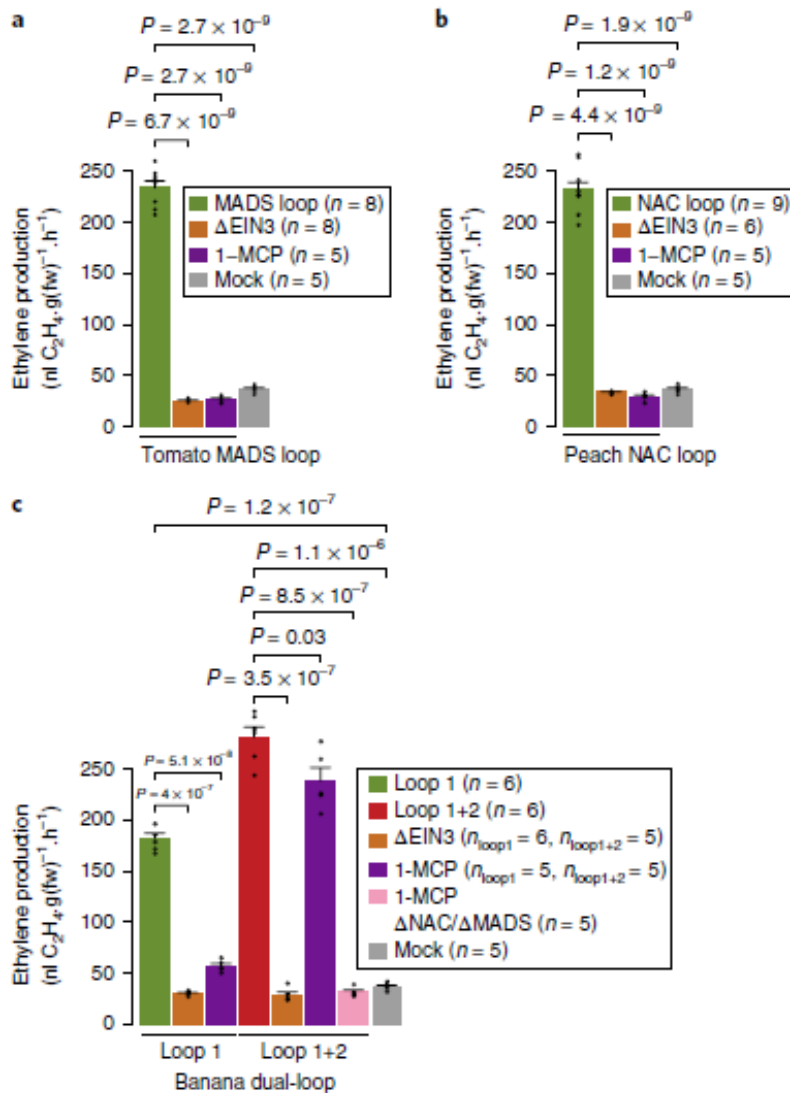
536  
 537

538 **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.



545  
546

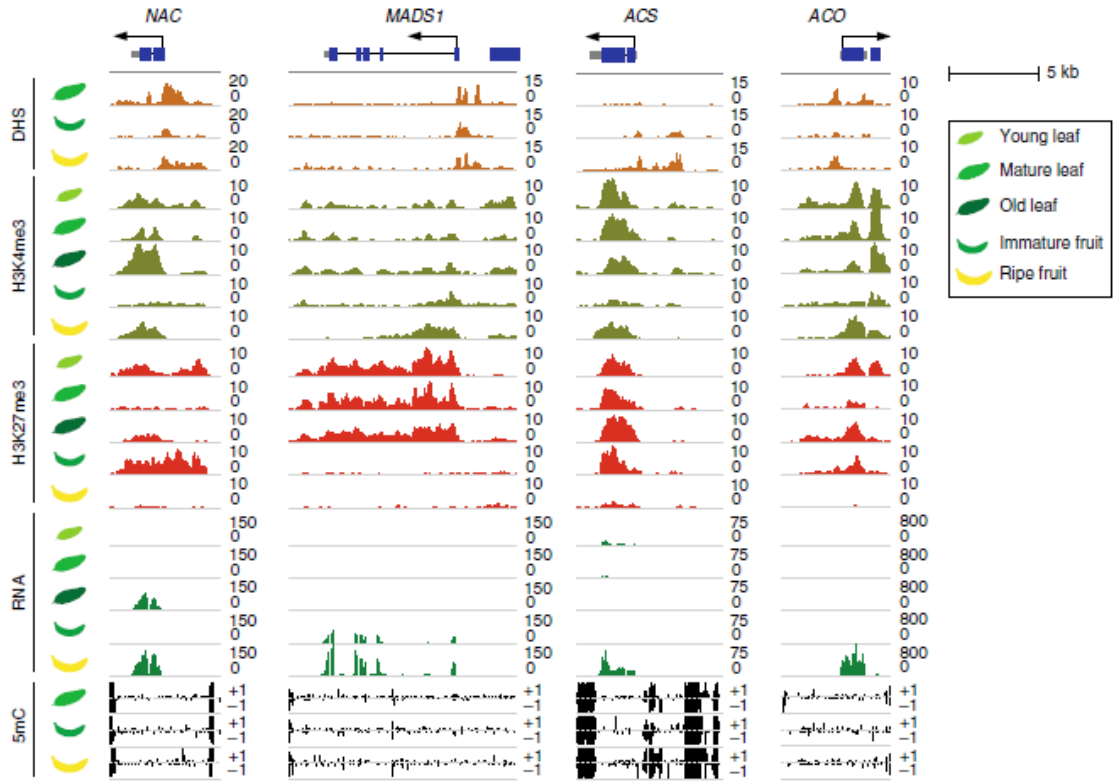
547 **Fig. 3** Recreation of the positive feedback circuits for autocatalytic ethylene synthesis in tobacco. a,  
 548 Expression of tomato MADS-loop components *RIN* and *ACS2* under their native promoter and *TAGL1*  
 549 under the 35 S promoters is sufficient to generate autocatalytic ethylene. Mutation of the EIN3 motif in the  
 550 *RIN* promoter or ethylene inhibitor 1-MCP treatment can disrupt the autocatalytic ethylene production. b,  
 551 Ectopic expression of the peach *NAC* and *ACS* genes under native promoter generated  
 552 autocatalytic ethylene. c, Ectopic expression of the *NAC* and *ACS* genes from the banana loop 1 generated  
 553 autocatalytic ethylene, which could be disrupted by inhibitor treatment or EIN3 motif deletion. When loop  
 554 1 is co-expressed with the three *MADS* from loop 2, the autocatalytic ethylene could not be blocked by 1-  
 555 MCP. Deletion of the *NAC* motif in the *MADS1* promoter and the *MADS* motif in the *NAC* promoter could  
 556 disrupt the second loop. Sample sizes are shown in the figure legends. Individual values  
 557 and their mean value are shown as dots and bars, respectively. Error bars represent  $\pm$  s.e.m. *P* values were  
 558 calculated using two side Student's *t*-test. Tobacco leaves infiltrated with empty vector were used as a mock control.  
 559



560

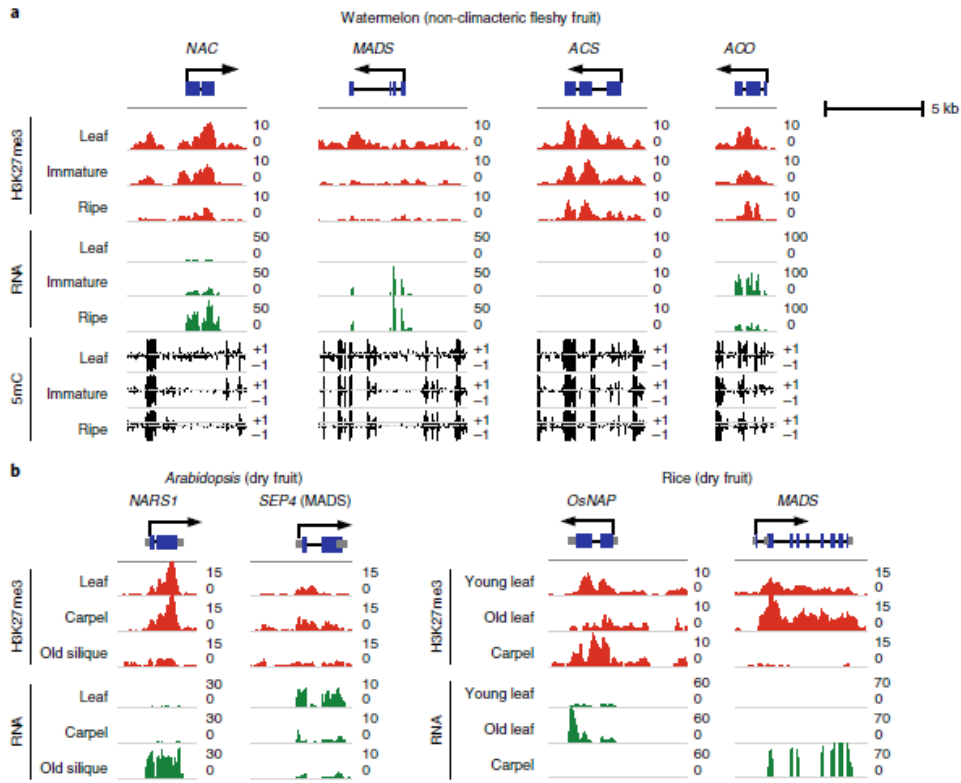
561

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



568  
569

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



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

