Running title: Melon flavonoid accumulation is regulated by CmKFB

* Corresponding author details:
Yaakov Tadmor
Cucurbit Section, Newe Ya’ar Research Center, ARO
Ramat Yishay 30095, ISRAEL
Fax: 972-4-9836936
Phone: 972-4-9539548
E-mail: tadmory@agri.gov.il

Research area: metabolism
A Kelch domain-containing F-box coding gene negatively regulates flavonoid accumulation in *Cucumis melo*


Cucurbit Section, Newe Ya’ar Research Center, ARO, Ramat Yishay 3009500, Israel (A. F., J. B., A. M., R. D., V. P., E. L., N. K., Y. T.); Faculty of Biotechnology and Food Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel (A. F., Y. K.); Plant Science Institute, ARO, The Volcani Center, P.O.B. 6, Bet Dagan, 50250, Israel (A. A. S.); Plant Protection Institute, ARO, The Volcani Center, P.O.B. 6, Bet Dagan, 50250, Israel (A. G.); Boyce Thompson Institute for Plant Research, Cornell University, Ithaca, NY 14853, USA (S.G., Z. F.)

One Sentence Summary:

Melon *CmFBK* expression regulates a shift of the flavonoid metabolic flux toward coumarin and general phenylpropanoids.

* - Address correspondence to tadmory@agri.gov.il
ABSTRACT

The flavonoids are phenylpropanoid-derived metabolites that are ubiquitous in plants, playing many roles in growth and development. Recently, we observed that fruit rinds of yellow casaba melons (Cucumis melo Inodorous Group) accumulate naringenin chalcone, a yellow flavonoid pigment. RNA-Seq analysis of bulksampled segregants, representing the tails of a population segregating for naringenin chalcone accumulation, followed by fine mapping and genetic transformation, we identified a Kelch domain-containing F-box protein coding gene (CmKFB) that, when expressed, negatively regulates naringenin chalcone accumulation. Further metabolite analysis indicated that downstream flavonoids are accumulated together with naringenin chalcone, while CmKFB expression diverts the biochemical flux towards coumarins and general phenylpropanoids. These results demonstrate that CmKFB functions as a post-transcriptional regulator which diverts flavonoid metabolic flux.
INTRODUCTION

Melons, *Cucumis melo* L. (Cucurbitaceae), exhibit extreme diversity in fruit traits, varying in size, shape, external color, aroma, sugar content, acidity, pigmentation and climacteric or non-climacteric fruit ripening (Burger et al., 2010; Cohen et al., 2014). Various classifications of melons have been proposed over the years. The most comprehensive yet easily understood classification distinguishes two subspecies, *agrestis* and *melo*, with 5 and 11 groups respectively (Pitrat et al., 2000).

The phenylpropanoid pathway gives rise to a large range of phenolic compounds, including flavonoids, lignins, lignans, coumarins, stilbenes and sinapate esters (Veitch and Grayer, 2008; Petersen et al., 2010), which are involved in many aspects of plant physiology. Phenylpropanoids have roles related to photoprotection (Kim et al., 2008), crosstalk with hormones (Besseau et al., 2007; Yuan et al., 2013), *Rhizobium* symbiosis (Wasson et al., 2006), pollination (Mo et al., 1992), structural components (Pizzi and Cameron, 1986), and plant-animal interactions (Haribal and Renwick, 1996). Recently flavonoids have gained increasing interest as health benefit agents in the human diet (Butelli et al., 2008; Romano et al., 2013).

Due to the wide interest in flavonoids, considerable efforts have focused on understanding the regulatory mechanisms of the different branches of their synthesis pathway. Our general understanding is that flavonoid accumulation is largely controlled at the transcriptional level, by transcription factors, regulating the expression of the pathway structural genes (Hartmann et al., 2005; Koes et al., 2005). However, pathway regulation downstream to the transcriptional level is less understood (Tanaka and Uritani, 1977; Pairoba and Walbot, 2003; Deguchi et al., 2013). Lately, Kelch domain containing F-box proteins (KFB) were shown to be involved in regulation of the phenylpropanoid pathway (Shao et al., 2012; Zhang et al., 2013; Zhang et al., 2015). F-box proteins serve as substrate recruitment domains in the SCF (Skp1, Cullin, F-box) E3 ubiquitin ligase complexes, targeting protein substrates to 26S proteasome dependent degradation, in which the F-box serves as the target recognition factor, mediating the specificity of the complex (Jonkers and Rep, 2009). In plants, the F-box
superfamily seems to have gone through vast duplication and specialization. Although the human and *Drosophila* genomes contain only 68 and 33 *F-box* genes, respectively (Ou et al., 2003; Jin et al., 2004), the *Arabidopsis* genome contains about 700 members of this superfamily (Gagne et al., 2002). One subgroup of this family, the KFB, found as a single ortholog in the human genome, is highly represented in the *Arabidopsis* genome, comprising at least 103 KFB sequences, most of them still uncharacterized (Schumann et al., 2011). A rice KFB homolog was found to negatively regulate flavonoid accumulation (Shao et al., 2012). Four members of the KFB subfamily in *Arabidopsis* were shown to physically interact with different phenylalanine ammonia-lyases (PAL) isozymes, mediating their degradation, and subsequently decreasing accumulation of an array of various phenylpropanoids, implicating these KFBs as general repressors of the phenylpropanoid pathway (Zhang et al., 2013; Zhang et al., 2015).

Little is known about the regulation of the flavonoid pathway in the Cucurbitaceae. Flavone derivatives are accumulated in leaves of *Cucumis sativus* (cucumber) and *C. melo* (Krauze-Baranowska and Cisowski, 2001). Kaempferol and quercetin glycosides (flavonol derivatives) accumulate in the reproductive organs of some cucurbits (Imperato, 1980). Recently, we showed that naringenin chalcone (NarCh) is accumulated as the main flavonoid in the fruit rind of some melon varieties, including the yellow casaba melon, ‘Noy Amid’ (NA; Tadmor et al., 2010).

In this study, we used a combination of RNA-Seq of pools of F₃ segregants together with fine mapping, to identify *CmKFB*, a protein coding gene which negatively regulates flavonoid accumulation in melon. The repression of the flavonoid pathway results in biochemical flux alternation towards the accumulation of coumarin and other general phenylpropanoids derivatives. Functional verification of *CmKFB* was performed by its transient expression in melon leaves and stable transformation of tomato resulting in altered phenylpropanoid patterns. Our results provide experimental evidence of regulation that alters the flow between the phenylpropanoid pathway branches.

**RESULTS**

Characterization of biochemical differences between two melon genotypes differentiating in
NarCh accumulation

To obtain a general view of the morphological and biochemical differences between the NarCh accumulating (NA) and non-accumulating (TVT) melons, microscopic examinations were performed. Slices of melon rind were stained with DPBA (diphenyl boric acid aminoethyl ester), which fluorescently dyes flavonoid glycosides (Stracke et al., 2007). Under visible light the upper cuticle layer of NA was yellow, probably due to the accumulation of NarCh, while in TVT this layer seems transparent. After DPBA staining and under UV illumination, NA samples exhibited a greenish fluorescence in the tissue that is located at approximately 2 mm under the upper cuticle layer, while the cuticle of TVT exhibited blue fluorescence and lacked the greenish fluorescence under the cuticle (Fig. 1, A-H). This differential fluorescence display suggests that both NA and TVT accumulate different flavonoid compounds in addition to NarCh. The cuticle of NA exhibited no fluorescence after DPBA staining, suggesting the NarCh is accumulated in the cuticle as aglycone and not a glycoside that would be fluorescent under UV light. To further investigate whether the additional fluorescent compounds accumulated in different tissues in NA and TVT are associated with NarCh accumulation, methanol extracts of mature fruit rinds of 19 F3 families were fractionated by thin layer chromatography (TLC), stained with DPBA and visualized under UV light. The 19 families were generated from a cross between NA and TVT (12 ‘yellow’ and 7 ‘white’ F3 families in “Materials and Methods”; Supplemental Fig. S1). Under these conditions, the fluorescence pattern of the ‘yellow’ F3 families (NarCh accumulating) was identical with that of the NA parent, while the ‘white’ F3 families exhibited the same fluorescence pattern as TVT, the non-NarCh accumulating parent, showing association of these fluorescent compounds with NarCh accumulation after two generations (Fig. 1I).

DPBA mediated fluorescence of additional tissues of the two melon types indicated that leaves and flowers, regardless the presence/absence of NarCh in their fruits, exhibited similar fluorescence patterns (Supplemental Fig. S2 A and B). The differences in fruit rind fluorescence between the two melon types were evident from 10 days after anthesis (DAA). During fruit development, the fluorescence intensity of yellow/green compounds decreased in ‘white’ and increased in ‘yellow’ F3s while blue fluorescent compounds gradually increased during ‘white’ fruit development and were absent in ‘yellow’ fruit (Supplemental Fig. S2 C, D and E).
Similar staining experiments performed in *Arabidopsis* identified yellow and green DPBA-mediated fluorescent compounds as quercetin and kaempferol (flavonols), and blue fluorescence emitting compounds as sinapate glycoside derivatives (Stracke et al., 2007; Stracke et al., 2009). Application of additional staining including Shinoda’s test and Zn/HCl (Pew’s test) on ‘yellow’ and ‘white’ F₃ methanol extracts (Supplemental Fig. S3) showed color accumulation in ‘yellow’ was correlated with the accumulation of flavonols and dihydroflavonols (Pew, 1948), respectively, in contrast to ‘white’ which was negative to these stainings, suggesting that additional flavonoids were accumulated in addition to NarCh.

To further identify these fluorescent compounds, methanolic extracts of both parents (NA and TVT) and bulks of F₃ families (‘white’ and ‘yellow’) were analyzed with LC-MS (Supplemental Fig. S4, A-D). A similar pattern to the previously presented fluorescence on TLC was observed: the NarCh accumulating parent and the ‘yellow’ F₃ bulk possessed similar characteristics, which are distinctive from the non-NarCh accumulating TVT and the ‘white’ F₃ group. An untargeted metabolic analysis was conducted to identify the main compounds that were differently accumulated in these two populations using Agilent's MPP metabolic tools. The means of the total ion counts of the different compounds found in ‘yellow’ and ‘white’ groups (each group is comprised of a parental line and a bulk of F₃ families) were compared by a volcano plot using a threshold of at least 10 fold difference between the means and a P value of 0.05. A total of 99 masses that significantly differentiated between these two groups were identified (Supplemental Fig. S4E). The software could predict molecular formula for 64 compounds of these 99 masses, many of them matching different flavonoid glycosides (Supplemental Table S2). To further identify the aglycone backbones which underlined these differences, enzymatic hydrolysis was performed on the samples, after which aglycones were isolated and analyzed with LC-MS. A corresponding similar pattern of statistically significant differential accumulation of compounds between the two groups was observed (Fig. 2; Supplemental Table S3). Identification of the aglycones was performed via comparison with authentic standards. Eight compounds were identified, among which caffeic acid, ferulic acid, and aesculetin showed significantly higher accumulation in the non-NarCh accumulating fruit rinds (TVT and ‘white’ F₃ families), while, naringenin, luteolin, kaempferol, quercetin, and traces...
of dihydrokaempferol were found to specifically accumulate in the NarCh accumulating tissues (NA and 'yellow' F₃ families; Supplemental Table S4; Supplemental Fig. S5). Thus, a clear change of the pathway flux could be noted, as evidenced by the metabolic patterns observed between the two groups (Fig. 3).

NarCh-accumulating lines accumulate additional downstream flavonoids including: luteolin (a flavone), dihydrokaempferol (a dihydroflavonol), quercetin and kaempferol (flavonols). The non-NarCh accumulating lines over-accumulate mainly caffeic and ferulic acids, in addition to aesculetin (a coumarin), which are branched prior to the flavonoid biosynthetic pathway. Although clear differences in accumulation of caffeic and ferulic acids was observed between the NarCh accumulating types, smaller levels of caffeic and ferulic acids were detectable in the NarCh accumulating lines (retention times 3.7, 6.3 minutes in Fig. 2, A and B; Supplemental table S4), whereas flavonoids were not detected in the non-NarCh accumulating types.

**Genetic mapping of the locus governing flavonoid accumulation**

Previous studies on fruit rind NarCh accumulation in an F₂ population derived from the cross NA × TVT, suggested that this trait is governed by a single gene (Tadmor et al., 2010). To better identify the genetic factor regulating flavonoid accumulation in melon fruit rind, we comparatively analyzed the transcriptome of the ‘yellow’ and ‘white’ bulks of F₃ families, a methodology that is based on the classical ‘bulk segregant analysis’ approach (Michelmore et al., 1991; Supplemental Fig. S1; “Materials and Methods”). RNA extracted from developing fruits of ‘yellow’ and ‘white’ F₃ families was bulked and sequenced (Supplemental Table S5). As at the time we carried out this portion of the research the melon genome sequence was not yet available, a *de-novo* assembled transcriptome was developed that included 87,336 unigenes. All RNA-Seq reads of the two bulks at three developmental stages were aligned to the unigenes for the identification of single nucleotide polymorphism (SNP) that existed between the ‘yellow’ and ‘white’ bulks. Upon the availability of the melon genome (Garcia-Mas et al., 2012) SNPs were mapped to the melon genome. One hundred and seventy nine SNPs were found to differentiate between the two bulks (Supplemental Table S6), 164 of them are located on
scaffold 16 (Linkage group 10) of the melon genome, and the majority of them are localized in a region of about 0.5Mb (Fig. 4A), suggesting that the gene controlling NarCh accumulation in melon fruit rind is located on scaffold 16 of chromosome 10. The remaining 15 SNPs were scattered randomly throughout the genome suggesting that they are not specifically related to flavonoid accumulation.

For a more detailed characterization of the genetic locus that governs NarCh accumulation, analysis was performed on additional 784 plants, derived from 14 F3 families segregating for fruit rind NarCh accumulation. Mature fruit of this additional plant material exhibited three phenotypes in the field: yellow rind, pale yellow rind and white rind (Fig. 4B). Phenotypic distribution of these F3 individuals supported a 1:2:1 ratio (yellow : pale yellow : white), expected for a single co-dominant gene (chi-square = 2.593, P = 0.27; Fig. 4C). Yellow fruit were assumed to be homozygous for the dominant allele of the gene governing NarCh accumulation in melon fruit rind, white fruit were assumed to be homozygous recessive and pale yellow, heterozygotes. DNA extracted from these F3 individuals was used to map the locus governing NarCh accumulation by the analysis of recombinants.

At first, three SNPs, UN22458, 06224 and 10338, were selected for genotyping of the segregating population (SNP locations are marked with arrows in Fig. 4A; Supplemental Tables S6 and S8) to screen the segregating population. UN10338 showed the highest genotype to phenotype association (Fig. 4E). Fine mapping using recombinants, which were identified by these three markers, was carried out using nine selected SNPs, dispersed in the genomic region restricted by UN10338 and UN06224 (Supplemental Fig. S6; Supplemental Table S8). A number of informative recombinants were identified, including two plants with white fruit (no. 1030 and 826), three pale yellow fruit plants (no. 539, 1041 and 1281), and two plants with yellow fruits (no. 1442 and 720; Fig. 4G). Combined phenotypic and genotypic data of the plants indicated that the genetic factor regulating flavonoid accumulation in the melon segregating population is located between high resolution melt (HRM) markers no. 2 and 4 (Fig. 4G). All the recombinant F3 plants that were genotyped showed a complete co-segregation between their genotype at this chromosomal region and their fruit color phenotype. Seventeen protein coding genes, from \textit{MELO3C01178} to \textit{MELO3C01194} (Supplemental Table S7), are located within this region.
approximately 100 Kb chromosomal region, from position 2,156,094 to position 2,255,982 on scaffold 16 of chromosome 10.

Characterization of a candidate gene regulating flavonoid accumulation

MELO3C011980, which is annotated as being similar to F-box/kelch-repeat protein At1g23390 (Arabidopsis thaliana) is one of the 17 genes that are located in the region restricted by HRM markers no. 2 and no. 4. At1g23390 KFB protein coding gene hasn’t been functionally characterized (Supplemental Fig. S7). MELO3C011980 was the most significantly differentially expressed gene (DEG) between the two bulks; while thousands of raw reads were generated in ‘white’ bulks during all three analyzed developmental stages, only a few reads were detected in ‘yellow’ bulks (Supplemental Fig. S8A). We named MELO3C011980 CmKFB. The RNA-Seq data could not detect any sequence polymorphism between CmKFB ‘yellow’ and ‘white’ alleles due to the extremely low number of reads derived from the ‘yellow’ allele. To identify DNA sequence polymorphism that might affect CmKFB differential expression, sequencing of the upstream portion of the gene was performed. A deletion of 12-bp in the 5’ UTR of CmKFB was found in ‘TVT’ parental lines (Supplemental Fig. S8B). In order to indicate function of this deletion we screened a range of C. melo accessions differing in peel color. This deletion was present in eight analyzed melon accessions that do not accumulate NarCh and is absent in all analyzed melon accessions that accumulate NarCh (Supplemental Fig. S8C, Supplemental Table S1).

As indicated earlier, DPBA staining of leaves exhibited similar yellow UV fluorescence patterns in both ‘white’ and ‘yellow’ F3 plants, suggesting both types accumulate flavonoids in their leaves (Supplemental Fig. S2A). To further associate the role of CmFBK with flavonoid accumulation, we conducted qRT-PCR analysis of CmKFB in leaves. As expected, a very low expression of CmKFB was found in leaves irrelevant to fruit rind NarCh accumulation (Supplemental Fig. S8D).

Further analysis of the effect of CmKFB expression on the flavonoid pathway was performed on individual F3 plants derived from the self-pollination of CmKFB heterozygous F2 plants. The fruit rinds of five F3 individuals from each of the visual phenotypic groups, white, pale yellow and yellow, were analyzed for CmKFB expression by qRT-PCR. All white fruits
showed strong expression of *CmKFB* and all yellow fruits showed very weak expression, similar to its expression in the F3 bulks derived from the RNA-Seq analysis (Supplemental Fig. S8A). The pale yellow fruit showed intermediate expression of *CmKFB* gene (Fig. 5B), which is correlated with the intermediate color phenotype and intermediate fluorescence pattern of DPBA stained fruit rind methanol extracts fractionated on TLC and visualized under UV light (Fig. 5C).

PCR analysis of the *CmKFB* 5′ UTR deletion area was performed on genomic DNA of the analyzed F3 plants, verified that ‘white’ are homozygotes to the TVT allele, ‘yellow’ are homozygotes to the NA allele, while the pale yellow are heterozygotes (Fig. 5D). Interestingly, similar PCR analysis with cDNA suggests heterozygous to accumulate only the TVT non-flavonoid accumulating parent allele (Fig. 5E).

We used the RNA-seq data to analyze the expression level (RPKM) of all gene family members of the genes coding for major enzymes of the flavonoid pathway, including *PAL*, cinnamate 4-hydroxylase (*C4H*), and chalcone synthase (*CHS*). Only one major ortholog in each family was found to be significantly up regulated during development (*MELO3C025786, MELO3C019585*, and *MELO3C014767* respectively; Supplemental Fig. S9), in correlation with NarCh accumulation pattern in NA (Supplemental Fig. S1), and phenylpropanoid accumulation in TVT, as observed by the increased blue fluorescence (Supplemental Fig. S2), suggesting only a single ortholog of these gene families participated in the fruit rind pathway in both melon types. On the other hand, 4-Coumarate:CoA ligase (*4CL*) gene family members exhibited various expression patterns during fruit development, one of which matched the expression pattern of the previous three genes (*MELO3C024724*, Supplemental Fig. S9). No significant differences in transcript abundance were found between ‘white’ and ‘yellow’ that could explain the biochemical differences between these two bulks.

The AtKFBs homologs (KFB01, KFB20, KFB39, and KFB50) which were reported to regulate the phenylpropanoid pathway, have been shown to be involved in PAL degradation resulting in down regulation of the phenylpropanoid pathway (Zheng et al 2013; Zheng et al 2015). The biochemical evidence presented in this study suggests CmKFB to regulate a downstream point
to these Arabidopsis KFBs. In order to access this point melon KFB homologs (Supplemental Table S9) and the Arabidopsis KFBs (Zhang et al. 2015) were phylogenetically analyzed (Supplemental Figure S11). This analysis shows the four AtKFBs to cluster with the melon ME3C014678 and MELO3C009596 homologs, while CmKFB clusters with the Arabidopsis At1g23390 in a subtree which is not tightly linked to the previous one.

**Functional analysis of CmKFB**

To functionally validate the ability of CmKFB to divert the flavonoid pathway biochemical flux, two different transformation systems were applied: A virus induced gene expression (VIGE) of CmKFB in melon leaves and a stable transformation of tomato MP1 plants with CmKFB under the constitutive 35S promoter.

VIGE is an established experimental procedure to analyze the effect of gene over-expression in cucurbits (Harpaz-Saad et al., 2007). Our chemical analysis of melon leaves suggested that flavonoid accumulation in leaves is irrelevant to flavonoid accumulation in melon fruit rind (Supplemental Fig. S2A), which is associated with very low expression of CmKFB (Supplemental Fig. S8D), thus overexpression of CmKFB in leaves appeared like a suitable experimental system to validate CmKFB function. Fourteen days after VIGE infection, CmKFB expression was significantly increased as indicated by qRT-PCR analysis (Fig. 6C). Leaf methanol extracts of VIGE-CmKFB and the two controls (virus only and VIGE-GFP) were fractionated with TLC, stained with DPBA and observed under UV light. CmKFB overexpression exhibited enhanced blue fluorescence and reduction in yellow-green fluorescence accumulation in comparison to the controls (Fig. 6E). Further LC-MS analysis of leaf methanol extracts showed that VIGE-CmKFB accumulated higher amounts of compounds (retention time 2.2-4.5 minutes) while lower amounts of other compounds eluting from the column during the time range of 5.5 to 9.5 minutes when compared to the two controls (Fig. 6, F-H). To identify the main aglycones differentiating between the controls and VIGE-CmKFB, enzymatic hydrolysis was performed and the hydrolyzed samples were analyzed by LC-MS and compared to authentic standards. Three aglycones were identified to be up-regulated by VIGE-CmKFB: caffeic acid, salycilic acid, and
aesculetin and three down-regulated: luteolin, naringenin, and quercetin (Fig. 6, I-J). These differences are less significant than those observed in fruits most probably due smaller difference in \textit{CmKFB} expression, and incomplete infection of the virus, as observed in the AG-GFP control (Fig. 6A). VIGE-\textit{CmKFB} down-regulated compounds were flavonoids while the upregulated compounds belong to side chains of the phenylpropanoid and coumarin pathways, upstream of the flavonoid pathway, indicating \textit{CmKFB} expression in melon leaves mediated a metabolic flux alternation as was observed in melon fruits.

To functionally analyze \textit{CmKFB} in a heterologous system we used a stable transformation of tomato MP-1 cultivars. Tomato plants accumulate anthocyanins in vegetative tissues in response to stress and accumulate NarCh, kaempferol and quercetin in their fruit peels (Mintz-Oron et al., 2008). A binary vector harboring \textit{CmKFB} and a \textit{GFP} reporter gene was used for stable transformation. Two transgenic plants were isolated. These plants did not accumulate visible amounts of anthocyanins in any observed tissues: leaves, stems, flowers (Supplemental Fig. S10). At the ripe fruit stage, fruit peels of both transformed \textsubscript{T0} plants accumulated significantly reduced NarCh in the cuticle, compared to the MP-1 control (Fig. 7 A and B). In addition, fruit sections under UV light after DPBA staining exhibited a shift from yellow to blue fluorescence in the \textit{CmKFB} transgenic fruit (Fig. 7C), similar to VIGE-\textit{CmKFB} leaves (Fig. 6E), and to melon fruit rinds (Fig. 1I). \textsubscript{T1} seeds from both transgenic plants were germinated and grown under cold stress conditions. After 15 days, a clear difference could be observed between anthocyanin accumulating seedlings and seedlings which did not accumulate anthocyanins. The \textsubscript{T1} plantlets with anthocyanins did not exhibit GFP fluorescence while those that did not accumulate anthocyanins exhibited GFP fluorescence (Fig. 7D), indicating overexpression of \textit{CmKFB} in tomato to repress flavonoid accumulation in a similar manner as observed in melon.

\section*{DISCUSSION}

\textbf{Bulk segregant analysis of RNA-Seq and genetic mapping}

Bulk segregant analysis (BSA) is a method developed to identify genetic markers linked to a specific trait (Michelmore et al., 1991). BSA has been widely used ever since, uncovering
genetic markers and genes in a wide variety of plants. The rapid development of new sequencing technologies and computation systems allows a very efficient use of BSA while using next generation sequencing technologies to discover polymorphism that discriminate between the bulks (Liu et al., 2012; Trick et al., 2012). Combining BSA with RNA-Seq and detailed metabolite analyses enabled a relatively rapid and accurate isolation and characterization of CmKFB, the gene of interest in this work.

A high level of recombination occurred in the 2 Mbp region in the F3 segregating population (Fig. 4, D-F), Such a recombination rate in this genomic region was a significant factor enabling high resolution mapping, which resulted in less than 200 SNP that differentiate between the bulks, most of them localized in an 0.5 Mbp genomic region in scaffold 16 of chromosome 10 (Fig. 4A), while using a relatively small number of participants in each bulk (12 F3 families in the ‘yellow’ bulk and only 7 in the ‘white’). Recombinant analysis of the segregating population enabled to further fine map the genetic loci associated with NarCh accumulation to a genomic region of 100 Kb, which consisted of 17 protein coding genes. CmKFB was one of these 17 coding genes, and showed the most significant difference in expression between the ‘yellow’ and the ‘white’ F3 bulks, suggesting CmKFB as a primary candidate gene.

RNA-Seq analysis of ‘yellow’ (accumulating NarCh) and ‘white’ (do not accumulate NarCh) developing fruit of F3 segregant pools indicated that CmKFB is expressed at a very low level in fruits that accumulate NarCh while it is highly expressed in NarCh non-accumulating fruits. Due to the low abundance of CmKFB RNA-Seq reads obtained in the ‘yellow’ bulks we could not define sequence variation between the ‘yellow’ and ‘white’ alleles of this gene. DNA sequence analysis identified a 12-bp deletion in the 5’ UTR of the ‘white’ bulk and TVT. The 5’ UTR 12-bp deletion was consistently associated with flavonoid accumulation in our segregating population (Fig. 5D) and in different melon accessions (Supplemental Fig. S8C). Interestingly, monitoring allelic expression on cDNA (Fig. 5E), showed heterozygotes F3 to accumulate only the TVT non-flavonoid accumulating parent allele, suggesting CmKBF allelic expression to be affected by allelic variation. Motif analysis using JASPAR (Mathelier et al., 2014) suggested this deletion to disrupted a putative binding site of HMG-1 (High Mobility Group box), which is a chromatin
remodeling factor (Grosschedl et al., 1994; Webster et al., 1997; Ikeda et al., 2011), suggesting differential epigenetic allelic states to result in differential allelic CmKFB expression. Nevertheless, the association of the 12-bp deletion with CmKFB expression, the tissue specificity of this expression and the co-dominant nature of CmKFB expression need further studies.

Metabolites analysis and CmKFB function.

There is growing interest in manipulating the metabolic flux through different phenylpropanoid pathway branches for a number of reasons, for example, their association with plant tolerance to different biotic and abiotic stresses, commercial and consumer health impact of increased bioactive compounds in agricultural products, and alteration of the lignin biosynthesis pathway (Liu et al., 2014). An example of such metabolic flux alternation is the silencing of Arabidopsis HCT (hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase) that led to repression of lignin biosynthesis and redirection of the metabolic flux toward the flavonoid pathway (Besseau et al., 2007).

Various transcription factors were shown to divert the phenylpropanoid metabolic flux. ZmMYB3, a direct repressor of COMT and AtF5H genes, over-expression in Arabidopsis results in reduction in sinaopylmalate and lignin accumulation while enhancing accumulation of ferulic acid, coumaric acid and anthocyanins (Fornalé et al., 2010). In addition, the colorless fruit peel y mutant in tomato was found to be a mutation causing down regulation of SlMYB12, a TF that acts as a positive regulator of an array of phenylpropanoid structural genes. Interestingly, this mutation exhibits elevated accumulation of caffeic and ferulic compounds, accompanied by reduction in flavonoid content (Adato et al., 2009), similarly to the flux alteration observed in this work.

Biochemical analysis of melon fruit rinds showed that NarCh accumulation is accompanied by the accumulation of downstream flavonoids, as observed by DPBA staining.
(Fig. 1), Pew’s and Shinoda’s tests (Supplemental Fig. S3), and LC-MS analysis (Fig. 2). This accumulation is associated with significant decline in \( CmKFB \) transcription in a dosage dependent manner (Fig. 5), and reduction in caffeic acid, ferulic acid and aesculetin accumulation, which are branched from the phenylpropanoid to flavonoid pathway (Fig. 3). No significant differences in the expression pattern of structural genes of the flavonoid pathway were found based on our RNA-Seq data analysis (Supplemental Fig. S9), which could explain the biochemical differences between ‘yellow’ and ‘white’. The flux alternation demonstrated here can be explained by a negative post translational regulation of CmKFB at the entry point of the flavonoid pathway. This biochemical flux alternation was observed both when CmKFB was transiently expressed in melon leaves, resulting in elevation of salicylic acid, caffeic acid and aesculetin accumulation, accompanied by reduction in naringenin, luteolin and quercetin accumulation (Fig. 6), and by stable expression in tomato, resulting in elevated blue fluorescence accompanied by reduction in NarCh accumulation in fruits, and anthocyanins in leaves and flowers (Fig. 7; Supplemental Fig. S10). The functionality of CmKFB in the tomato system indicates the possible general role of this gene in plants. CmKFB is most probably involved in proteasome dependent protein degradation mechanism mediated by the SCF E3 ubiquitin ligase complex, as has been previously reported in \textit{Arabidopsis} (Zhang et al., 2013; Zhang et al., 2015). In these studies, four closely related KFB orthologs (AtKFB01, AtKFB20, AtKFB39 and AtKFB50) were found to target different PAL isozymes, which serve as the entry point to the phenylpropanoid pathway, to proteasome dependent degradation, resulting in reduced accumulation of a wide array of compounds: flavonols, anthocyanins, condensed tannins, sinapic acid esters and lignin. In this study, PAL doesn’t appear to be a suitable candidate as a target of CmKFB. The expression pattern of the single PAL ortholog (\textit{MELO3C025786}, Figure S9), which matches both the flavonoid and phenylpropanoid accumulation patterns during development (Supplemental Figs. S1 and S2) suggests that this gene is functionally active in both melon types. Phylogenetic analysis of melon and \textit{Arabidopsis} KFB proteins shows the four characterized AtKFBs and CmKFB to cluster in two distinct phylogenetic clades (Supplemental Figure S11), suggesting the presence of two distinctive functional groups, which are conserved between the \textit{Brassicaceae} and the \textit{Cucurbitales} orders,
further suggesting different substrate specificity. Observing the biochemical branch point between the two melon types, a natural candidate as a target of CmKFB is 4CL. This enzyme is encoded in all plants by multiple genes, which differ in substrate specificity of the paralogs, and is considered one of the main regulatory points in directing the phenylpropanoid metabolic pathway (Ehlting et al., 1999; Hamberger and Hahlbrock, 2004). This option is further supported by significant expression of multiple paralogs of 4CL in melon fruit (Supplemental Fig. S9). On the other hand, the degradation of MELO3C014767, which is the only significant ortholog of CHS expressed in melon fruit rind might also result in the biochemical differences which were observed. This work suggests that CmKFB regulates either 4CL or CHS enzymatic function via a protein degradation mechanism, either directly by targeting one of these enzymes to proteolytic degradation, or by modulating an additional factor, which regulates one of these enzymes function.

In summary, we have identified a Kelch repeat F-Box coding gene (CmKFB) which negatively regulates flavonoid accumulation in melon fruit by shifting the metabolic flux towards general phenylpropanoid products, including caffeic and ferulic acids and coumarin (aesculetin). Characterization of this natural variation in melon might prove useful for plant breeders in tailoring new cultivars.

MATERIALS AND METHODS

Plant material

A set of 18 melon cultivars were selected to represent rind color variation in Cucumis melo. They included the parental lines of the analyzed segregating population and representatives of five major taxonomic groups of melon, which accumulate or lack NarCh, with both climacteric and non-climacteric modes of fruit ripening (Supplemental Table S1). Plants were grown under standard conditions in the field and greenhouse at the Newe Ya’ar Research Center.
**F₃ population establishment**

A population segregating for fruit NarCh accumulation was established by crossing two casaba melons, ‘Noy Amid’ (NA) – a NarCh accumulating melon inbred line with a canary yellow fruit rind and ‘Tendral Verde Tardio’ (TVT) – a NarCh lacking inbred line with a green fruit rind. Two hundred and thirteen F₂ plants derived from this cross visually segregated to four fruit rind color phenotypes: green-yellow, green, yellow and white fitting the expected 9:3:3:1 Mendelian ratio for two independent genes, green and yellow. All F₂ fruit rinds were analyzed by HPLC for NarCh content (Tadmor et al., 2010). To define F₃ families that are homozygous to either NarCh fruit accumulation or non-accumulation, 50 F₃ families were selected, 30 with a yellow fruit rind F₂ progenitors and 20 with a white fruit rind F₂ progenitors. Thirty plants from each family were grown in the winter of 2010 at Ein Tamar in the Northern Arava Valley, Israel. Of these 50 families, 12 accumulating NarCh were selected phenotypically, the fruits of all plants in each family having yellow rinds. In addition, seven F₃ families, non-accumulating NarCh were selected phenotypically, for the fruits of all plants in each family having white rinds. Thirty plants of each of these 19 families were grown in open field of Newe Ya’ar in the spring of 2010. Female flowers were tagged on the day of anthesis and three biological replicates of fruit rind from each F₃ family were sampled at three fruit developmental stages: 10 days after anthesis (DAA), 20DAA, and ripe.

**RNA-Seq, data processing and de novo assembly**

Fruit rind RNA extraction for RNA-Seq analysis was performed according to Portnoy et al. (2011). Twelve samples (two replicates of three developmental stages: 10 DAA, 20 DAA and mature fruit rind of 'yellow' and 'white' F₃ families) were used to construct RNA-Seq libraries following the standard Illumina protocol. The libraries were sequenced on the Illumina HiSeq2000 system with the paired-end mode. The sequencing was performed at the Biotechnology Center at the University of Illinois at Urbana-Champaign. The raw Illumina RNA-Seq reads were first processed to remove low quality regions and adaptor sequences using an in-house perl script. The RNA-Seq reads were further aligned to the ribosomal RNA database (Quast et al., 2013) using Bowtie (Langmead et al., 2009) allowing two mismatches and the
mappable reads were discarded. The resulting high-quality cleaned reads were assembled de novo into contigs using Trinity with “min_kmer_cov” set to 2 (Grabherr et al., 2011). To remove the redundancy of Trinity-generated contigs, they were further assembled de novo using iAssembler with minimum percent identify (-p) set to 99 (Zheng et al., 2011).

Gene expression quantification and differential expression analysis
The high-quality cleaned RNA-Seq reads were aligned to the assembled melon transcripts with the Bowtie program (Langmead et al., 2009) allowing one mismatch. Following alignments, raw counts for each melon transcript and each sample were derived and normalized to reads per kilobase of exon model per million mapped reads (RPKM). Differentially expressed genes between white and yellow fruits at the same stages were identified with the DESeq package (Anders and Huber, 2010). Raw p-values of multiple tests were corrected using FDR (Benjamini and Hochberg, 1995).

SNP and small indel identification from the RNA-Seq data
To identify SNPs and small indels between yellow and white genotypes, identical RNA-Seq reads from each library were first collapsed into a single sequence. The resulting unique reads were then aligned to the assembled melon transcripts using BWA (Li and Durbin, 2009) allowing one mismatch, one gap opening and one gap extension. Following mapping, SNPs and small indels were identified based on the mpileup files generated by SAMtools (Li et al., 2009). The identified SNPs and small indels were supported by at least five independent reads and had an allele frequency of at least 0.8.

Phylogenetic analysis
Melon KFB protein sequences were retrieved through BLAST searches in the melon genome website (https://melonomics.net/) and conserved domains were verified against the IterPro protein database. Arabidopsis KFB homologs were adopted from Zhang et al. (2015). Protein sequences were aligned by the ClustalW program. Phylogenetic analysis was performed using the MEGA package version 6 (Tamura et al., 2013) using the neighbor-joining method, with the pairwise deletion option.
Fluorescent probes

UN22458 SNP (SNP no. 86, Supplemental Table S6) genotyping was performed in ABI PRISM 7000 sequence detection system (Applied Biosystems). Each reaction contained 0.25 µL Assay mix (SNP genotyping mix, Biosearch Technologies), 5 µL reaction mix (Genotyping ToughMix, ROX, Quanta Biosystems), 1 µL DNA, and 3.75 µL water. UN06224 and UN10338 SNP (SNPs no. 39 and 54, Supplemental Table S6) genotyping was performed in StepOnePlus Real-Time PCR system (Applied Biosystems), and each reaction contained 4 µL Mix (TaqMan genotyping master mix, Applied Biosystems), 0.4 µL Assay mix (TaqMan SNP genotyping Assay, Applied Biosystems), 0.8 µL DNA, and 2.8 µL water. Primers are listed in Supplemental Table S8.

High Resolution Melt (HRM) analysis

DNA samples were diluted into equal concentrations of 100 ng/µL using NanoDrop. Samples were amplified on an Eco Real-Time PCR system (Illumina). Each sample contained 1 µL DNA, 0.2 µL of 10 mM of each primer, 5 µL reaction mix (FastSYBR green master mix, Applied Biosystems), and 3.6 µL ddH₂O, programmed as specified by the enzyme manufacturer, and analyzed with the Eco version 4 software. Primers for each HRM SNP marker are specified in Supplemental Table S8.

Virus-induced gene expression

A ZYMV-AGII (AGII) potyvirus-based vector system (Arazi et al., 2001) was used. To overexpress CmKFB, a GFP AGII vector was digested with PstI and Sall enzymes to remove the GFP fragment. CmKFB coding sequence (CDS) was amplified by PCR from TVT melon cDNA with primers containing linkers to the corresponding PstI and Sall restriction sites (Supplemental Table S8), and ligated to the digested vector (T4 DNA ligase, Thermo Scientific). Particle bombardment inoculation of AG-GFP and AG-CmKFB cDNA was performed using a handgun device (Gal-On et al., 1997) on cotyledons of 14 day old NA melon seedlings. After inoculation, seedlings were placed in a growth chamber, with a 16h photoperiod and at 25°C for 14 days.

Tomato stable transformation

CmKFB was amplified from TVT melon cDNA in two PCR steps, isolating the CDS with additional
39-bp linkers on each side of the gene, generating truncated AttL sites for direct recombination to pK7WG2D Gateway binary destination vector (Invitrogen) by LR clonase (Fu et al., 2008; Ischebeck et al., 2011; Supplemental Table S8). Cotyledons of the ‘MP1’ tomato were transformed with *Agrobacterium tumefaciens* GV3101 that harbored the vector. After 48 h of co-cultivation in the dark (3% glucose, 0.1 mg/L IAA, 1 mg/L Zeatin, 100 µM AS), explants were transferred to a shoot-induction medium (3% glucose, 0.1 mg/L IAA, 1 mg/L Zeatin, 100 mg/L kanamycin, 500 mg/L cloran), under light. Regenerated shoots were moved to a shoot-elongation medium (3% glucose, 0.1 mg/L Zeatin, 100 mg/L kanamycin, 500 mg/L cloran). Elongated shoots were transferred to a rooting medium (1/2 MS, 2% sucrose, 2 mg/L IBA).

**RT-PCR**

Reactions were performed on an Eco RT-PCR system (Illumina). Each sample contained 1µL cDNA, 0.2 µL of each primer (10 mM), 5 µL reaction mix (FastSYBR green master mix, Applied Biosystems), and 3.6 µL ddH₂O, programmed as specified by manufacturer, and analyzed with Eco version 4 software.

**Flavonoid extraction and staining methods**

Fresh fruit rind or leaf samples were collected, immediately frozen in liquid nitrogen, and stored at -80°C. Before extraction, samples were ground to a fine powder with a mortar and pestle in the presence of liquid nitrogen. Two grams of ground tissue were extracted in 5 mL methanol (MeOH), vortexed for 10 seconds, incubated for 10 minutes in 60°C and vortexed again. Samples were then centrifuged at 10,000 g for 5 minutes and the upper methanol phase was collected. For DPBA (diphenyl boric acid aminoethyl ester) staining, 1 µL of the methanolic extract was applied on an HPTLC (silica gel 60, Merck) and developed with ethyl acetate:acetic acid:H₂O in ratio 67:17.1:17.1 (v/v). After separation, the plate was dried with a hair dryer, sprayed with DPBA (Sigma) dissolved in MeOH (1% w/v), dried again, and visualized under a UV light (320 nm).

Additional flavonoid staining protocols performed on the methanol extracts included the Zinc/HCl test (Pew’s Test), for which Zinc powder and a drop of 5N HCl were added to the plant
flavonoid extract, and the Magnesium/HCl test (Shinoda’s Test) which was applied in the same way as the Zinc/HCl test, except that magnesium wire was used instead of Zinc powder.

**LC-MS**

The method used was modified from De Vos et al. (2007). Methanol extraction was carried out as described above. Enzymatic hydrolysis was conducted using 1 mL of the evaporated methanol extracts. Samples were speed-vac evaporated and supplemented with an enzymatic solution containing 1 mg/mL of each: cellulase, pectinase, and \( \beta \)-glucosidase (Sigma), dissolved in 0.1M NaAc pH 5.3 solution, and incubated at 37°C overnight. Aglycones were extracted with 500 mL ethyl acetate. The upper ethyl acetate phase was removed, evaporated, and pellet re-dissolved in 300 µL MeOH. Both samples, before and after hydrolysis, were supplemented with ddH₂O and formic acid (final concentrations of 25% and 0.1% respectively) and filtered through a PFE 0.2 µm filter (Pall Life Sciences).

Mass spectral analysis was conducted using an LC-MS-TOF (LC: 1290 infinity system Agilent Technologies, MS-TOF: 6224-TOF-LC-MS Agilent Technologies), in ESI negative mode. Separation of compounds was performed using Zorbax Extended C18 RRHT 2.1*50 mm, 1.8 µm column. The mobile phase consisted of 0.1% formic acid in water (Phase ‘A’), and 0.1% formic acid in acetonitrile (phase ‘B’). A gradient program was used as follows: 95% to 65% ‘A’ over 15 minutes, 65% to 25% ‘A’ over 40 seconds, 25% ‘A’ held over 2 minutes, 25% to 95% ‘A’ over 1.5 minutes, 95% held over 2 minutes at constant flow rate of 0.19 ml/min. Analysis of the masses was performed with the MassHunter qualitative analysis software (version B.05.00, Agilent Technologies). Determination of differential mass expression between samples was performed using the Mass Profiler Professional (MPP) software (Agilent Technologies).

**Microscopy**

DPBA-stained tissue of handmade slices were observed with an Olympus BX61 microscope coupled with a U-HGLGPS illumination system, Olympus 20x/0.50 objective, using a U-MNBV2 filter (excitation 420-440nm, emission >475nm). Images were collected with a digital camera (DP73, Olympus), and processed with cellSens Dimension software (Olympus).
Tomato seedling fluorescence was observed with a Leica M205FA stereomicroscope, Leica 10472649 Planapo 0.63x objective, under a GFP3 filter. Images were collected with a digital camera (DFC495, Leica) and processed with the Leica Application Suit V3.8 software.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Center for the Improvement of Cucurbit Fruit Quality, ARO, Israel. We also acknowledge Uzi Sa’ar and Fabian Baumkoler for expert maintenance of the plants in the field and the greenhouse and Dalia Wolf for preparation of the transgenic tomato plants. We thank Harry Paris for editing the manuscript. The authors have no conflict of interest to declare. Publication No. xx/2015 of the Agricultural Research Organization, Bet Dagan, Israel.

AUTHOR CONTRIBUTIONS

A.F., J. B., Y. T. designed the experiments; A. F. performed the majority of the experiments as partial fulfillment of the requirements for the degree of Doctor of Philosophy; E. L. provided biochemical experimental support; N. K., A. A. S., V. P. provided genomic experimental support; A. M., R. D. provided metabolomics experimental support; A. G. provided VIGE experimental support; S.G., Z. F. conducted bioinformatics analyses; Y. K. co-supervised A. F.; A. F. and Y. T. wrote the paper. All authors discussed the results, and edited the manuscript.
**FIGURE LEGENDS**

**Figure 1.** DPBA staining of melon fruit rinds. Microscopic view of fruit rinds of the parental lines NA (A, B, E and F) and TVT (C, D, G and H). A, C, E and G are fruit rind cross section, B, D, F and G are upper view of fruit rind. BL, bright visible light; UV, UV light after DPBA staining. I, TLC of methanol extract stained with DPBA of NA (left lane), TVT (right lane), marked by a yellow and a white lines are the 12 F3 yellow families and 7 F3 white families, respectively.

**Figure 2.** LC-MS analysis of fruit rind methanol extracts after enzymatic hydrolysis. A-D, Base peak chromatograms (BPC) of mass range m/z 100-1000, in ESI negative mode. A, NA parent. B, Bulk of ‘yellow’ F3. C, TVT parent. D, Bulk of ‘white’ F3. E, A volcano plot indicating specific compounds that differentially accumulated (P< 0.05, minimum 10 fold difference) in the NA & ‘yellow’ F3 group (red) or in the TVT & ‘white’ F3 group (blue).

**Figure 3.** Phenylpropanoid compounds accumulated in ‘yellow’ and ‘white’ melon fruit rinds and their corresponding biosynthesis pathway. Six compounds were specifically accumulated in the NarCh accumulating lines (framed in yellow): naringenin chalcone, naringenin, luteolin, dihydrokaempferol, kaempferol, and quercetin. Three compounds were found to be specifically accumulated in the non-NarCh accumulating lines (framed in blue): caffeic acid, ferulic acid, and aesculetin.

**Figure 4.** Genetic mapping of the locus governing flavonoid accumulation. A, Physical positions of the 164/179 SNP which significantly differentiated between the bulks on scaffold 16 of the melon genome. Red triangles represent position of nine selected HRM markers. B, Phenotypes of individual F3 fruits included in the recombinants analysis: white (W), pale yellow (Int) and yellow (Y) fruit. C, Segregation (%) of Y, Int and W phenotypes within each of the 14 F3 families and among all analyzed individual F3 plants (Total). D-F, Distribution of TVT (T), heterozygote (H) and NA (N) alleles of the three fluorescent SNP probes (UN22458, UN10338 and UN06224) within white (W), pale yellow (Int) and yellow (Y) phenotypes of individual F3 plants, physical position of each probe is indicated by a blue arrow. G, Fine mapping with the nine HRM markers (red triangles 1-9). Below are individual plants segregating to these HRM markers and exhibiting defined phenotypes (white, pale yellow and yellow). N, NA allele; H, heterozygous; and T, TVT allele. Bottom red line indicates the 100 kbp genomic region that shows 100% co-segregation of genotype and phenotype.
Figure 5. Analysis of \textit{CmKFB} gene and biochemical differences in representative segregating F\textsubscript{3} plants. A, Three visual phenotype groups, yellow, light yellow and white fruit. B, qRT-PCR of \textit{CmKFB} performed on five individual F\textsubscript{3} plants within each phenotypic group. C, Methanol extraction of each individual fractionated by TLC and visualized under UV light after DPBA staining. D, Agarose gel of PCR products of genomic DNA fragments containing the 5’ UTR deletion. E, Similar PCR analysis performed on cDNA.

Figure 6. Virus induced \textit{CmKFB} expression in melon leaves. A, Control (uninfected and VIGE-GFP) and VIGE-C\textit{mKFB} leaves observed through a GFP filter under UV light. B, PCR amplification of the inserted cDNA products. C, \textit{CmKFB} expression analyzed by qRT-PCR. D, PCR amplification of \textit{CmKFB} CDS cDNA. E, TLC fractionation of leaf methanol extracts, observed under UV after DPBA staining. F-H, base peak chromatograms (BPC), after enzymatic hydrolysis, of mass range m/z 100-1000, measured in LC-MS. F, Uninfected control leaf. G, VIGE-GFP leaf. H, VIGE-C\textit{mKFB} leaf. I-J, Quantitative comparison of differentiating compounds in leaf methanol extract. Compounds were identified by comparisons with authentic standards. I, Up-regulated in VIGE-C\textit{mKFB} (red) compared to controls (black). J, Down-regulated in VIGE-C\textit{mKFB} (red) compared to controls (black).

Figure 7. Over expression of \textit{CmKFB} under the control of the CaMV 35S promoter in stable transformed tomato plants. A, Visual comparison of a control and a T\textsubscript{0} fruit peels. B, Microscopic view of a control and a T\textsubscript{0} cuticles. C, Microscopic view of a control and a T\textsubscript{0} fruit cut under UV light after DPBA staining (arrow marks the cuticle upper layer that did not emit any fluorescence). D, Leaves of T\textsubscript{1} segregants of \textsuperscript{1}st transgenic plant and MP1 control grown under cold stress conditions, upper panel-visible light, lower panel – GFP fluorescence.
Supplemental Data

LARGE DATA SETS

The raw RNA-Seq data has been deposited in NCBI SRA under the accession number SRP060002.

Supplemental Figure S1. F3 population establishment. Developing fruit of parental lines NA and TVT and representative F3 families: ‘White’ – represents developing fruit of 7 families with low chlorophyll and low NarCh contents, ‘Yellow’ – represents developing fruit of 12 families with high NarCh and low chlorophyll contents. Three developmental stages (framed in red), consisting of 3 biological repetitions of each family fruit rind, were subjected to further biochemical and RNA-Seq analyses.

Supplemental Figure S2. Visualization of DPBA stained TLC of methanol extracts under UV light. A, parental lines (NA and TVT) and corresponding F3 bulks (Y (yellow) and W (white)) of leaf tissue on the left, and 1 mm flowers on the right. B, similar TLC performed on one day after anthesis flower petals, on the left, and receptacle on the right. C-E, fruit rind of 12 ‘yellow’ (Y) F3 families on the left and 7 ‘white’ (W) F3 families on the right, at: C, 10DAA. D, 20ADD. E, ripe stage.

Supplemental Figure S3. Additional differentiating staining methods. Methanol extracts of ‘yellow’ F3 bulk (Y F3) and ‘white’ F3 bulk (W F3) analyzed for a visual color change with Zn/HCl (Pew’s test) or Mg/HCl (Shinoda’s test) in ELISA plate wells.

Supplemental Figure S4. LC-MS analysis of fruit rind methanol extracts. A-D, base peak chromatograms (BPC) of mass range m/z 200-800, in ESI negative mode. A, NA parent. B, Bulk of ‘yellow’ F3. C, TVT parent. D, Bulk of ‘white’ F3. E, a volcano plot indicating specific compounds that differentially accumulated (P< 0.05, minimum 10 fold difference) in the NA & ‘yellow’ F3 group (red) or in the TVT & ‘white’ F3 group (blue).

Supplemental Figure S5. Fruit rind compound identification. Identification of eight compounds, significantly differentiating the two NarCh accumulating groups. A, Three compounds over-accumulated in both’ TVT’, and ‘white’ F3 bulk: caffeic acid, ferulic acid and aesculetin. Each specific mass identified (upper panel) with an analytical grade authentic standard (lower
B, Five compounds over-accumulated in the NarCh accumulating group (NA parent, and ‘yellow’ F₃ bulk): naringenin, kaempferol, luteolin, dihydrokaempferol (traces), and quercetin. Each specific mass identified (upper panel) with analytical authentic standard (lower panel).

**Supplemental Figure S6.** The melting curves of HRM markers. High resolution melting curves of control: NA, TVT and heterozygous samples to each of the nine markers. Heterozygous colored in green.

**Supplemental Figure S7.** CmKFB and At1g23390 protein sequence alignment.

**Supplemental Figure S8.** Molecular analysis of the candidate gene *CmKFB*. A, *CmKFB* digital expression (RPKM) pattern of ‘white’ and ‘yellow’ bulks at three fruit developmental stages. B, The 12-bp deletion found in the *CmKFB* 5'-UTR of TVT (the non-flavonoid accumulating parent). C, Agarose gel fractionation of PCR amplification of DNA fragments containing the 5'-UTR deletion performed on 14 melon accessions in addition to NA and TVT. NarCh accumulating varieties are indicated in yellow. D, Relative expression analysis of *CmKFB* by qRT-PCR in fruits and in leaves of yellow (Y-F₃) and white (W-F₃) F₃ bulks.

**Supplemental Figure S9.** Digital expression pattern of the gene families leading to the flavonoid pathway. Gene family members of PAL, C4H, 4CL, CHS, were predicted by the genome annotations (http://www.melonomics.net). For each gene the average digital expression (RPKM) during three fruit developmental stages (10 DAA, 20 DAA and ripe fruit) in ‘white’ (white bars) and ‘yellow’ (yellow bars) bulks is presented.

**Supplemental Figure S10.** Over-expression of *CmKFB* in stable transgenic T₀ tomato plants. A-D, T₀ transgenic plant (right) and the control (left). A, Leaves under visible light after cold stress. B, GFP fluorescence of same leaves. C, Developing shoot meristem of mature plant with flower buds. D, Flower petals (note the purple line in the middle of control's flower petals).

**Supplemental Figure S11.** Phylogenetic analysis of KFB proteins in melon and Arabidopsis. Subtree colored in blue contains the four Arabidopsis KFB proteins, KFB01 (At1g15670), KFB20 (At1g80440), KFB39 (At2g44130), and KFB50 (At3g59940), which were shown to regulate PAL degradation (Zhang et al., 2015), in addition to the melon ME3C014678 and MELO3C009596. Subtree colored in red contains CmKFB (MELO3C011980) and the Arabidopsis At1g23390.
Supplemental Table S1. List of melon varieties included in this study. Taxonomic groups were based on Pitrat (2000).

Supplemental Table S2. List of differentiating compounds between the two melon groups (NA+Y vs. TVT+W). Cpd - compound number; RT - the retention time; Diff - difference between theoretical and found m/z values in ppm.

Supplemental Table S3. List of differentiating compounds after hydrolysis between the two melon groups (NA+Y vs. TVT+W). Cpd - compound number; Ret. Time - the retention time; Found m/z - mass found; Theor. m/z - theoretical monoisotopic mass calculated for the ion (M-H)-; Diff(ppm) - difference between theoretical and found m/z values in ppm.

Supplemental Table S4. List of fruit rind compounds identified by comparison with authentic standards. Ret. Time- retention time (minutes), ESI(-) m/z found – exact mass recorded, ESI(-) m/z theor – exact mass predicted, Diff(ppm) – deviation from the predicted mass, Formula – predicted formula generated by masshunter software.

Supplemental Table S5. RNA-Seq statistics.

Supplemental Table S6. List of 179 SNPs differentiating between 'yellow' and 'white' bulks. ‘UN number’ is the unigene number designated to each gene during the de novo assembly of the melon transcriptome; 'UN SNP position' designates the position of the SNP in the UN; 'melo' designates the name of the UN in the melon genome web site https://melonomics.net/; 'SNP position' designates the SNP position on the scaffold; "Y" and "W" base are the yellow and white bases at the SNP location in parenthesis is the number of the specific 'base' reads; "Y" and "W" coverage are the number of reads of the specific 'UN' in the yellow and white bulks.

Supplemental Table S7. List of genes expressed in the physical region governing flavonoid accumulation. Gene name (according to https://melonomics.net) and our de novo assembly (UN). RPKM of 'yellow' (Y) and 'white' (W) fruit bulks, at 10, 20 DAA and at the mature fruit stage (R). At each time point, two replicates were compared and adjusted P value was calculated. Each gene was annotated.

Supplemental Table S8. List of primers and probes.

Supplemental Table S9. List of F-box Kelch repeat protein coding genes in melon.
LITERATURE CITED


Methods 64: 103–110


Imperato F (1980) Five plants of the family Cucurbitaceae with flavonoid patterns of pollens different from those of corresponding stigmas. Experientia 36: 1136–1137


alignment of short DNA sequences to the human genome. Genome Biol 10: R25


Trick M, Adamski N, Mugford SG, Jiang CC, Febrer M, Uauy C (2012) Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. BMC Plant Biol 12: 14


Webster CI, Packman LC, Pwee KH, Gray JC (1997) High mobility group proteins HMG-1 and HMG-I/Y bind to a positive regulatory region of the pea plastocyanin gene promoter. Plant J 11: 703–711


Figure 1. DPBA staining of melon fruit rinds. Microscopic view of fruit rinds of the parental lines NA (A, B, E and F) and TVT (C, D, G and H). A, C, E and G are fruit rind cross section, B, D, F and G are upper view of fruit rind. BL, bright visible light; UV, UV light after DPBA staining. I, TLC of methanol extract stained with DPBA of NA (left lane), TVT (right lane), marked by a yellow and a white lines are the 12 F$_3$ yellow families and 7 F$_3$ white families, respectively.
Figure 2. LC-MS analysis of fruit rind methanol extracts after enzymatic hydrolysis. A-D, Base peak chromatograms (BPC) of mass range m/z 100-1000, in ESI negative mode. A, NA parent. B, Bulk of ‘yellow’ F₃. C, TVT parent. D, Bulk of ‘white’ F₃. E, A volcano plot indicating specific compounds that differentially accumulated (P< 0.05, minimum 10 fold difference) in the NA & ‘yellow’ F₃ group (red) or in the TVT & ‘white’ F₃ group (blue).
Figure 3. Phenylpropanoid compounds accumulated in ‘yellow’ and ‘white’ melon fruit rinds and their corresponding biosynthesis pathway. Six compounds were specifically accumulated in the NarCh accumulating lines (framed in yellow): naringenin chalcone, naringenin, luteolin, dihydrokaempferol, kaempferol, and quercetin. Three compounds were found to be specifically accumulated in the non-NarCh accumulating lines (framed in blue): caffeic acid, ferulic acid, and aesculetin.
Figure 4. Genetic mapping of the locus governing flavonoid accumulation. A, Physical positions of the 164/179 SNP which significantly differentiated between the bulks on scaffold 16 of the melon genome. Red triangles represent position of nine selected HRM markers. B, Phenotypes of individual F₃ fruits included in the recombinants analysis: white (W), pale yellow (Int) and yellow (Y) fruit. C, Segregation (%) of Y, Int and W phenotypes within each of the 14 F₃ families and among all analyzed individual F₃ plants (Total). D-F, Distribution of TVT (T), heterozygote (H) and NA (N) alleles of the three fluorescent SNP probes (UN22458, UN10338 and UN06224) within white (W), pale yellow (Int) and yellow (Y) phenotypes of individual F₃ plants, physical position of each probe is indicated by a blue arrow. G, Fine mapping with the nine HRM markers (red triangles 1-9). Below are individual plants segregating to these HRM markers and exhibiting defined phenotypes (white, pale yellow and yellow). N, NA allele; H, heterozygous and T, TVT allele. Bottom red line indicates the 100 kbp genomic region that shows 100% co-segregation of genotype and phenotype.
Figure 5. Analysis of *CmKFB* gene and biochemical differences in representative segregating F$_3$ plants. A, Three visual phenotype groups, yellow, light yellow and white fruit. B, qRT-PCR of *CmKFB* performed on five individual F$_3$ plants within each phenotypic group. C, Methanol extraction of each individual fractionated by TLC and visualized under UV light after DPBA staining. D, Agarose gel of PCR products of genomic DNA fragments containing the 5’ UTR deletion. E, Similar PCR analysis performed on cDNA.
Figure 6. Virus induced *CmKFB* expression in melon leaves. A, Control (uninfected and VIGE-GFP) and VIGE-*CmKFB* leaves observed through a GFP filter under UV light. B, PCR amplification of the inserted cDNA products. C, *CmKFB* expression analyzed by qRT-PCR. D, PCR amplification of *CmKFB* CDS cDNA. E, TLC fractionation of leaf methanol extracts, observed under UV after DPBA staining. F-H, base peak chromatograms (BPC), after enzymatic hydrolysis, of mass range m/z 100-1000, measured in LC-MS. F, Uninfected control leaf. G, VIGE-GFP leaf. H, VIGE-*CmKF* leaf. I-J, Quantitative comparison of differentiating compounds in leaf methanol extract. Compounds were identified by comparisons with authentic standards. I, Up-regulated in VIGE-*CmKFB* (red) compared to controls (black). J, Down-regulated in VIGE-*CmKFB* (red) compared to controls (black).
Figure 7. Over expression of CmKFB under the control of the CaMV 35S promoter in stable transformed tomato plants. A, Visual comparison of a control and a T₀ fruit peels. B, Microscopic view of a control and a T₀ cuticles. C, Microscopic view of a control and a T₀ fruit cut under UV light after DPBA staining (arrow marks the cuticle upper layer that did not emit any fluorescence). D, Leaves of T₁ segregants of 1st transgenic plant and MP1 control grown under cold stress conditions, upper panel-visible light, lower panel – GFP fluorescence.


Haribal M, Renwick JAA (1996) Oviposition stimuliants for the monarch butterfly: flavonol glycosides from Asclepias curassavica. Phytochemistry 41: 139-144


Imperato F (1980) Five plants of the family Cucurbitaceae with flavonoid patterns of pollens different from those of corresponding stigmas. Experientia 36: 1136-1137


Trick M, Adamski N, Mugford SG, Jiang CC, Febrer M, Uauy C (2012) Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. BMC Plant Biol 12: 14


Webster CI, Packman LC, Pwee KH, Gray JC (1997) High mobility group proteins HMG-1 and HMG-I/Y bind to a positive regulatory region of the pea plastocyanin gene promoter. Plant J 11: 703-711


