# Regulatory structural variation upstream of a NAC transcription factor controls fruit maturity timing in peach


## Authors

Tagliabue Alessandro Giulio[1], Friel James[1], Chiozzotto Remo[1], Baccichet Irina[1], Da Silva Linge Cassia[1], Calastri Elisa[1], Biffi Gianluca[1], Zaracho Nathalia[2], Gattolin Stefano[3], Sabrina Micali[4], Eduardo Iban[2], Bassi Daniele[1], Rossini Laura[1] and Marco Cirilli[1]

## Affiliation

[1]Department of Agricultural and Environmental Sciences (DISAA), University of Milan, 20133 Milan, Italy

[2]Centre for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB-UB, Campus UAB, Edifici CRAG, Cerdanyola del Vallès (Bellaterra), 08193 Barcelona, Spain

[3]Institute of Agricultural Biology and Biotechnology (IBBA), CNR - National Research Council of Italy, 20133 Milan, Italy

[4]Consiglio per la Ricerca in Agricoltura e l'analisi dell'Economia Agraria, Research Centre for Olive, Fruit, and Citrus Crops, 00134 Rome, Italy

## Corresponding authors

*email: marco.cirilli@unimi.it; laura.rossini@unimi.it


## Highlights

Structural changes in DNA upstream of a key gene (*NAC1*) control when peaches ripen. These variants shift gene activation timing, explaining early or late harvests and enabling breeders to predict fruit maturity.


## Abstract

Fruit maturity date is a key developmental trait in fruit trees, as it determines the timing at which fruits enter the ripening phase, thereby shaping harvest calendar and fruit quality. In peach, ripening time varies over several months, yet the genetic and regulatory basis underlying this extensive phenological diversity has remained largely unresolved. Here, we dissect the long-standing chromosome 4 maturity date locus (*qMD4.1*) and show that a multi-allelic series of



© The Author(s) 2026. Published by Oxford University Press on behalf of Society for Experimental Biology. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.




ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

structural variants in the upstream regulatory region of the transcription factor NAC1 explains most of the observed variation in ripening time in cultivated peach. Across multiple independent populations and diverse genetic backgrounds, these variants (collectively termed *Md*) stratify maturity date from ultra-early to late cultivars with largely additive and dosage-dependent effects. Using recombinant-based genetic dissection, long-read sequencing, transcriptomic analyses and epigenomic data integration, we show that *Md* alleles are consistently associated with allele-dependent shifts in the temporal dynamics of *NAC1* transcript accumulation, and that the affected region overlaps with chromatin features characteristic of regulatory activity. By resolving the genetic architecture of a major phenological locus, our results support a central role for cis-regulatory structural variation in modulating developmental timing in peach and provide biologically interpretable markers for predicting maturity date. More broadly, this work illustrates how non-coding structural variation can contribute to variation in fruit developmental trajectories, with implications for breeding strategies aimed at modulating harvest windows in peach and other stone fruit species.

**Keywords**

QTL; fruit trees; maturity date; peach; ripening; structural variants

## Introduction

Across the genus Prunus, differences in fruit maturation timing underlie much of the agronomic diversity observed among stone fruit cultivars, influencing adaptation, harvest scheduling and breeding decisions. Understanding the genetic and regulatory basis of these differences is therefore central to fruit biology and improvement.

Peach [*P. persica* (L.) Batsch] is the reference species for stone fruits due to its economic value, small genome, low heterozygosity and close phylogenetic relationship with other major *Prunus* crops (**Shulaev et al., 2008**). As a climacteric drupe, peach undergoes a well-defined developmental and ripening program, making it an excellent model for investigating the genetic regulation of fruit maturation. Fruit development in peach typically follows the characteristic double-sigmoid pattern of climacteric drupes, in which coordinated physiological, hormonal, and molecular events determine final size, flesh texture, composition and sensory quality (**Tonutti et al., 2023**). The early stage (S1) is characterized by rapid growth driven by intense cell division (**Pavel and DeJong, 1993**; **Lombardo et al., 2011**). During S2, endocarp lignification occurs via activation of the phenylpropanoid and lignin biosynthetic pathways, temporarily slowing mesocarp enlargement (**Dardick and Callahan, 2014**). This stage is reduced or absent in early-ripening

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

cultivars. In S3, growth resumes through cell expansion, accompanied by the accumulation of sugars and organic acids (**Etienne et al., 2002**) until the transition to S4, when fruit reaches its final size and ripening initiates (**Brummell et al., 2004**). Auxin promotes the onset of S4 by enhancing ethylene biosynthesis (**Trainotti et al., 2007**), leading to climacteric rise in respiration and autocatalytic ethylene production mediated by the ACS1-ACO1 pathway, which triggers texture changes, colour development, and flavour formation (**Rasori et al., 2002; Tatsuki et al., 2013**).

Fruit maturity date (MD) is a key trait in peach, directly affecting orchard management, harvesting scheduling, post-harvest handling and fruit quality (**Crisosto, 1997**). Peach cultivars exhibit wide variability, ranging from very early (May) to ultra-late (October) maturity in the Northern hemisphere (**Baccichet et al., 2025**). Although MD has a strong genetic basis, it is also shaped by environmental and agronomic factors, including temperature regimes affecting the accumulation of growing degree hours (**Souza et al., 2019**), as well as crop load, water availability and canopy microclimate (**Reighard et al., 2017; Yamaguchi et al., 2002**).

Because of its agronomic relevance, MD has been the target of numerous genetic studies, all reporting high heritability and a predominant genetic component. Early mapping work identified a major-effect QTL on linkage group 4 (*qMD4.1*), with pleiotropic effects on ripening date and quality traits (**Dirlewanger et al., 1999; Eduardo et al., 2011**). The presence and stability of this locus have been confirmed across multiple peach families, where it explains a large fraction of phenotypic variation, often with discrete bimodal or trimodal distributions (**Bassi et al., 1988; Fresnedo-Ramírez et al., 2015; Rawandoozi et al., 2021**). Fine mapping of the locus narrowed the interval to a ~220 kb genomic region, containing two NAC transcription factor genes: *Prupe.4G187100* (ppa007577m, *NAC1*), highly similar to the tomato *NON-RIPENING* (*NOR*) gene, and *Prupe.4G186800* (ppa008301m, *NAC5*), closely related to *ANAC072/055/019*, a gene primarily involved in stress responses (**Pirona et al., 2013**). A 9-bp INDEL in the third exon of *NAC5* (alleles $NAC_9$ and $NAC_0$) was proposed as the causal variant for *qMD4.1*. Despite the predictive value of this variant across different germplasm, a substantial portion of phenotypic variability remained unexplained (**Cao et al., 2023; da Silva Linge et al., 2024**), suggesting the presence of additional allelic diversity. A large deletion encompassing the entire *NAC5* gene up to the proximal *NAC1* promoter (~500 bp from the TSS) was also mapped in the same genomic region and underlies the recessive slow-ripening (*sr*) phenotype, in which fruits fail to ripen normally (**Eduardo et al., 2015; Nuñez-Lillo et al., 2015**). Functional evidence increasingly pointed to *NAC1* as primary regulator of ripening, as it showed fruit-specific expression and a climacteric-associated transcriptional peak (**Cao et al., 2023**). Nevertheless, no sequence

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

variants in *NAC1* regulatory or coding regions have been conclusively associated with MD, leaving the molecular basis of *qMD4.1* unresolved.

Insights into the genetic regulation of fruit ripening have largely originated from tomato, where natural mutants identified key transcriptional regulators such as NOR, CNR and RIN, belonging to the NAC, MADS-box, and SBP-box families, respectively (**Vrebalov et al., 2002; Manning et al., 2006; Giovannoni et al., 2004**). Among these, the NOR functions as a master regulator coordinating ethylene biosynthesis, cell-wall modification, pigment accumulation and aroma formation during ripening (**Osorio et al., 2011**). The central role of NAC transcription factors in ripening regulation appears conserved across species, as demonstrated for example in apple (*MdNAC18.1*), strawberry, banana and kiwifruit (**Martín-Pizarro et al., 2021**; **Li et al., 2024**; **Yue et al., 2024**; **Wu et al. 2025**).

In this study, we dissect the genetic architecture of maturity date in peach and resolve the long-standing ambiguity surrounding the *qMD4.1* locus. We demonstrate that the previously proposed INDEL does not explain MD variation. Instead, we identify a multi-allelic series of structural variants in the upstream regulatory region of *NAC1* (named *Md*) that perfectly co-segregate with ripening time across different mapping families and germplasm. Beyond establishing the causal variants, this work provides robust diagnostic markers enabling accurate allele-based prediction for marker-assisted breeding.

## Materials and Methods

### Plant materials and phenotyping

Accessions, breeding selections and 151 F1 progenies from 'Dulcebo66' × 'Pulchra' belong to the University of Milan collection located in the experimental farm 'M. Neri' of RiNOVA in Imola (Emilia-Romagna region, Italy). Seedlings were own-rooted and planted at distance of 1 x 4 meters (within and between rows, respectively), trained as free central leader and managed according to standard cultural practices for irrigation, fertilization and pruning. Accessions and breeding selections were grafted on GF677 rootstock, trained according to open vase system and regularly spaced at a distance of 4 x 2.5 meters. Fruits were thinned within 40-60 days after bloom, setting a crop load proportional to tree vigor. Maturity date (MD) trait was recorded when 2-3% of the fruits on the tree had reached full physiological maturity, assessed by sensory and visual inspection, also recording IAD index (DA-meter, TR-Turoni, Forlì, Italy) and fruit firmness (puncture test by a digital penetrometer). MD was expressed as the number of days from 1 January (Julian days, JD). MD was evaluated for 2 years (2014 and 2016). Additional segregating populations previously used for maturity date (MD) QTL mapping included Bt×Ak (n = 72) and

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

Bt×Nr (n = 40) (**Serra et al., 2017**); $F_2$ E×E (n = 72) (**Kalluri et al., 2022**); $F_2$ W×By (n = 102) and $F_2$ C×A (n = 302) (**Pirona et al., 2013**); M×R (n = 69) (**Ciacciulli et al., 2018**). Detailed phenotypic data and haplotypes at the Md locus are provided in Supplementary Table S1.

**Genotyping, linkage map construction and QTL-mapping**

The 18K peach SNP array was used to genotype the DxP population and the accessions panel, as previously described (**Cirilli et al., 2021**). Genotyping data were initially filtered for marker missing rate <10%. Non-informative markers, including monomorphic or non-segregating SNPs, as well as redundant loci with identical segregation pattern (genotype similarity > 0.95) were excluded. Markers showing distorted segregation from the expected 1:1 ratio for pseudo-testcross progeny were removed using chi-square goodness-of-fit tests ($p < 0.05$). A total of 2,033 and 2,826 informative markers were retained for D66 and P, whose physical positions were retrieved from the peach reference genome assembly V2.0. Single-marker analysis was performed in Tassel 5.2.15 (**Bradbury et al. 2007**) using a general linear model (GLM), testing marker genotype as a fixed effect. DxP linkage maps were constructed following the two-way pseudo-testcross strategy for outcrossing species (CP) using JoinMap v4.1 (**Van Ooijen, 2006**), coding SNP as lm×ll and nn×np configurations, respectively for the seed and pollen parents. Linkage groups were defined using a minimum logarithm of odds (LOD) value of 10.0 using the regression algorithm (Kosambi mapping function), with a recombination frequency threshold of 0.4, LOD value of 1.0 and a goodness-of-fit jump of 5.0. The MapQTL v6.0 software was used for detecting QTLs. Significance thresholds were calculated by random permutation test (PT) with 10,000 replicates considering the genome-wide LOD scores corresponding to $p = 0.05$. The interval mapping (IM) function was employed for QTLs detection with 95% significance and estimation of the percentage of phenotypic variation. MapChart v2.1 software was used to draw the mapped QTLs and the LOD plots.

**DNA extraction and resequencing**

For each sampled accession, fresh young tissue was collected during spring, immediately snap-frozen in liquid nitrogen, and stored at –80 °C. High-molecular-weight (HMW) genomic DNA was extracted following the protocol described by **Kang et al. (2023)** with optimizations tailored for Oxford Nanopore Technologies (ONT) applications. Library preparation was performed using the Ligation Sequencing Kit v14 (SQK-LSK114, ONT). Each DNA sample meeting the quality requirements underwent a modified ligation-based library preparation workflow. A minimum input of 1.4 µg of HMW genomic DNA was used for the end-preparation step. Incubation times at 20 °C and 65 °C were extended to 30 minutes. Magnetic bead volumes were adjusted for each sample based on DNA concentration and initial fragment size distribution to maximize recovery

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

of long fragments. Ethanol washing steps were repeated for samples showing lower purity. After the quality-check step, DNA concentration and fragment size distribution were reassessed to determine whether to proceed with adapter ligation and clean-up. For adapter ligation, incubation with T4 DNA ligase and ligation adapters was extended to 30 minutes. Bead and washing buffer volumes were further adjusted to improve long-fragment retention. Different proportions of Long Fragment Buffer (LFB) and Short Fragment Buffer (SFB) were used during bead washing depending on fragment size profiles. The final quantification step was used to assess DNA concentration and estimate library molarity. A loading range of 10–20 femtomoles of DNA was considered suitable for sequencing. Molarity was calculated from DNA mass assuming a standard average molecular weight per base pair. Each sample was evaluated considering its initial electrophoretic profile. A MinION Mk1C device (Oxford Nanopore Technologies), running MiniKNOW software, was used for sequencing. In parallel, whole-genome sequencing (WGS) was performed for selected accessions using Illumina short-read sequencing to support variant discovery and genotyping analyses.

**Genomics analyses**

Genomic analyses were performed using both long- and short-read sequencing approaches. For ONT long-reads libraries, raw signals from the MinION device were basecalled using Dorado, and FASTQ reads were assessed for quality, read-length distribution and total yield using NanoPack. Reads were filtered to retain those with a minimum Q-score of 8 and a minimum read length of 1,000 bp, and mapped to the *P. persica* 'Lovell' reference genome v2.0 (**Verde et al., 2017**) using *minimap2* (**Li, 2018**). For Illumina short-read libraries, reads were quality-filtered and mapped to the same reference genome using bwa-mem algorithm. Variant calling was performed using FreeBayes v1.3.9, and downstream file handling and filtering steps were carried out using *samtools*, *bcftools* and *vcftools*. All analyses were run using Python v3.7.

**Validation and genotyping of *Md* alleles**

Genomic DNAs were extracted from young leaves by CTAB protocol. Several primer pairs were designed and tested to genotype each *Md* allele. The final primer sets and annealing temperatures are indicated in **Supplementary Table S1**. PCR reactions were carried out using Go-Taq™ in a volume of 20 µl, containing 20 ng of peach genomic DNA. PCR was carried out under the following conditions: $Md_{WT}$ (WT) and $Md_{DP}$ (DP): 2 min at 95 °C, 30 cycles of 30 s at 95 °C, 30 s at 60 °C, and 90 s at 72 °C, followed by a 5 min final extension at 72 °C; $Md_{M}/Md_{WT}$ (M/WT): 2 min at 95 °C, 30 cycles of 30 s at 95 °C, 3 min at 60 °C, and 90 s at 72 °C, followed by a 15 min final extension at 72 °C; $Md_{DP3}$ (DP3): 2 min at 95 °C, 30 cycles of 30 s at 95 °C, 1 min at 60 °C, and 90 s at 72 °C, followed by a 5 min final extension at 72 °C. PCR products were

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

examined using agarose-gel electrophoresis. Primer sets were applied sequentially in the following order: WT, DP, M/WT, and DP3.

**Fruit Sampling and RNA-seq analyses**

To characterise NAC1 transcriptional dynamics during fruit development, an RNA-seq time course was generated using the early- and late-ripening cultivars 'Springcrest' and 'O' Henry'. Fruits were collected at 13, 25, 36, 57, 71, 84, and 96 days after full bloom (DAFB) and at 16, 28, 39, 50, 64, 85, 106, 143, and 157 DAFB, respectively, during growing season 2023. All time points were represented by two or three biological replicates (**Supplemental Table S1**). Mesocarp tissue was immediately dissected, flash-frozen in liquid nitrogen, and stored at −80 °C prior to RNA extraction. Total RNA was extracted using a CTAB-LiCl protocol and treated with DNase I to remove genomic DNA contamination. RNA-seq libraries were prepared using the Illumina TruSeq mRNA protocol. Raw reads were trimmed with TrimGalore (adapter and quality trimming). Transcript abundance was quantified using Salmon (v1) (**Patro et al., 2017**) in quasi-mapping mode, indexing the *P. persica* v2.0 transcriptome with a k-mer size of 31. Expression values were imported into R using *tximport* and summarised to gene level as TPM. Public RNA-seq datasets from additional peach genotypes were also retrieved and processed following the same pipeline (**Supplementary Table S1**). These datasets include annotations of sampling period (DAFB) and developmental stages as reported in the original studies and were used as external validation to assess the generality of observed expression patterns across different *Md* allelic backgrounds. Because datasets differ in sampling schemes and growth conditions, comparative analyses focused on the relative timing of transcriptional activation rather than absolute transcript abundance, allowing a qualitative interpretation of expression dynamics across *Md* genotypes.

**Epigenomic datasets and data processing**

Public epigenomic datasets used in this study are enlisted in **Supplementary Table S1** and include ChIP-seq and DAP-seq experiments targeting the peach transcription factor NAC1, histone modifications (H3K4me3, H3K27ac, H3K27me3), DNase I hypersensitivity assays, and whole-genome bisulphite sequencing (WGBS) data. All datasets were derived from the fruit ENCODE project (**Lu et al., 2018**), except for DAP-seq (**Cao et al., 2023**), and were generated from peach mesocarp tissues. Although cultivar background was heterogeneous, developmental stages are consistently annotated according to peach fruit developmental stages: ChIP-seq and DAP-seq were sampled at S4 stage; H3K4me3, H3K27ac, H3K27me3, DNase I hypersensitivity and WGBS at S1 and S4 stages. Epigenomic signals were analysed in a locus-centred manner, focusing on the genomic region upstream of NAC1 on chromosome 4 (Pp04: 11,118,000–11,145,000 bp). This interval encompasses the entire upstream intergenic region extending from

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

the annotated NAC1 TSS to the neighbouring gene boundary. Chromatin features were examined in their genomic context by assessing spatial co-occurrence patterns between NAC1 binding, chromatin accessibility, histone modification profiles and DNA methylation across developmental stages. When available, biological or technical replicates were averaged, and profiles were summarized by grouping samples into immature (S1) and mature classes (S4). Raw sequencing reads were uniformly processed using standardized pipelines. Adapter sequences and low-quality bases were removed using Trim Galore (v0.6.10) and quality-filtered reads aligned to the peach reference genome 'Lovell' v2.0, For ChIP, DAP, DNase-seq, and histone modifications (H3K4me3, H3K27ac and H3K27me3), reads were aligned using *minimap2*. Signals were summarized as per-base read coverage across the upstream regulatory interval of NAC1 on chr 4 (Pp04: 11,118,000–11,145,000 bp). Coverage profiles were computed using pileup-based counting, smoothed using a Gaussian filter ($\sigma$ = 20-50 bp depending on dataset), and averaged across biological replicates within each assay and developmental stage. For visualization, signals were locally normalized within each chromatin mark to emphasize relative enrichment patterns along the locus. For WGBS, reads were aligned using *Bismark* (**Krueger et al., 2011**) with *Bowtie2* backend (**Langmead and Salzberg, 2012**). Cytosine methylation levels were extracted for CG, CHG and CHH contexts and quantified as the fraction of methylated reads over total reads at each cytosine position, yielding values ranging from 0 (fully unmethylated) to 1 (fully methylated). Custom scripts used for epigenomic data processing and visualization are provided as supplementary codes.

## Results

### MD mapping in DxP progeny

An F1 population of 151 seedlings derived from the cross 'Dulcebo66' (D66) × 'Pulchra' (P) (hereafter referred to as D×P) was phenotyped for maturity date (MD) over two years. D66 is a mid-season peach selection (205 JD), while P is an ultra-early peach (155 JD). In D×P, MD segregated across a range of approximately 50 days (from 150 to 205 JD) with a strong year-to-year correlation. A single-marker QTL analysis was performed separately in the two parents to test associations with MD variation. A major QTL for MD was detected in chromosome (chr) 4 in both parents, although not fully overlapping (**Supplementary Figure S1**). Linkage maps of LG4 covered 47 cM in D66 and 57 cM in P (**Supplementary Table S1**). Interval mapping identified significant QTL peaks at SNP_AO_0443007 (LOD = 38.9, position 10.15 Mb, $R^2$ = 0.69) in P and at SNP_AO_0444920 (LOD = 6.5, position 11.9 Mb, $R^2$ = 0.21) in D66 (**Figure 1A**). Haplotype analysis around the region revealed four allelic classes, indicating additive effects contributed by

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

both parents, with the $D_1P_1$ haplotype associated with earlier maturity (**Figure 1B, Supplementary Table S1**). Recombinant individuals (#063 and #186 in P; #064 and #136 in D66) enabled narrowing the candidate intervals. In D66, the QTL was confined to the 10.7-11.9 Mb region, while in P recombinant genotypes excluded the upstream region up to SNP_IGA_408654 (10.2 Mb). The downstream boundary in P could not be fully resolved due to a large homozygous block with no informative recombination.

This genomic region is already known to harbour a major QTL for MD (*qMD4.1*) and a candidate 9-bp INDEL in the *NAC5* gene, associated with earlier maturity (**Pirona et al., 2013**). Hereafter, the 9-bp and the reference alleles were denoted as $NAC_9$ and $NAC_0$, respectively. D66 was heterozygous $NAC_9/NAC_0$ and P homozygous $NAC_9$. In the D×P progeny, the NAC INDEL only co-segregated with D66-derived haplotypes *D1* and *D2*, indicating that a biallelic model cannot account for the observed segregation pattern and suggesting the presence of additional functional variation at *qMD4.1*, or of a tightly linked QTL contributed by 'Pulchra'.

**Genomic variation at the *Md* locus**

Genomic variations within the mapped region were characterised using 7.25 Gb of Illumina short-reads and 5.74 Gb of ONT long-reads from 'Pulchra', which was heterozygous for the putative causal mutation (assembly statistics in **Supplementary Table S1**). The combination of short- and long-read sequencing provided base-pair level resolution and confirmed the presence of a largely homozygous region extending from 10.27 to approximately 17.0 Mb and several heterozygous SNPs and INDEL variants within the upstream 70 Kb interval (10.21 to 10.27 Mb). Interestingly, annotation revealed a single heterozygous locus within the extended homozygous block, harbouring two complex structural variants (SV) compared to the 'Lovell' v2 reference genome (**Supplementary Figure S2**). Both SVs mapped to approximately the same genomic region between the two NACs, encompassing *Prupe.4G186900*, a predicted long non-coding RNA gene. We referred to this locus as *Md*, assigning specific abbreviations to distinguish the different SVs. The first SV, named $Md_M$, comprises a composite rearrangement occurring at ~11,126,654 (near the 3' UTR of *Prupe.4G186900*) and consisting of an upstream 94-bp tandem duplication and a downstream 413-bp deletion compared to 'Lovell' genome reference (named $Md_{WT}$) (**Figure 2 and Supplementary Figure 3**). The second SV, named $Md_{DP3}$, corresponds to a tandem triplication of the $Md_M$ structural unit, beginning upstream of $Md_M$ rearrangement, at ~11,125,958 bp (including part of the 5' UTR of *Prupe.4G186900*), resulting in three consecutives ~1.5-kb copies (**Figure 2**).

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

The presence of these SVs was further explored in other accessions sequenced with combined short- and long-read approaches. 'Lucrezia', an ultra-early season selection (~145 JD) derived from self-pollination of 'Maycrest' (one of the two parents of 'Pulchra') was homozygous $Md_{DP3}$; 'Springcrest', an early cultivar (~170 JD) from which 'Maycrest' originated as a bud sport, carried an additional SV consisting of a simple duplication of the $Md_M$ unit, and named $Md_{DP}$ (**Figure 2**), being heterozygous $Md_{DP}/Md_{WT}$. Two early-ripening nectarine cultivars, 'Zhong You 4' (ZY4) and its bud sport mutant 'Li Xia Hong' (LXH), which ripens 16 days earlier (**Zhou et al., 2023**) also carried *Md* SVs: ZY4 was heterozygous $Md_{DP}/Md_{WT}$, while LXH homozygous $Md_{DP}$. A PCR assay was developed to detect *Md* SVs (**Supplementary Figure S4**) and used to genotype the D66 parent (which resulted heterozygous $Md_M/Md_{WT}$) and other early-ripening materials, including: 'Mayfire' - one of the earliest ripening nectarines and parent of ZY4 and LXH (**Li et al., 2013**) - and its progenitor 'Armking', as well as 'Springtime', 'Spring Baby' and 'Maycrest': all were homozygous or heterozygous for $Md_{DP}$ except 'Maycrest', which carries $Md_{DP3}$ (**Table 1**).

The recurrent association between earliness and *Md* alleles across different early-ripening genetic backgrounds supports their involvement in determining maturity date.

**Association of *Md* alleles with maturity date across multiple segregating progenies**

To provide an additional layer of validation for the association between *Md* alleles and maturity date, and to assess their relationship with *NAC5* INDEL, we re-evaluated segregation patterns in several populations previously used for *qMD4.1* mapping. The entire progeny (when available) or a subset of individuals for each cross were genotyped to infer phase and co-segregation. Haplotypes, recombinant individuals and statistical analyses across all populations are summarised in **Supplementary Table S1**. Segregation patterns across the different progenies are illustrated in **Figure 3** and **Table 2**.

In D×P population, according to the $Md_{WT}/Md_M \times Md_{DP3}/Md_M$ cross configuration, the $Md_{DP3}$ was linked to the early-ripening $P_1$ haplotype, $Md_M$ to *P2* and *D1*, and $Md_{WT}$ to the late-ripening *D2*, in perfect agreement with the four MD classes. However, the linkage of $NAC_9$ - $Md_M$ and $NAC_0$ - $Md_{WT}$ in D66 parent does not allow to determine whether the observed effect corresponds to a single multi-allelic QTL (*Md*) or to two tightly linked QTLs (*Md* and *NAC*) (**Table 2**).

In 'Big Top' × 'Armking' (Bt×Ak) (**Serra et al., 2017**), the Ak parent was $Md_{DP}/Md_M$ and homozygous $NAC_9$ whereas Bt was homozygous $Md_M$ and heterozygous $NAC_9/NAC_0$, an allelic configuration confirmed through ONT re-sequencing of Bt. Given the close proximity of *NAC* and *Md* (~8 Kbp), the alternative Bt haplotype $NAC_0$ - $Md_M$ provided a highly informative recombinant background for evaluating co-segregation between these loci. As in 'Pulchra', the Ak map was



ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

truncated at ~10.5 Mbp due to a homozygous block of non-segregating SNPs extending to the end of chromosome, with closest recombination occurring at ~9.4 Mb. Consistent with the presence of a QTL only on LG4 of the Ak map, phasing of Ak-derived haplotypes confirmed the association of $Md_{DP}$ with the earlier-ripening class, whereas *NAC* alleles showed no significant effect (**Table 2**).

In 'Big Top' × 'Nectaross' (Bt×Nr), a QTL for MD was previously positioned on LG4 of the Nr map. The Nr parent was heterozygous for $Md_M$ and for the *slow-ripening* allele (*sr*), a ~26 Kb deletion previously described (**Eduardo et al., 2015**). Segregation creates multiple configurations among *Md*, *sr* and *NAC* alleles, but only *sr* and $Md_M$ clearly distinguished early- and late-ripening classes ($Md_M$ = 198, *sr* = 223 JD; $p < 0.01$), without significant effect of NAC (**Table 2**). Thus, segregation was fully consistent with *Md* alleles while the recombination breakpoint observed in 'Big Top' haplotype excludes a causal role for the *NAC5* INDEL variant.

We then re-analysed three $F_2$ populations carrying a major MD QTL on LG4: 'Contender' × 'Ambra' (C×A), 'Earlygold' self-pollinated (E×E) and 'NJ Weeping' × 'Bounty' (W×By) (**Pirona et al., 2013; Kalluri et al., 2022**). The $F_1$ C×A parent was heterozygous $Md_{DP}/Md_{WT}$, because inherited the $Md_{DP}$ allele from 'Mayfire' (the male parent of 'Ambra'), 'Earlygold' heterozygous $Md_{DP}/Md_M$ and the $F_1$ W×By heterozygous $Md_M/Md_{WT}$, all in agreement with the observed segregations. In these progenies, the perfect linkage between $NAC_9$ and $Md_{DP}$ or $Md_M$ (and between $Md_{WT}$ and $NAC_0$) likely masked the identification of the *Md* locus. Notably, the different SVs combinations coherently match the distribution of ripening classes observed in the three progenies, from ultra-early individuals (homozygous $Md_{DP}$, in E×E) to late ones (homozygous $Md_{WT}$ in C×A and W×By) (**Figure 3**). PCR assays also identified several recombinant individuals in E×E (**Supplementary Table S1**). Moreover, heterozygous individuals displayed partial dominance, particularly evident for $Md_{DP}$ (versus either $Md_M$ or $Md_{WT}$ in E×E and C×A, respectively) and to a lesser extent for $Md_M$ *vs.* $Md_{WT}$ (in W×By), causing heterozygous to shift towards the earlier allele.

Finally, in $F_1$ 'Max10' × 'Rebus28' (M × R) (**Ciacciulli et al., 2018**), M carried the $Md_{WT}/sr$ genotype while R was $Md_{DP}/Md_M$, a configuration fully consistent with the four observed MD classes (**Figure 3**). However, R (originated from 'Big Top' x 'Mayfire') inherited the alternative $Md_M$ haplotype from Bt, which results in partial linkage between *Md* and *NAC* for three of the four allelic classes. This recombinant background explains why the *Md-NAC* association is maintained only for a subset of genotype combinations (as also reflected in the segregation patterns and linkage notes in **Table 2**), while being disrupted in others.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

In summary, all mapped QTLs across the different populations share a common genomic segment encompassing the *Md* locus, whose multiple alleles consistently co-segregate with MD variation, while recombination events in some genetic backgrounds exclude a causal role for the 9-bp INDEL mutation.

**Association of *Md* alleles with maturity date in germplasm collections**

To further substantiate the association observed across segregating populations, we realigned the assembled genomes of all parental lines to identify additional polymorphisms in linkage with *Md* alleles. By integrating QTL evidence from independent families with parental genome realignment and haplotype phasing, we were able to refine the *Md* locus and retain only variants truly compatible with the observed segregation patterns. Phase information from these haplotypes was also used to examine linkage patterns in other materials (cultivars and breeding selections) with known pedigrees, using founder haplotypes as reference.

Within the $Md_{DP}$ interval, fine-mapped to 10.88-12.10 Mbp in 'F1 CxA', the parents 'Armking', 'Rebus28' and 'Earlygold' were homozygous across the entire region except for the heterozygous $Md_{DP}/Md_{M}$ allele. For $Md_{DP3}$, an informative recombination in the ultra-early selection 'Borgia' (from selfing of 'Maycrest'), delimited the downstream boundary at 11.3 Mb (**Supplementary Table S1**). Likewise, across this region, 'Pulchra' was homozygous except for the $Md_{DP3}$ allele. For $Md_{M}$, the interval was relatively narrow (about 200 Kb, from 10.98-11.20 Mbp in WxBy) and contained (aside from $Md_{M}$) only a limited number of variants consistent with segregation, including the *NAC5* INDEL, already excluded as causal (**Supplementary Table S1**). Collectively, genomic data reinforce the presence of a single multi-allelic causal locus at *qMD4.1*.

Given the abundance of publicly available short-reads datasets, we tested a simplified approach to assess the detectability of Md SVs using short-reads, which often fail to resolve complex structural rearrangements. In our case, standard alignment to the reference genome allows reliable detection of $Md_{M}$ allele, due to the presence of ~94 bp and ~400 bp length polymorphisms relative to $Md_{WT}$. However, this approach does not allow discrimination of $Md_{DP}$, because tandem duplications do not generate an obvious loss of coverage or unique mapping pattern when reads are aligned to the reference sequence. To overcome this limitation, (i) the full-length $Md_{WT}$ sequence and (ii) the duplication junction were used as BLAST queries against short-read libraries. Using this targeted approach, $Md_{M}$, $Md_{WT}$ and $Md_{DP}$ could be reliably distinguished, even in heterozygosity (**Supplementary Figure S5**). Nevertheless, $Md_{DP}$ could not be distinguished from $Md_{DP3}$ based on short reads alone, as they share an identical duplication junction (although the latter is present only in progenies derived from 'Maycrest'). Detection of *sr*,



ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

instead, required whole-genome alignment, where it can be inferred from a local reduction in read coverage.

Using publicly available WGS data and phenotypic records, we further assessed the association between *Md* alleles and maturity date in a broad set of accessions (**Supplementary Table S1**). Despite some residual variability, likely due to segregation of other QTLs, *Md* alleles clearly and additively separate maturity classes from ultra-early to late, collectively explaining nearly the full harvesting calendar in peach germplasm (140 to 250 JD) (**Figure 4**). In predictive comparison, *Md* SVs ($R^2$ = 75.4%, MAE = 9.4) outperformed both the *NAC* ($R^2$ = 48.6%, MAE = 14.1) and the most strongly associated SNP on the peach array ($R^2$ = 20.6%, MAE = 19.06).

We also explored the entire *NAC1* upstream region in search of additional candidate polymorphisms. Aside from a previously reported 232-bp deletion located ~500 bp upstream of the transcription start site, which does not co-segregate with MD (**Pirona et al., 2013**), no other variants showed patterns consistent with a functional role in maturity variation.

**Functional evidence of the effects of *Md* structural variants on *NAC1* expression**

To investigate the transcriptional dynamics of the *NAC1* gene, fruits of 'Springcrest' (SP) ($Md_{DP}/Md_{WT}$) and 'O' Henry' ($Md_{WT}/Md_{WT}$) were sampled at multiple developmental stages, from 20 days after full bloom (DAFB) until ripening, occurring at 96 and 157 DAFB, respectively. In both genotypes, *NAC1* transcript accumulation displayed an exponential trajectory culminating in a distinct peak at the onset of fruit ripening. However, the expression peak appeared earlier in SP, consistent with the presence of the $Md_{DP}$ allele (**Figure 5**). To extend this observation, we analysed publicly available RNA-seq datasets for additional genotypes carrying distinct *Md* alleles (**Supplementary Table S1**). Those harbouring the $Md_{DP}$ allele (e.g., 'LXH', 'ZY4', and 'Jin Chun') exhibited an accelerated rise in *NAC1* expression, reaching the maturation-associated peak sooner than $Md_M$ or $Md_{WT}$ backgrounds. Although overall expression levels varied among genotypes, the timing of *NAC1* expression was consistently shifted according to allelic composition, indicating a dosage-dependent modulation of its transcription during fruit development. Conversely, *NAC5* expression was less clear, displaying strong genotype-dependent variability and no coherent temporal pattern, although transcript levels were generally higher at early developmental stages. *NAC1* transcription was markedly reduced in a slow-ripening genotype homozygous for the *sr* deletion, which spans *NAC5* and extends into the upstream regulatory region of *NAC1*, including the *Md* locus (**Supplementary Table S1**). In parallel, we noted that the adjacent gene *Prupe.4G186900*, likely encoding a long non-coding RNA, is transcriptionally active (albeit at low levels) predominantly in $Md_{DP}$ genotypes, including

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

SP, LXH, ZY4 and JC. By contrast, this transcript is absent or detected only sporadically at late maturation stages in *Md*$_{WT}$ or *Md*$_{M}$ backgrounds, suggesting a possible link between allele-specific regulatory structure and lncRNA expression. Also, the other transcript annotated in the region (*Prupe.4G187000*) was barely detectable.

Publicly available epigenomic datasets were explored to obtain a qualitative view of the regulatory landscape at the Md locus during fruit development. Although derived from different genotypes and experimental contexts, these datasets collectively provide a useful framework for interpreting the chromatin landscape in this region and for identifying regulatory features associated with the Md locus. They include ChIP-seq and DAP-seq experiments targeting the transcription factor NAC1, profiles of histone modifications (H3K4me3, H3K27ac and H3K27me3), DNase I hypersensitivity assays and bisulphite sequencing, all generated from peach mesocarp tissue sampled at immature and mature fruit stages (**Figure 5 and Supplementary Table S1**). ChIP-seq data from 'Babygold' (*Md*$_{M}$/*Md*$_{WT}$) revealed two NAC1 regions of binding enrichment: one proximal to the coding sequence (~11.138 Mb) and a second within the upstream regulatory regions (~11.126 Mb), coinciding with the insertion site of the *Md* structural variants. Independent DAP-seq analysis in 'Hu Jing Mi Lu' (*Md*$_{M}$/*Md*$_{WT}$) confirmed a strong peak at the same upstream position, supporting the presence of a conserved regulatory interaction. Interestingly, the *Md*$_{M}$ rearrangement generates a duplication of the putative *NAC1* binding motif ACG(T/C)(A/C) in this region, suggesting a potential modulation of autoregulatory interaction. Beyond NAC1 binding, chromatin features display hallmarks of regulatory activity, including DNase I hypersensitivity, enrichment of activating histone marks (H3K4me3 and H3K27ac), and a relative depletion of repressive mark H3K27me3. DNA methylation profiles show a locally structured pattern with alternating hypo- and hypermethylated segments, and a reproducible hypomethylated valley around the *Md* insertion site. No major qualitative differences were observed between immature and ripe fruit stages, indicating that this region maintains a permissive chromatin state throughout development.

**Origin and spread of *Md* alleles**

In order to explore and trace the origin of the various *Md* SVs, genome sequences of wild peach relatives were investigated. *Md*$_{WT}$ sequence is substantially conserved across species, except for several SNPs scattered throughout the region and a few INDELs, the most notable of which is approximately 30 bp long and present only in peach, *P. kansuensis* and apricot (**Supplementary Figure S6**). Within peach, *Md*$_{M}$ is widespread in many Chinese landraces as well as in European and North American accessions. More intriguing is the case of the early-



ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

ripening *Md*$_{DP}$, whose introduction traces back to the old cultivar 'Mayflower', of uncertain origin and likely identified around 1860 in North Carolina (**Hedrick, 1917**). Analyses of the 'Mayflower' accession (CREA-Rome collection) confirmed the presence of the *Md*$_{DP}$, as well as in breeding materials initially developed by Armstrong Nurseries (Ontario, California) and University of Florence (Italy): 'Robin' (from Babcock × Mayflower), heterozygous *Md*$_{DP}$/*Md*$_{WT}$, and 'Morettini1' (from Cumberland × Mayflower, 1938) (**Figure 6**). 'Robin' was the founder of important early varieties such as 'Springtime' (1953), 'Earlygold' (1958) and 'Maytime' (1958), followed later by the first early nectarine 'Armking' (1969) (**Okie, 1998**) In turn, 'Springtime' was the parent of 'Springcrest' (1969), but also of 'Springold' (1966), 'Flordasun' (1969) and 'Royal Gold', a bud sports isolated by C. F. Zaiger (1967). According to pedigree information, 'Maycrest' originated as a bud sport of Springcrest discovered in 1977 and ripening 3 - 5 days earlier. Our genomic analyses show that only 'Maycrest' carries *Md*$_{DP3}$, representing an additional copy-number expansion of the *Md*$_{M}$-derived structural unit. This suggests that the bud sport event leading to 'Maycrest' likely involved a further duplication event at the *Md* locus. This is consistent with a somatic duplication event in the bud-sport lineage. However, whether *Md*$_{DP3}$ systematically confers a measurable phenological shift relative to *Md*$_{DP}$ will require validation through controlled crosses and segregation analyses.

## Discussion

### Allelic determinants of maturity date in peach

The identification of causal variants underlying key agronomic traits is crucial to accelerate breeding progress in perennial fruit trees. In peach, maturity date (MD) is particularly important for defining harvest calendar and market windows, and for its pleiotropic effects on internal and external fruit quality and postharvest behaviour. Consequently, numerous studies have aimed to characterize the genetic architecture of MD, identifying major QTLs associated with early or late maturation in both biparental segregating populations and germplasm collections. The central region of chromosome 4 has repeatedly emerged as an MD hotspot due to the frequency with which QTLs have been mapped there and the large phenotypic variance they explain across diverse genetic backgrounds (**Quilot et al., 2004**; **Dirlewanger et al., 2012; Romeu et al., 2014**; **Hernández Mora et al., 2017**; **Rawandoozi et al., 2021**; **da Silva Linge et al., 2021**), segregating as a Mendelian trait in certain progenies such as CxA (Eduardo et al. 2011). A 9-bp INDEL located in the third exon of the *NAC5* gene was initially proposed as a candidate causal variant for MD in peach (**Pirona et al., 2013**). Association studies have consistently shown that this variant

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

provides reasonable discriminatory power across different germplasm sets, though showing substantial phenotypic variability within genotypic classes (**Cao et al., 2023**).

In this study, we mapped a large-effect locus for MD to the central region of chromosome 4 in both parents of the DxP population. The segregation pattern could not be explained solely by the 9-bp INDEL, suggesting the presence of additional functional alleles at *qMD4.1* or a tightly linked QTL. The presence of a large fully homozygous IBD-like region in the P parent (donor of the early-maturity allele) enabled the identification of a structural variant in the region (~11.126 Mb) between *NAC5* and *NAC1* consistent with the observed segregation. By expanding the analysis to important breeding and germplasm materials, we identified a multi-allelic series at the *Md* locus, comprising the reference allele $Md_{WT}$ ('Lovell') and three structurally distinct derived alleles ($Md_{M}$, $Md_{DP}$ and $Md_{DP3}$).

Reassessment of multiple segregating populations previously used for MD QTL mapping clearly showed a consistent co-segregation of *Md* allelic series with maturity date, predicting phenotypes from ultra-early to late ripening. The allelic effects were generally additive, with partial dominance particularly evident in the early alleles $Md_{DP3}$ and $Md_{DP}$. Identification of an alternative haplotypes in 'Big Top' allowed genetic decoupling of *Md* from *NAC* variant, confirming the former as the causal locus at *qMD4.1*.

In addition to *Md* alleles, a series of large independent deletions (~26 kb) encompassing the entire *NAC5* gene and most of the *NAC1* promoter have been associated with a slow-ripening (*sr*) trait when homozygous, and delayed maturation when heterozygous (**Eduardo et al., 2015; Nuñez Lillo et al., 2015**). The *sr* phenotype is characterized by a block in the ripening process, resulting in fruits that fail to mature properly (**Brecht et al., 1984**). Recently, palindromic sequences flanking the deletion have been implicated in this phenomenon in a de novo somatic mutant (**Pietrella et al., 2025**). The *sr* allele helps to complement the *Md* allelic series and provides a mechanistic explanation for the full spectrum of phenotypic variation observed in segregating populations, accounting for late- or atypical-ripening classes in certain families (e.g., 'Bt×Nr' and 'M×R'). We also assessed the association of *Md* alleles with MD leveraging publicly available germplasm data, confirming their strong predictive power despite the likely segregation of additional QTLs.

Furthermore, we reconstructed the evolutionary and breeding history of *Md* alleles. The distribution of $Md_{DP}$ across germplasm and breeding materials appears recent and geographically limited compared with $Md_{WT}$ and $Md_{M}$, which are widespread in traditional European, American, and Chinese germplasm. Comparisons with wild relatives indicate that the $Md_{WT}$ allele is ancestral and highly conserved (though with some polymorphisms) even in more distantly related *Prunus*

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

species such as apricot. Pedigree analyses of major North American germplasm suggests that the early-ripening *Md*$_{DP}$ allele most likely entered modern breeding through the historical accession 'Mayflower' - also known under synonyms such as 'Fior di Maggio' (Italy) and 'Avant Pêche' (France) (**Okie et al., 1998**). This accession is genetically closely related to other 19$^{th}$-century early-ripening cultivars, including 'Amsden' (selected by L.C. Amsden in Missouri, 1868) and 'Ribet' (**Morettini and Bellini, 1971**). Within this genealogical background, the emergence of *Md*$_{DP3}$ in 'Maycrest' represents a more recent copy-number expansion derived from the *Md*$_{DP}$ haplotype carried by 'Springcrest'. Although *Md*$_{DP3}$ segregates with ultra-early ripening, whether it confers an effect distinct from *Md*$_{DP}$ under equivalent genetic backgrounds has not yet been formally demonstrated. The extensive IBD block shared among several *Md*$_{DP}$-bearing cultivars is consistent with the apparent monophyletic origin of this allele and reflects strong selection for earliness across decades of breeding. Such a combination of single-origin introgression and directional selection has resulted in a highly conserved haplotype surrounding the *Md*$_{DP}$ allele, which persists across multiple independent breeding lines.

**Functional evidence and transcriptional dynamics of *Md* locus**

Structural variation at the *Md* locus highlights a link between upstream regulatory architecture and NAC1 transcriptional control during fruit maturation. In particular, the structural rearrangements associated to *Md* alleles overlap an upstream region that is bound by NAC1 in both ChIP-seq and DAP-seq experiments. The co-occurrence of NAC1 binding with open chromatin, enrichment of activating histone marks and local DNA hypomethylation indicates that this interval may represent an active regulatory region. These observations are compatible with a model in which NAC1 may engage in upstream autoregulatory interactions within a permissive chromatin environment, contributing to the fine-tuning of its transcript accumulation in an allele-dependent manner. Similar upstream autoregulatory architectures have been documented in apple for a NAC transcription factor (MdNAC18.1) (**Zhang et al., 2025**). Previous studies have examined the relative contributions of *NAC5* and the closely linked *NAC1* to peach fruit development and ripening. Based on their expression patterns, the two genes appear to perform distinct functions, displaying different tissue specificities and temporal dynamics. *NAC5* shows broad tissue expression, with highest transcript abundance in flowers, roots, phloem, fruits, seeds, and leaves (**Cao et al., 2023**). In fruits, *NAC5* is predominantly expressed at immature stages, and its abundance decreases as ripening progresses. Moreover, when comparing expression profiles between mutant sports differing in ripening time and their donor genotypes, *NAC5* does not consistently show differential expression (**Cao et al., 2023; Zhou et al., 2023**). While *NAC5*

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

is excluded as the causal determinant at *qMD4.1*, it may still contribute to other aspects of fruit development or ripening-related processes. In contrast, *NAC1* displays fruit-specific expression, peaking at the onset of climacteric ripening. Our results extend these observations by showing that the timing of the NAC1 expression peak is allele-dependent: $Md_{DP}$ alleles are associated with an earlier *NAC1* peak, $Md_{M}$ alleles with intermediate timing, and $Md_{WT}$ alleles with later gene expression. In this context, the strong reduction of NAC1 transcription observed in the slow-ripening background may represent one of several genetic mechanisms contributing to the modulation of its expression.

Studies on NAC downstream genes in peach are limited, but yeast one-hybrid and dual-luciferase assays showed that NAC1 is able to interact with promoter sequences of key ripening genes, including those involved in ethylene-biosynthesis (*ACS1*, *ACO1*), cell-wall modification (*PG1*, *PME1*, *PL1*) or aroma production (**Jin et al., 2022; Cao et al., 2023; Zhang et al., 2024**). NAC1 binding to NAC-motif-rich regulatory regions of ripening genes is consistent with a direct transcriptional activation mechanism (**Cao et al., 2023**). By contrast, NAC5 has far fewer validated targets. Recent evidence shows that NAC5/NAC1 tandem can activate *PGF*, a gene involved in peach softening, but its broader regulatory network remains unclear (**Zhang et al., 2024**).

Peach remains recalcitrant to genetic transformation, limiting species-specific functional validation. Heterologous complementation experiments in tomato have been widely used, although they provide contrasting evidence. **Cao et al. (2023)**, using 'Ailsa Craig', reported no clear phenotypic changes in *NAC5*-overexpression lines (with or without the 9-bp insertion), whereas those over-expressing *NAC1* caused fruit to turn red about 6 days earlier than wild type. By contrast, **Zhang et al. (2023)** reported that overexpression of both genes accelerated ripening in the same tomato background, although *NAC1* produced a more pronounced effect. Overall, *NAC1* is emerging as core regulator of the ripening network in peach, whereas *NAC5* appears to play an auxiliary role in developmental (or stress-related) processes.

Additional loci influencing fruit maturation in peach represent genetic mechanisms distinct from the *Md* locus but acting within ripening contexts in which NAC1 may be functionally implicated. A clear example is the DBF locus, which controls anthocyanin accumulation in blood-fleshed peaches, and is also associated to advanced ripening. Functional analyses have shown that the DBF-encoded NAC transcription factor (BL) can interact with NAC1 and promote the expression of ripening-related genes (**Wang et al., 2024**). These observations indicate that NAC1 may operate within a broader regulatory network in which distinct genetic factors converge on common transcriptional regulators to modulate the timing and progression of fruit ripening.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

## Breeding implications and marker-assisted selection

We developed a set of diagnostic markers capable of discriminating the four structural variants at the *Md* locus. Together with existing *sr* markers, they provide a comprehensive and reliable genotyping toolkit for predicting MD in peach. Given the near-Mendelian effect of *Md* alleles (as clearly shown in **Figure 4** in the germplasm panel), this marker set enables accurate classification of genotypes into distinct ripening classes and provides a strong basis for marker-assisted selection. From an operational breeding perspective, the availability of these markers allows early prediction of maturity timing, facilitating rapid exclusion of undesired ripening classes and more efficient allocation of resources. By directly capturing the full allelic complexity of a major-effect locus, *Md*-targeted selection offers a robust and transferable alternative to SNP-based assays. At present, *Md* genotyping relies on PCR amplicon-based assays that are not readily scalable to high-throughput SNP platforms. Consistently, the SNP most strongly associated with $Md_{DP/DP3}$ (Peach_AO_0443007) showed limited predictive power, particularly in nectarines, and failed to recapitulate the multi-allelic structure of *Md* (**Supplementary Figure S7**). Thus, while *Md*-targeted genotyping offers a practical and highly predictive MAS tool, its widespread implementation will require routine use of specific primers and future development of more scalable assays capable of capturing the full allelic complexity of this locus.

## Conclusions

Our results identify *Md* as the major causal locus underlying *qMD4.1* in peach, with a multi-allelic series of SVs in the upstream regulatory region that appear to influence the timing and amplitude of *NAC1* transcription. From a breeding standpoint, *Md* genotyping provides a powerful predictor of MD, far outperforming either *NAC5* indel or SNP-based assays that cannot capture the underlying allelic complexity. The diagnostic primer set developed provides a practical tool for marker-assisted selection and design of tailored harvest windows. Although genetic, structural, and transcriptional evidence converges on *Md*, definitive functional validation will require allele-swap or regulatory-editing experiments (for example, introducing $Md_{DP}$ upstream regulatory region into the wild-type background, and reciprocally, via CRISPR-mediated promoter replacement). Such assays would directly test whether the structural variants alone are sufficient to alter *NAC1* transcriptional dynamics and ripening time. These experiments, along with the characterization of additional minor MD loci and potential G×E effects, represent the next steps towards a more complete mechanistic understanding of fruit maturity in peach.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

**Acknowledgements**

We thank S. Foschi, M. Lama and M. Samorè for their contribution to orchard management. This work was supported by the project MAS.PES: apricot and peach breeding by molecular-assisted selection.

**Funding Statement**

The study was carried out within the Agritech National Research Center and funded by the European Union – NextGenerationEU under the National Recovery and Resilience Plan (PNRR, Mission 4, Component 2, Investment 1.4; CN00000022). Additional funding was provided by the Italian Ministry of University and Research (MUR) under the National Recovery and Resilience Plan (Mission 4, Component 2, Investment 1.1; Project IMPEACHMENT, CUP J53D23006820006). Open access funding was provided by the University of Milan within the CRUI-CARE Agreement.

**Author contributions**

T.A.G, F.J. conducted the investigation. B.G, C.R., B.I., D.S.L.C., C.E., Z.N., G.S. performed the analyses. M.C. and T.A.G wrote the manuscript. I.E., B.D and S.M. provided data and plant materials. M.C. and L.R. conceived and designed the study, acquired the funding and revised the manuscript. All authors contributed to data interpretation.

**Competing interests**

The authors declare no competing interests

**Data availability**

All data supporting the findings of this study are available in the NCBI Sequence Read Archive PRJNA1435823. Custom scripts used for data processing, analysis and visualization are available at the GitHub repository: https://github.com/marcoc83/peachMD

## References

Baccichet I, Patuelli F, Pozzetto B, Tagliabue AG, da Silva Linge C, Biffi G, Calastri E, Chiozzotto R, Bassi D, Sgarbi P et al. 2025. Peach fruit quality on the fresh market: a long-term evaluation series from Northern Italy is pushing the overturn of commercial standards within the mass-supply chain. *Postharvest Biology and Technology* 224, 113460.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

Bassi D, Gambardella M, Negri P. 1988. Date of ripening and two morphological fruit traits in peach progenies. *Acta Horticulturae* 254, 59–66.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635.

Brecht JK, Kader AA, Ramming DW. 1984. Description and postharvest physiology of some slow-ripening nectarine genotypes. *Journal of the American Society for Horticultural Science* 109, 596–600.

Brummell DA, Dal Cin V, Crisosto CH, Labavitch JM. 2004. Cell wall metabolism during maturation, ripening and senescence of peach fruit. *Journal of Experimental Botany* 55, 2029–2039.

Cao K, Pan H, Zhao Y, Bie H, Wang J, Zhu G, Wang L. 2023. Discovery of a key gene associated with fruit maturity date and analysis of its regulatory pathway in peach. *Plant Science* 330, 111735.

Cao X, Li X, Su Y, Zhang C, Wei C, Chen K, Grierson D, Zhang B. 2024. Transcription factor PpNAC1 and DNA demethylase PpDML1 synergistically regulate peach fruit ripening. *Plant Physiology* 194, 2049–2068.

Ciacciulli A, Cirilli M, Chiozzotto R, Attanasio G, da Silva Linge C, Pacheco I, Rossini L, Bassi D. 2018. Linkage and association mapping for the slow softening (SwS) trait in peach (Prunus persica L. Batsch) fruit. *Tree Genetics and Genomes* 14, 93.

Cirilli M, Baccichet I, Chiozzotto R, Silvestri C, Rossini L, Bassi D. 2021. Genetic and phenotypic analyses reveal major quantitative loci associated to fruit size and shape traits in a non-flat peach collection (Prunus persica L. Batsch). *Horticulture Research* 8, 232.

Crisosto CH, Johnson RS, DeJong TM, Day KR. 1997. Orchard factors affecting postharvest stone fruit quality. *HortScience* 32, 820–823.

Dardick C, Callahan AM. 2014. Evolution of the fruit endocarp: molecular mechanisms underlying adaptations in seed protection and dispersal strategies. *Frontiers in Plant Science* 5, 284.

da Silva Linge C, Cai L, Fu W, Clark J, Worthington M, Rawandoozi Z, Byrne DH, Gasic K. 2021. Multi-locus genome-wide association studies reveal fruit quality hotspots in the peach genome. *Frontiers in Plant Science* 12, 644799.

da Silva Linge C, Fu W, Calle A, Rawandoozi Z, Cai L, Byrne DH, Worthington M, Gasic K. 2024. Ppe.RPT/SSC-1: from QTL mapping to a predictive KASP test for ripening time and soluble solids concentration in peach. *Scientific Reports* 14, 1453.

Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet C. 1999. Mapping QTLs controlling fruit quality in peach. *Theoretical and Applied Genetics* 98, 18–31.

Dirlewanger E, Quero-García J, Le Dantec L, Lambert P, Ruiz D, Dondini L, Illa E, Quilot-Turion B, Audergon JM, Tartarini S, et al. 2012. Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three Prunus species: peach, apricot and sweet cherry. *Heredity* 109, 280–292.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

Eduardo I, Pacheco I, Chietera G, Bassi D, Pozzi C, Vecchietti A, Rossini L. 2011. QTL analysis of fruit quality traits in two peach intraspecific populations and importance of maturity date pleiotropic effect. *Tree Genetics and Genomes* 7, 323–335.

Eduardo I, Picañol R, Rojas E, Batlle I, Howad W, Aranzana MJ, Arús P. 2015. Mapping of a major gene for the slow ripening character in peach: co-location with the maturity date gene and development of a candidate gene-based diagnostic marker for its selection. *Euphytica* 205, 627–636.

Etienne C, Rothan C, Moing A, Plomion C, Bodénès C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E. 2002. Candidate genes and QTLs for sugar and organic acid content in peach (Prunus persica L. Batsch). *Theoretical and Applied Genetics* 105, 145–159.

Fresnedo-Ramírez J, Bink MCA, van de Weg E, Famula TR, Crisosto CH, Frett TJ, Gradziel TM. 2015. QTL mapping of pomological traits in peach and related species breeding germplasm. *Molecular Breeding* 35, 166.

Giovannoni JJ, Tanksley SD, Vrebalov J, Noensie E. 2004. NOR gene for use in manipulation of fruit quality and ethylene response. United States Patent 5,234,834.

Hedrick UP. 1917. The peaches of New York. Albany, NY: J.B. Lyon Company, 541 pp.

Hernández Mora JR, Micheletti D, Bink M, van de Weg E, Cantín C, Nazzicari N, Caprera A, Dettori MT, Micali S, Banchi E, et al. 2017. Integrated QTL detection for key breeding traits in multiple peach progenies. *BMC Genomics* 18, 404.

Jin Z, Wang J, Cao X, Wei C, Kuang J, Chen K, Zhang B. 2022. Peach fruit PpNAC1 activates PpFAD3-1 transcription to provide omega-3 fatty acids for the synthesis of short-chain flavor volatiles. *Horticulture Research* 9, uhac085.

Kalluri N, Eduardo I, Arús P. 2022. Comparative QTL analysis in peach 'Earlygold' F2 and backcross progenies. *Scientia Horticulturae* 293, 110726.

Kang M, Chanderbali A, Lee S, Soltis DE, Soltis PS, Kim S. 2023. High-molecular-weight DNA extraction for long-read sequencing of plant genomes: an optimization of standard methods. *Applications in Plant Sciences* 11, e11528.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* 27, 1571–1572.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100.

Li XW, Meng XQ, Jia HJ, Yu ML, Ma RJ, Wang LR et al. 2013. Peach genetic resources: diversity, population structure and linkage disequilibrium. BMC Genetics 14, 84.

Li H, Chen Z, Zhu W, Ni X, Wang J, Fu L, Chen J, Li T, Tang L, Yang Y et al. 2024. The MaNAP1–MaMADS1 transcription factor module mediates ethylene-regulated peel softening and ripening in banana. The Plant Cell 37, koae282.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

Lombardo VA, Osorio S, Borsani J, Lauxmann MA, Bustamante CA, Budde CO, Andreo CS, Lara MV, Fernie AR, Drincovich MF. 2011. Metabolic profiling during peach fruit development and ripening reveals the metabolic networks that underpin each developmental stage. *Plant Physiology* 157, 1696–1710.

Lü P, Yu S, Zhu N, Chen RY, Zhou B, Pan Y, Tzeng D, Fabi JP, Argyris J, Garcia-Mas J, et al. 2018. Genome encode analyses reveal the basis of convergent evolution of fleshy fruit ripening. *Nature Plants* 4, 784–791.

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB. 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nature Genetics 38, 948–952.

Martín-Pizarro C, Vallarino JG, Osorio S, Meco V, Urrutia M, Pillet J, Casañal A, Merchante C, Amaya I, Willmitzer L, et al. 2021. The NAC transcription factor FaRIF controls fruit ripening in strawberry. *The Plant Cell* 33, 1574–1593.

Morettini A, Bellini E. 1971. Le cultivar precocissime di pesco. Rivista di Orticoltura, Floricoltura e Frutticoltura Italiana 55, 3–44.

Nuñez-Lillo G, Cifuentes-Esquivel A, Troggio M, Micheletti D, Infante R, Meneses C. 2015. Identification of candidate genes associated with mealiness and maturity date in peach (Prunus persica L. Batsch) using QTL analysis and deep sequencing. *Tree Genetics and Genomes* 11, 86.

Okie WR. 1998. Handbook of peach and nectarine varieties: performance in the southeastern United States and index of names. United States Department of Agriculture, Agriculture Handbook 714.

Osorio S, Alba R, Damasceno CMB, Lopez-Casado G, Lohse M, Zanor MI, Tohge T, Usadel B, Rose JKC, Fei Z, al. 2011. Systems biology of tomato fruit development: combined transcript, protein, and metabolite analysis of tomato transcription factor (nor, rin) and ethylene receptor (Nr) mutants reveals novel regulatory interactions. Plant Physiology 157, 405–425.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods 14, 417–419.

Pavel EW, DeJong TM. 1993. Relative growth rate and its relationship to compositional changes of nonstructural carbohydrates in the mesocarp of developing peach fruits. Journal of the American Society for Horticultural Science 118, 503–508.

Pietrella M, Ferrari L, Dondini L, Alessandri S, Pondini L, Tasini F, Brandi F, Sirri S, Giuliano G, Eduardo I, et al. 2025. A palindromic sequence on chromosome 4 of peach (Prunus persica L. Batsch) is flanking a large deletion related to a slow ripening phenotype. Scientia Horticulturae 353, 114494.

Pirona R, Eduardo I, Pacheco I, da Silva Linge C, Miculan M, Verde I, Tartarini S, Dondini L, Pea G, Bassi D et al. 2013. Fine mapping and identification of a candidate gene for a major locus controlling maturity date in peach. *BMC Plant Biology* 13, 166.

Quilot B, Wu BH, Kervella J, Génard M, Foulongne M, Moreau K. 2004. QTL analysis of quality traits in an advanced backcross between Prunus persica cultivars and the wild relative species P. davidiana. Theoretical and Applied Genetics 109, 884–897.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

Rasori A, Ruperti B, Bonghi C, Tonutti P, Ramina A. 2002. Characterization of two putative ethylene receptor genes expressed during peach fruit development and abscission. Journal of Experimental Botany 53, 261–271.

Rawandoozi ZJ, Hartmann TP, Carpenedo S, Gasic K, da Silva Linge C, Cai L, Van de Weg E, Byrne DH. 2021. Mapping and characterization QTLs for phenological traits in seven pedigree-connected peach families. *BMC Genomics* 22, 187.

Reighard GL, Souza FBM, Pio R. 2017. Peach cell number and size is affected by crop load and cultivar. Acta Horticulturae 1160, 357–359.

Romeu JF, Monforte AJ, Sánchez G, Granell A, García-Brunton J, Badenes ML, Ríos G. 2014. Quantitative trait loci affecting reproductive phenology in peach. BMC Plant Biology 14, 52.

Serra O, Giné-Bordonaba J, Eduardo I, Bonany J, Echeverria G, Larrigaudière C, Arús P. 2017. Genetic analysis of the slow-melting flesh character in peach. Tree Genetics and Genomes 13, 77.

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS, Folta KM, Iezzoni A, Main D, Arús P, Dandekar AM et al. 2008. Multiple models for Rosaceae genomics. Plant Physiology 147, 985–1003.

Souza F, Alves E, Pio R, Castro E, Reighard G, Freire AI, Mayer NA, Pimentel R. 2019. Influence of temperature on the development of peach fruit in a subtropical climate region. Agronomy 9, 20.

Tonutti P, Bonghi C, Drincovich MF, Trainotti L. 2023. Peach fruit growth, development, ripening and postharvest physiology. In: Peach. Wallingford, UK: CABI, 148–174.

Trainotti L, Tadiello A, Casadoro G. 2007. The involvement of auxin in the ripening of climacteric fruits comes of age: the hormone plays a role of its own and has an intense interplay with ethylene in ripening peaches. Journal of Experimental Botany 58, 3299–3308.

Van Ooijen JW. 2006. JoinMap 4: software for the calculation of genetic linkage maps in experimental populations. Wageningen, Netherlands: Kyazma BV.

Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, Paris R, Aramini V, Gazza L, Rossini L, et al. 2017. The peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics* 18, 225.

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J. 2002. A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science* 296, 343–346.

Yamaguchi M, Haji T, Miyake M, Yaegaki H. 2002. Varietal differences in cell division and enlargement periods during peach (Prunus persica Batsch) fruit development. Journal of the Japanese Society for Horticultural Science 71, 155–163.

Yue Q, Xie Y, Yang X, Zhang Y, Li Z, Liu Y, Cheng P, Zhang R, Yu Y, Wang X, et al. 2024. An indel variant in the promoter of the NAC transcription factor MdNAC18.1 plays a major role in apple fruit ripening. The Plant Cell 37, koaf007.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

Wang R, Lammers M, Tikunov Y, Bovy AG, Angenent GC, de Maagd RA. 2020. The rin, nor and Cnr spontaneous mutations inhibit tomato fruit ripening in additive and epistatic manners. Plant Science 294, 110436.

Wang J, Cao K, Li Y, Wu J, Li W, Wang Q, Zhu G, Fang W, Chen C, Wang X, et al. 2024. Genome variation and LTR-RT analyses of an ancient peach landrace reveal mechanism of blood-flesh fruit color formation and fruit maturity date advancement. Horticulture Research 11, uhad265.

Wu Y, Liu J, Sheng X, Wang W, Wang T, Martinez-Sanchez M, Wang S, Tu M, Deng J, Allan AC, et al. 2025. Spatial regulation of chlorophyll degradation in kiwifruit: AcNAC2–AcSGR1/2 cascades mediate rapid de-greening in the inner pericarp. Plant Biotechnology Journal 23, 2554–2569.

Zhang RX, Liu Y, Zhang X, Chen X, Sun J, Zhao Y, Zhang J, Yao JL, Liao L, Zhou H, et al. 2023. Two adjacent NAC transcription factors regulate fruit maturity date and flavor in peach. *New Phytologist* 241, 632–649.

Zhang LL, Wang XF, Dong K, Tan B, Zheng XB, Ye X, Wang W, Cheng J, Feng J. 2024. Tandem transcription factors PpNAC1 and PpNAC5 synergistically activate the transcription of PpPGF to regulate peach softening during fruit ripening. *Plant Molecular Biology* 114, 46.

Zhang B, Wang X, Yue Q, Zhang W, Liu H, Zhang T, Zhao L, Guan Q, You C, An J, et al. 2025. Autosuppression of MdNAC18.1 endowed by a 61-bp promoter fragment duplication delays maturity date in apple. Plant Biotechnology Journal 23, 1216–1229.

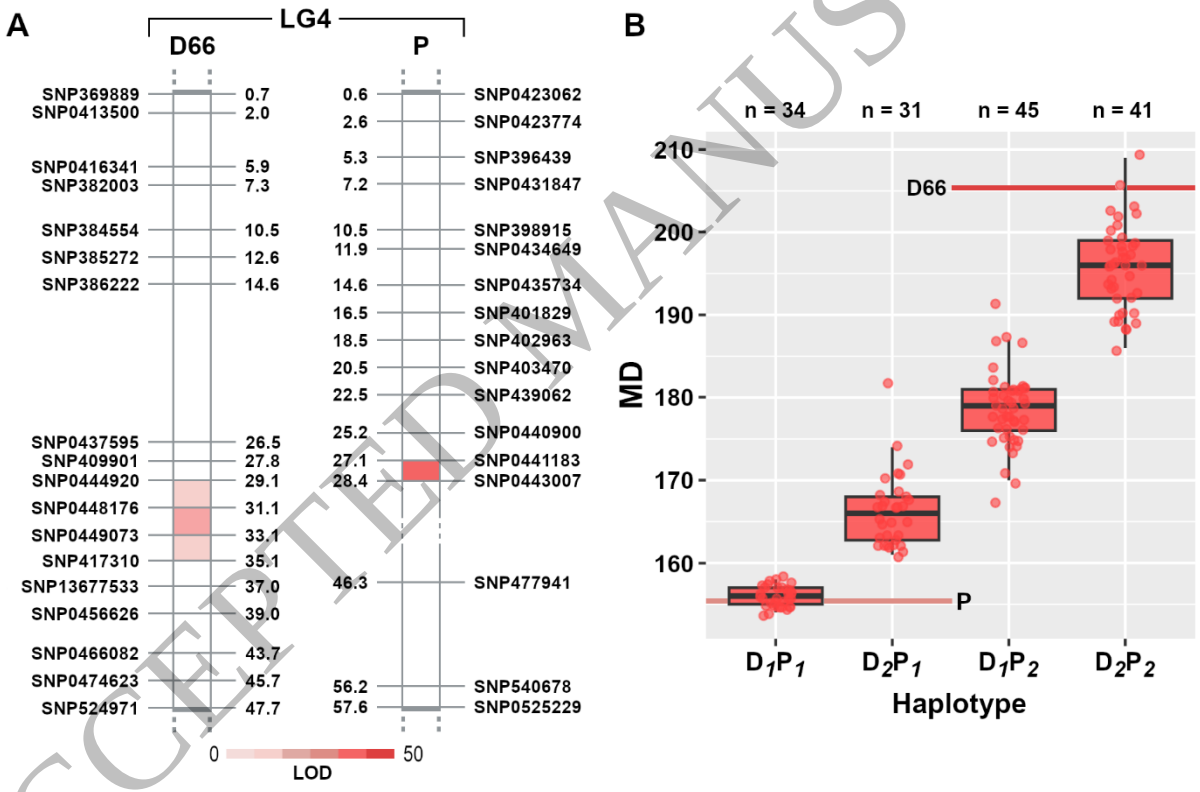
Zhou H, Zhang W, Sheng Y, Qiu K, Liao L, Shi P, Xie Q, Pan H, Zhang J, Han Y. 2023. A large-scale behavior of allelic dropout and imbalance caused by DNA methylation changes in an early-ripening bud sport of peach. *New Phytologist* 239, 13–18.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

## Figure legends

**Figure 1. QTL-mapping for maturity date (MD) in the $F_1$ DxP population**. (A) Linkage maps of LG4 for 'Dulcebo66' (D66) and 'Pulchra' (P) parents, showing SNP positions and QTL peaks detected by interval mapping: in P the strongest signal corresponds to SNP_AO_0443007 (LOD = 38.9, 10.15 Mb), while in D66 the peak maps at SNP_AO_0444920 (LOD = 6.5, 11.9 Mb). Recombinant-defined candidate intervals are highlighted in red. (B) Effects of the four parental haplotype classes on MD, illustrating additive contributions from both parents; the *D1P1* haplotype is associated with earlier maturity. Boxplots display MD values recorded over two years, and differences among haplotypes are all significant according to ANOVA ($p < 0.01$).

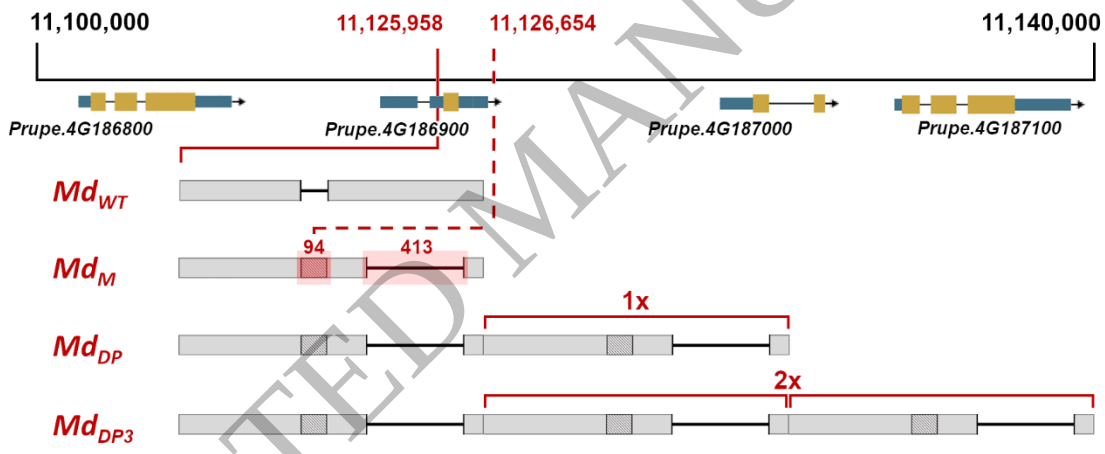



**Alt text**. Linkage maps on chromosome 4 identify a major QTL for maturity date in two parental lines, with highlighted candidate regions. Boxplots show four haplotype classes with additive effects, where specific allele combinations are associated with earlier ripening.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

**Figure 2**. **Structural variation at the *Md* locus**. Four allelic variants were identified in the region between *Prupe.4G186800* (*NAC5*) and *Prupe.4G187100* (*NAC1*) and near the 3'UTR of *Prupe.4G186900*, a predicted long non-coding RNA. The reference-like allele ($Md_{WT}$) corresponds to the structure present in the 'Lovell' reference genome. The $Md_{M}$ variant contains a composite rearrangement consisting of a 94 bp upstream tandem duplication insertion and a 413 bp downstream deletion (highlighted by the pink boxes) starting at ~11,126,654 bp (red dashed line). The early-ripening alleles $Md_{DP}$ and $Md_{DP3}$ consist of a full duplication or triplication of the $Md_{M}$ structural unit, respectively, beginning upstream of $Md_{M}$ rearrangement, at 11,125,958 bp, resulting in two or three consecutives ~1.5-kb copies (red continuous line). Grey boxes represent homologous blocks, hatched areas indicate the duplicated segment, and black bars mark deletion boundaries.

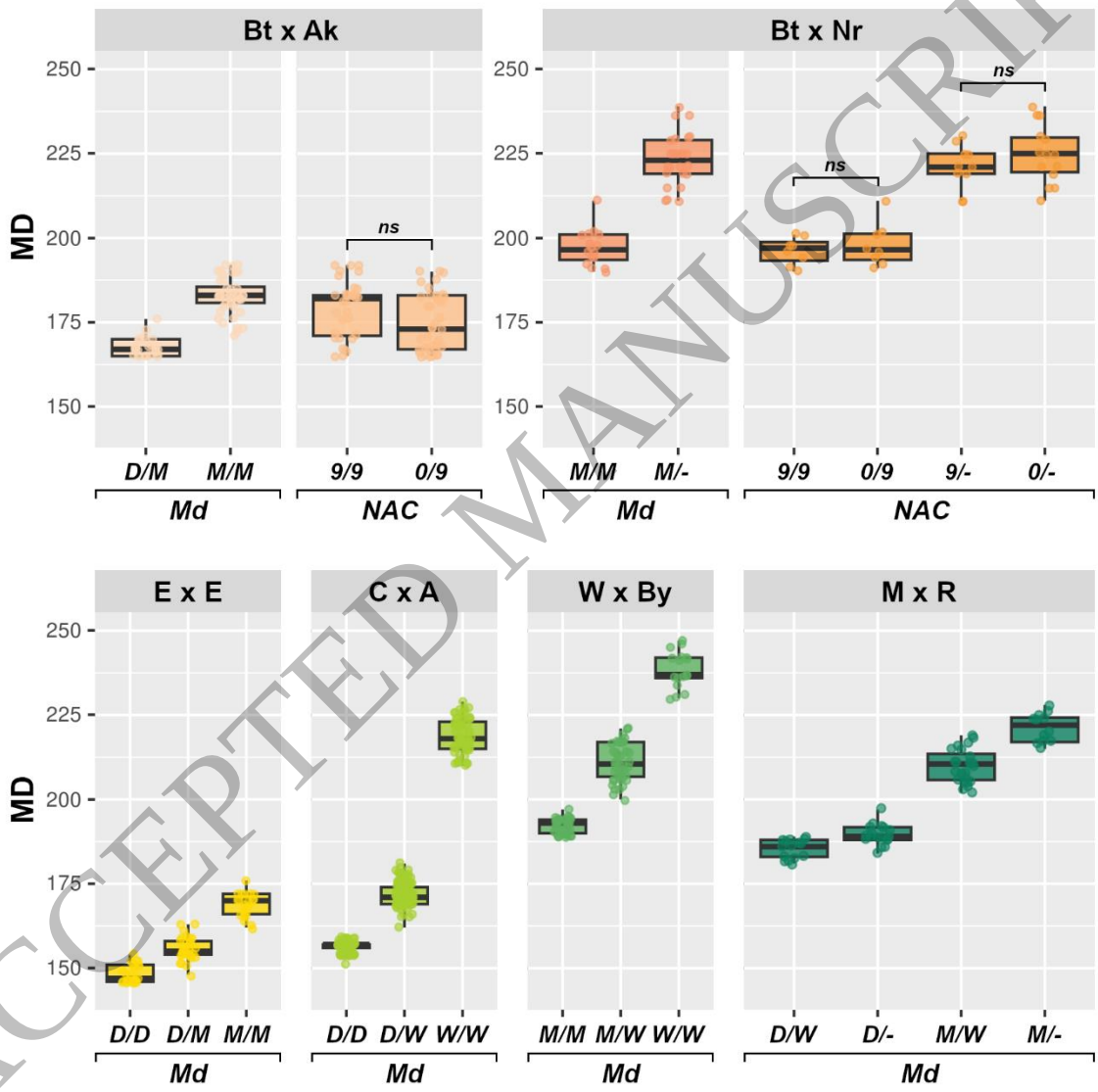



**Alt text**. Diagram showing structural variants in the region between NAC5 and NAC1, including the reference allele $Md_{WT}$ and derived alleles ($Md_{M}$, $Md_{DP}$, $Md_{DP3}$) with insertions, deletions, and tandem duplications that vary in the number of repeated sequence units.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

**Figure 3. Association of *Md* alleles with maturity date (MD) across multiple segregating progenies.** MD distributions are shown for *Md* and *NAC* genotypes in Bt × Ak, Bt × Nr, E × E, C × A, W × By and M × R. Md alleles are coded as 'D' ($Md_{DP}$), 'M' ($Md_{M}$), 'W' ($Md_{WT}$) and '-' (*sr,* slow ripening allele). Means, standard deviations, and significance scores for each progeny are summarized in Table 2.

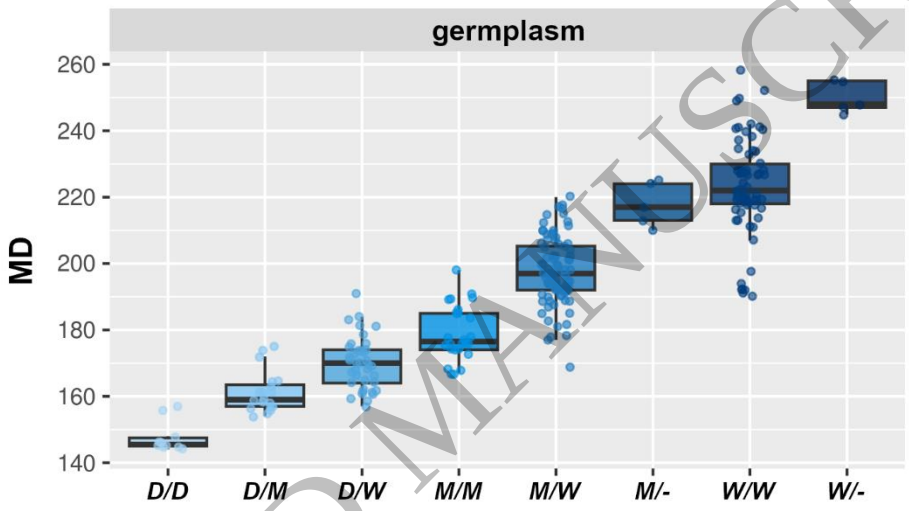



**Alt text**. Boxplots across multiple genetic populations show maturity date distributions for different *Md* genotypes. Consistent differences among allele classes demonstrate that *Md* variants strongly influence ripening time across diverse crosses.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

**Figure 4. Boxplot of the association between *Md* alleles and maturity date (MD) in peach germplasm.** *Md* genotypes stratify MD into distinct and additive classes across the germplasm panel, with all pairwise contrasts showing statistically significant differences (ANOVA, $p < 0.01$). The observed phenotypic separation spans the full harvest window (~140 - 250 JD). *Md* alleles are coded as 'D' ($Md_{DP}$), 'M' ($Md_{M}$), 'W' ($Md_{WT}$) and '-' (*sr,* slow ripening allele). Sample sizes per genotype class are as follows: M/W (n = 81), W/W (n = 73), D/W (n = 38), M/M (n = 30), D/M (n = 21), D/D (n = 10), W/- (n = 5) and M/- (n = 5).



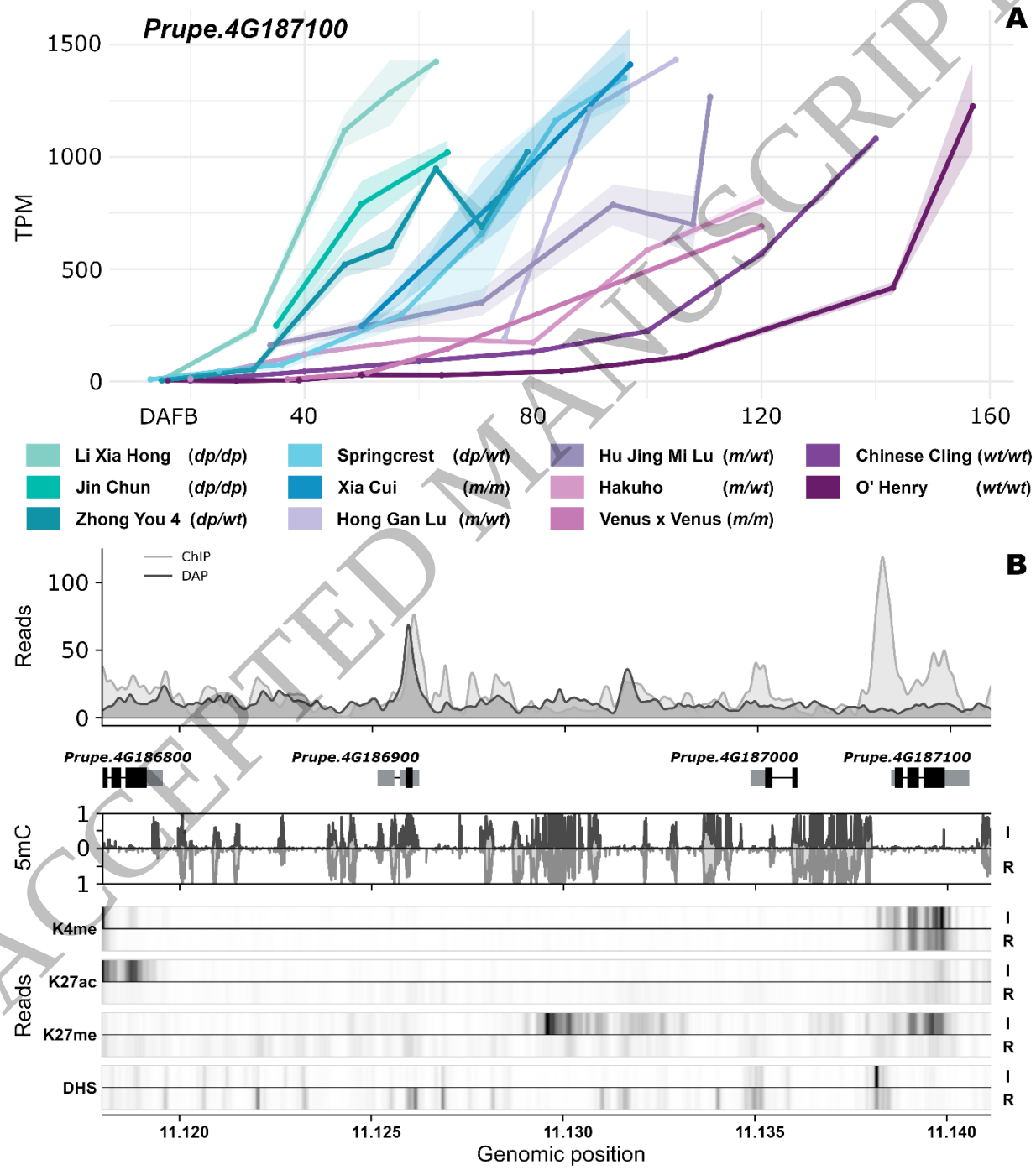

**Alt text.** Boxplot of maturity date in peach germplasm showing clear separation into distinct classes according to *Md* genotype. Alleles have additive effects and collectively span the full range of ripening times from early to late.

**Figure 5. Functional evidence linking *Md* structural variants to *NAC1* expression dynamics and chromatin state.** (A) Expression dynamics of *NAC1* (*Prupe.4G187100*) during fruit development in peach genotypes carrying distinct *Md* alleles. RNA-seq profiles are shown across developmental time, from early fruit development to ripening (days after full bloom, DAFB), as reported in the original studies. Curves represent the mean expression trend, with shaded areas indicating variability across available datasets. (B) Epigenomic profiles across the upstream regulatory region of NAC1. ChIP-seq and DAP-seq profiles highlight two regions of NAC1 binding: one proximal to the coding sequence and a second within the upstream regulatory region (~11.126 Mb), coinciding with the insertion site of *Md* structural variants. DNA methylation (5mC) from fruit mesocarp, together with chromatin accessibility (DNase I hypersensitivity) and histone modification profiles (H3K4me3, H3K27ac and H3K27me3), are shown for immature (I) and

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

mature (R) fruit stages and indicate a chromatin environment permissive to transcription at this locus. Because datasets differ in sampling schemes, developmental staging and growth conditions, both transcriptomic and epigenomic data are interpreted in terms of relative patterns and developmental timing.

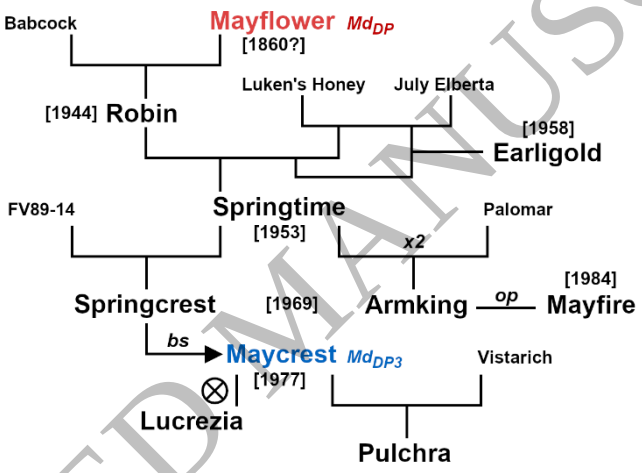



ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

**Alt text**. Line plots show NAC1 expression over fruit development, with earlier activation observed in genotypes carrying early-ripening alleles. Genomic tracks below display chromatin features and NAC1 binding in the upstream region.

**Figure 6. Pedigree of early-ripening peach cultivars tracing the origin of *Md$_{DP}$* and *Md$_{DP3}$* alleles.** 'Mayflower' introduces the early-ripening *Md$_{DP}$* haplotype into modern germplasm through 'Robin' and 'Springtime'. 'Springcrest' later gives rise to the bud sport 'Maycrest', in which an additional somatic duplication generates the *Md$_{DP3}$* allele, transmitted to 'Lucrezia' and 'Pulchra'. Years of release are shown in brackets; bs = bud sport; op = open pollination; ×2 indicates repeated use of 'Springtime' as a parent.




**Alt text**. Pedigree diagram tracing the origin and inheritance of early-ripening alleles in peach cultivars. The *Md$_{DP}$* allele originates from 'Mayflower' and spreads through breeding lines, while a further mutation event generates *Md$_{DP3}$* in a derived cultivar.

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

## Tables

**Table 1**. Summary of *Md* allelic composition across peach germplasm, breeding founder and mapping parents. Maturity date (MD) and fruit developmental period (FDP) are expressed in Julian days (JD) and days from blooming to maturity (D), respectively; *sr* indicates the slow-ripening allele.

| accession | origin | *Md* | MD (JD) | FDP (D) |
|---|---|---|---|---|
| | **Germplasm** | | | |
| Lucrezia | Maycrest self | $Md_{DP3}/Md_{DP3}$ | 145 | 72 |
| Spring Baby | Springcrest op x Springcrest op | $Md_{DP}/Md_{DP}$ | 153 | 80 |
| Mayfire | Armking op | $Md_{DP}/Md_{DP}$ | 157 | 84 |
| Springtime | (Lukens's Honey x July Elberta) x Robin | $Md_{DP}/Md_{M}$ | 160 | 87 |
| Maycrest | Springcrest mutation | $Md_{DP3}/Md_{WT}$ | 165 | 92 |
| Springcrest | FV89-14 x Springtime | $Md_{DP}/Md_{WT}$ | 170 | 97 |
| Large White | Germplasm | $Md_{M}/Md_{M}$ | 178 | 110 |
| Contender | Winblo x (Norman x NCA2679) | $Md_{M}/Md_{WT}$ | 215 | 140 |
| Elegant Lady | Early O'Henry x July Lady | $Md_{WT}/Md_{WT}$ | 218 | 143 |
| O'Henry | Merril Bonanza op | $Md_{WT}/Md_{WT}$ | 235 | 160 |
| Zhong You 4 | from Mayfire | $Md_{DP}/Md_{WT}$ | - | 79 |
| Li Xia Hong | Zhong You 4 mutation | $Md_{DP}/Md_{DP}$ | - | 63 |
| Redhaven | Halehaven x Kalhaven | $Md_{M}/Md_{WT}$ | 195 | 125 |
| | **cross-parents** | | | |
| F1 CxA | Contender x Ambra | $Md_{DP}/Md_{WT}$ | 180 | 110 |
| Rebus028 | Big Top x Mayfire | $Md_{DP}/Md_{M}$ | 170 | 97 |
| Max10 | unknown pedigree (Minguzzi) | $Md_{WT}/sr$ | 255 | 180 |
| BigTop | unknown pedigree (Zaiger) | $Md_{M}/Md_{M}$ | 190 | 120 |
| Armking | (Palomar x Springtime) x (Palomar x Springtime) | $Md_{DP}/Md_{M}$ | 174 | 105 |
| Ambra | Stark Red Gold x Mayfire | $Md_{DP}/Md_{WT}$ | 180 | 110 |
| Dulcebo066 | sel. x (Contender x Elegant Lady) | $Md_{M}/Md_{WT}$ | 210 | 137 |
| Pulchra | Vistarich x Maycrest | $Md_{DP3}/Md_{M}$ | 162 | 89 |
| Nectaross | Stark Red Gold x Legrand | $Md_{M}/sr$ | 221 | 148 |
| Earligold | complex from Robin | $Md_{DP}/Md_{M}$ | 165 | 92 |
| Bounty | [(Halberta op) x Redskin] x (Loring x FV89-14) op | $Md_{M}/Md_{WT}$ | 203 | 130 |
| NJ Weeping | ornamental germplasm (PI) | $Md_{WT}/Md_{WT}$ | 235 | 160 |

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026

**Table 2.** Summary of *Md* and *NAC* genotypic classes across multiple populations, their association with maturity date (MD) and significance levels. The table reports haplotype relationships between loci, highlighting cases of perfect and partial linkage, as well as non-informative progenies. Abbreviations: MD, maturity date; n, sample size; sr, slow-ripening allele; $NAC_0$ (0) and $NAC_9$ (9), for *NAC* allele; $Md_{WT}$ (WT), $Md_M$ (M), $Md_{DP}$ (DP) and $Md_{DP3}$ (DP3) for *Md* alleles; *ns*, not significant.

| Pop | *Md-NAC* segregation | genotype | # | MD | *p*-value | *Md-NAC* linkage | Note |
|---|---|---|---|---|---|---|---|
| D×P | *M/WT × DP3/M* | $Md_{DP3}/Md_M$ | 33 | 155.7±1.1 | $p < 0.001$ | D: $Md_M \leftrightarrow NAC_9$; $Md_{WT} \leftrightarrow NAC_0$ | partial linkage |
| | | $Md_{DP3}/Md_{WT}$ | 32 | 166.1±4.5 | | | |
| | | $Md_M/Md_M$ | 45 | 178.3±4.5 | | | |
| | | $Md_M/Md_{WT}$ | 41 | 195.5±5.1 | | P: $Md_{DP3} \leftrightarrow NAC_9$; $Md_{WT} \leftrightarrow NAC_0$ | |
| | *0/9 × 9/9* | $NAC_0/NAC_9$ | 73 | 182.6±15.4 | $p < 0.001$ | | |
| | | $NAC_9/NAC_9$ | 78 | 168.7±11.7 | | | |
| Bt×Ak | *M/M × DP/M* | $Md_{DP}/Md_M$ | 28 | 167.4±2.6 | $p < 0.001$ | Bt: $Md_M \leftrightarrow NAC_9$; $Md_M \leftrightarrow NAC_0$ | segregation of Bt alternative $Md_M$-$NAC_0$ haplotype |
| | | $Md_M/Md_M$ | 44 | 182.6±5.5 | | | |
| | *9/0 × 9/9* | $NAC_0/NAC_9$ | 33 | 178.9±8.5 | $p > 0.039$ *ns* | Ak: $Md_{DP} \leftrightarrow NAC_9$; $Md_M \leftrightarrow NAC_9$ | |
| | | $NAC_9/NAC_9$ | 39 | 174.7±8.6 | | | |
| Bt×Nr | *M/M × M/sr* | $Md_M/Md_M$ | 16 | 197.0±5.3 | $p < 0.001$ | Bt: $Md_M \leftrightarrow NAC_9$; $Md_M \leftrightarrow NAC_0$ | segregation of Bt alternative $Md_M$-$NAC_0$ haplotype Nr segregate for *sr* allele |
| | | $Md_M/sr$ | 24 | 223.2±7.8 | | | |
| | *9/0 × 9/sr* | $NAC_0/sr$ | 14 | 224.7±8.5 | $p > 0.136$ *ns* | | |
| | | $NAC_9/sr$ | 11 | 220.1±6.9 | | | |
| | | $NAC_9/NAC_0$ | 7 | 196.1±4.2 | $p > 0.993$ *ns* | Nr: $Md_M \leftrightarrow NAC_9$ | |
| | | $NAC_9/NAC_9$ | 8 | 195.1±4.1 | | | |
| $F_2$ E×E | *DP/M × DP/M 9/9 × 9/9* | $Md_{DP}/Md_{DP}$ | 25 | 148.3±2.3 | $p < 0.001$ | $Md_{DP} \leftrightarrow NAC_9$; $Md_M \leftrightarrow NAC_9$ | no NAC segregation |
| | | $Md_{DP}/Md_M$ | 26 | 155.6±3.1 | | | |
| | | $Md_M/Md_M$ | 21 | 168.8±3.2 | | | |
| $F_2$ W×By | *M/WT × M/WT 9/0 × 9/0* | $Md_M/Md_M$ | 41 | 191.8±2.1 | $p < 0.001$ | $Md_M \leftrightarrow NAC_9$; $Md_{WT} \leftrightarrow NAC_0$ | perfect linkage |
| | | $Md_M/Md_{WT}$ | 44 | 210.8±5.4 | | | |
| | | $Md_{WT}/Md_{WT}$ | 17 | 238.2±5.5 | | | |
| $F_2$ C×A | *DP/WT × DP/WT 9/0 × 9/0* | $Md_{DP}/Md_{DP}$ | 94 | 156.2±1.4 | $p < 0.001$ | $Md_{DP} \leftrightarrow NAC_9$; $Md_{WT} \leftrightarrow NAC_0$ | perfect linkage |
| | | $Md_{DP}/Md_{WT}$ | 143 | 171.4±3.2 | | | |
| | | $Md_{WT}/Md_{WT}$ | 65 | 218.5±4.8 | | | |
| M×R | *WT/sr × DP/M 0/sr × 9/0* | $Md_{DP}/sr$ | 18 | 189.4±2.7 | $p < 0.001$ | M: $Md_{WT} \leftrightarrow NAC_0$ | partial linkage M segregate for *sr* R $Md_M$-$NAC_0$ haplotype (not informative) |
| | | $Md_M/sr$ | 12 | 221.4±4.3 | | | |
| | | $Md_M/Md_{WT}$ | 24 | 209.8±4.9 | | R: $Md_{DP} \leftrightarrow NAC_9$; $Md_M \leftrightarrow NAC_0$ | |
| | | $Md_{DP}/Md_{WT}$ | 15 | 185.3±2.7 | | | |

ACCEPTED MANUSCRIPT

Downloaded from https://academic.oup.com/jxb/advance-article/doi/10.1093/jxb/erag210/8664534 by Universita degli Studi di Milano user on 30 April 2026