Generation and characterisation of an *Arabidopsis thaliana* *f3h*/*fls1*/*ans* triple mutant that accumulates eriodictyol derivatives

Hanna Marie Schilbert [1]* ORCID: 0000-0003-0474-7753

Mareike Busche [1]* ORCID: 0000-0002-2114-7613

Vania Sáez [2] ORCID: 0000-0002-6477-6671

Andrea Angeli [2] ORCID: 0009-0003-8882-5858

Bernd Weisshaar [1] ORCID: 0000-0002-7635-3473

Stefan Martens [2] ORCID: 0000-0001-9913-4882

Ralf Stracke [1#] ORCID: 0000-0002-9261-2279

[1] Bielefeld University, Faculty of Biology, Genetics and Genomics of Plants, 33615 Bielefeld, Germany

[2] Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige (TN), 38098, Italy

* These two authors contributed equally to this work.

[#] Corresponding author: ralf.stracke@uni-bielefeld.de

Abstract

Background: Flavonoids are plant specialised metabolites, which derive from phenylalanine and acetate metabolism. They possess a variety of beneficial characteristics for plants and humans. Several modification steps in the synthesis of tricyclic flavonoids cause for the amazing diversity of flavonoids in plants. The 2-oxoglutarate-dependent dioxygenases (2-ODDs) flavanone 3-hydroxylase (F3H, synonym FHT), flavonol synthase (FLS) and anthocyanidin synthase (ANS, synonym leucoanthocyanidin dioxigenase (LDOX)), catalyse oxidative modifications to the central C ring. They are highly similar and have been shown to catalyse, at least in part, each other's reactions. FLS and ANS have been identified as bifunctional enzymes in many species, including *Arabidopsis thaliana*, stressing the capability of plants to bypass missing or mutated reaction steps on the way to flavonoid production. However, little is known about such bypass reactions and the flavonoid composition of plants lacking all three central flavonoid 2-ODDs.

Results: To address this issue, we generated a *f3h*/*fls1*/*ans* mutant, as well as the corresponding double mutants and investigated the flavonoid composition of this mutant collection. The *f3h*/*fls1*/*ans* mutant was further characterised at the genomic level by analysis of a nanopore DNA sequencing generated genome sequence assembly and at the transcriptomic level by RNA-Seq analysis. The mutant collection established, including the novel double mutants *f3h*/*fls1* and *f3h*/*ans*, was used to validate and analyse the multifunctionalities of F3H, FLS1, and ANS *in planta*. Metabolite analyses revealed the accumulation of eriodictyol and additional glycosylated derivatives in mutants carrying the *f3h* mutant allele, resulting from the conversion of naringenin to eriodictyol by flavonoid 3'-hydroxylase (F3'H) activity.

Conclusions: We describe the *in planta* multifunctionality of the three central flavonoid 2-ODDs from *A. thaliana* and identify a bypass in the *f3h*/*fls1*/*ans* triple mutant that leads to the formation of eriodictyol derivatives. As (homo-)eriodictyols are known as bitter taste maskers, the annotated eriodictyol (derivatives) and in particular the observations made on

their *in planta* production, could provide valuable insights for creating of novel food supplements.

Keywords: 2-oxoglutarate-dependent dioxygenases, enzyme promiscuity, flavonoid biosynthesis, flavonoids, multifunctionality, triple 2-ODD mutant

Background

Flavonoids are a large group of specialised metabolites with an estimated number of 9,000 flavonoid derivatives and a wide distribution throughout the plant kingdom (1). Flavonoids share a basic structure with two aromatic C6 rings and a heterocyclic ring (2). Reorganisation or modification, such as oxidation or glycosylation, is mainly responsible for the wide variety of flavonoids. They can be classified into different subgroups including flavones, flavanones, flavonols, dihydroflavonols, flavandiols, anthocyanins, and proanthocyanidins (2). These subgroups exhibit manifold functions. Among others, anthocyanins promote pollination and seed dispersal by colouring flowers and fruits (3-5), flavonols play a role in auxin transport, plant fertility, and UV protection (6-8), and proanthocyanins serve as deterrent metabolites for defence against bird predators (9). Moreover, flavonoids are beneficial for human nutrition due to their antioxidative, antigenotoxic and anticarcinogenic properties (10, 11). These health-promoting effects have led to a strong interest in increasing the flavonoid content in plants (12, 13). However, as some flavonoids exhibit antinutritional properties, like a bitter-off taste (14), the engineering of flavonoid composition is of high relevance for the breeding of plants with desirable flavonoid profiles.

Flavonoid biosynthesis (Figure 1) is one of the best characterised pathways in plant specialised metabolism. Due to the visible phenotypes of respective mutants which often display a transparent testa resulting in yellow to pale-brown seed (15) and the large collections of T-DNA- and transposon-tagged lines for *Arabidopsis thaliana* (16-18), the genes involved in flavonoid biosynthesis are well studied. The flavonoid pathway is initiated by its key enzyme chalcone synthase (CHS), which condensates *p*-coumaroyl-CoA and three molecules of malonyl-CoA followed by cyclisation, leading to the formation of naringenin chalcone (19). A heterocyclic ring is introduced spontaneously or through the catalysis of chalcone isomerase (CHI), forming naringenin (20). Flavanone 3-hydroxylase (F3H or FHT) converts naringenin, a flavanone, to dihydrokaempferol, a dihydroflavonol, by hydroxylation at the C-3 position (21). In alternative branches, the flavonoid 3'-hydroxylase (F3'H) converts naringenin to eriodictyol and dihydrokaempferol to dihydroquercetin (22). Naringenin can also be converted to flavones

82 by flavone synthaseI or II (FNSI or II) (23, 24). However, flavones were thought to be absent
83 in *A. thaliana* until two derivatives (7-2″,3″-diacetylglucoside and pentamethoxydihydroxy
84 flavone) were identified in Columbia-0 (Col-0) leaves (25, 26). DOWNY MILDEW
85 RESISTANT6 (*Ath*DMR6), a 2-oxoglutarate-dependent dioxygenase (2-ODD), has been
86 proposed as the enzyme responsible for their biosynthesis (26). The enzyme flavonol synthase
87 (FLS) uses dihydroflavonols to form flavonols by introducing a double bond between C-2 and
88 C-3 (27). Alternatively, dihydroflavonol 4-reductase (DFR) can reduce dihydroflavonols to the
89 corresponding leucoanthocyanidins, thus competing with FLS for substrates (28, 29).
90 Anthocyanidin synthase (ANS, also known as leucoanthocyanidin dioxygenase (LDOX))
91 converts leucoanthocyanidins to anthocyanidins (30). Anthocyanidins can be modified by
92 glycosyltransferases (GTs) to form anthocyanins, or they can be converted to
93 proanthocyanidins by anthocyanidin reductase (ANR) (31, 32). Besides anthocyanidins, other
94 flavonoids like flavonols or flavanones can be modified by GTs to form the respective
95 glycosides (33-35).
96 F3H, FNSI, FLS and ANS belong to the 2-ODD enzyme family (36). 2-ODDs represent the
97 second largest enzyme family in plants (37). They are ubiquitously spread across the plant
98 kingdom but also microorganisms and are involved in a variety of primary and specialised
99 metabolism, including the synthesis of specialised metabolites, hormones, signalling
100 molecules, amino acid modification and DNA-repair (summarised in 38, 39). Plant 2-ODDs can
101 be classified into three classes, DOXA, DOXB and DOXC, based on amino acid sequence
102 similarity, with DOXC harbouring 2-ODDs involved in the synthesis of specialised metabolites,
103 such as flavonoids, including F3H and FNSI (DOXC28) and FLS and ANS (DOXC47) (37).
104
105

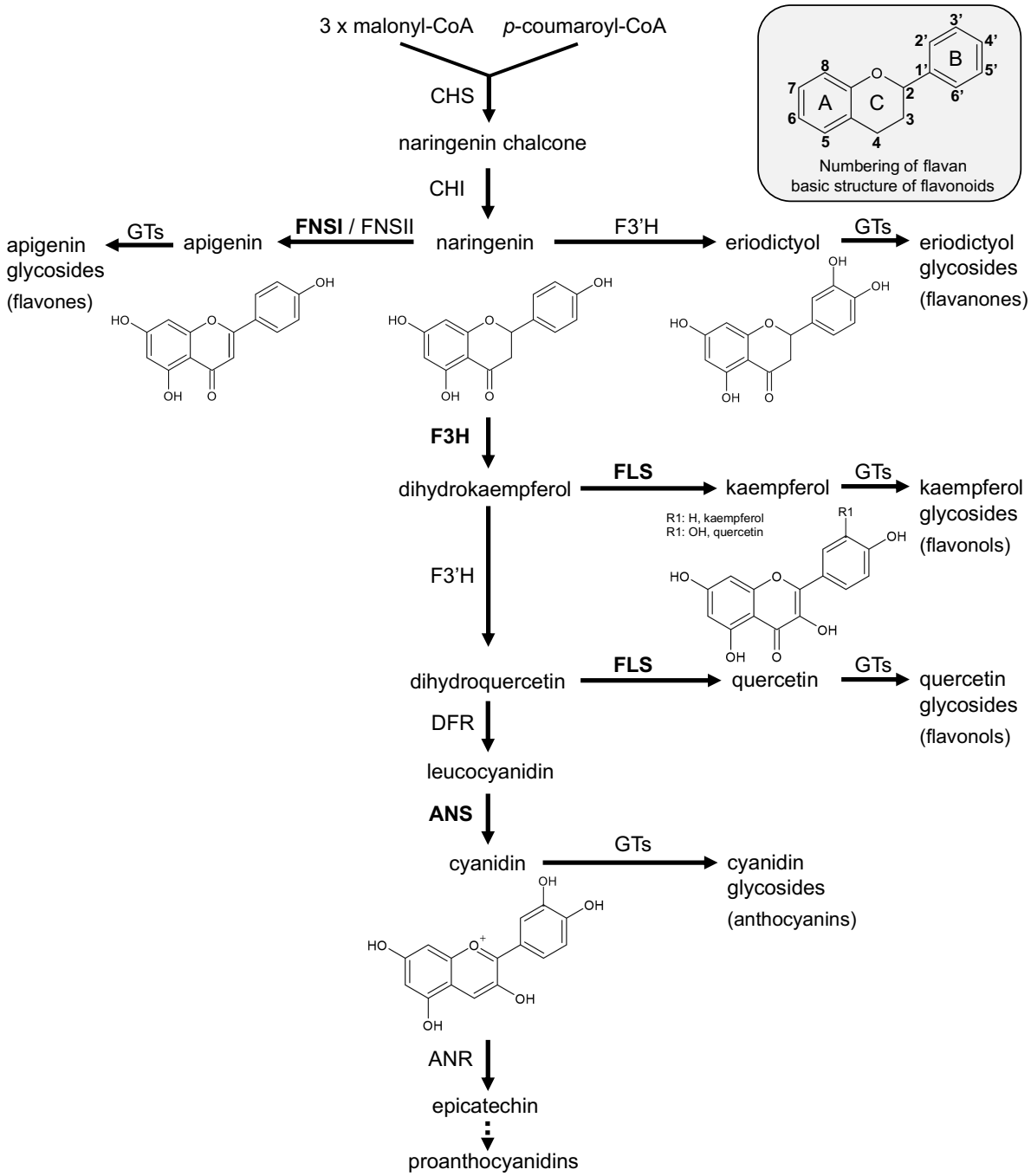

Figure 1: Simplified representation of the flavonoid biosynthesis pathway in plants. Details and abbreviations are given in the text. The 2-ODDs of the flavonoid biosynthesis F3H, FNSI, FLS, and ANS are shown in bold.

Characteristic hallmarks of 2-ODDs are generally (but not exclusively) the requirement of 2-oxoglutarate as co-reductant and iron as co-factor (36). Elucidation of the crystallographic structure of *Ath*ANS has identified the conserved amino acid residues involved in substrate-, co-factor-, and co-substrate binding (30). In addition, residues important for proper folding and substrate binding have been described (30).

117 As F3H, FLS and ANS all carry the conserved amino acid residues essential for 2-ODD activity,
118 it has been proposed that they share a common catalytic mechanism, possibly allowing
119 multifunctionality (40, 41). While FNSI and F3H often shows a narrow substrate spectrum (40,
120 42), FLS and ANS have been characterised as bi- or even multifunctional enzymes in several
121 species (40, 43-46). From an evolutionary perspective, it seems feasible that the 2-ODD
122 enzyme proposed to have evolved first, F3H, may have given rise to the other 2-ODDs involved
123 in flavonoid biosynthesis (summarised in 47) and that *Ath*FLS1 and *Ath*ANS retain *in vitro* F3H
124 (40, 48) and, in the case of *Ath*ANS, additional FLS side activity *in planta* (49). Due to the
125 multifunctional nature of flavonoid 2-ODD enzymes, comprehensive biochemical and
126 functional characterisation is required for unequivocal annotation, especially *in planta*. Only
127 seven amino acid residues distinguish parsley F3H from FNSI, providing an impressive
128 example that amino acid alignments are not sufficient for full functional characterisation (24).
129 The multifunctionalities of flavonoid 2-ODDs can be seen in the "leaky" phenotypes of *f3h*, *fls1*
130 and *ans* mutants of *A. thaliana* (21, 27, 49, 50), attributed to the partial compensation by other
131 2-ODDs. The *ans* mutant shows a significant reduction in anthocyanin and proanthocyanidin
132 content with almost unchanged flavonol levels compared to the wildtype (49, 51, 52), while the
133 *fls1* mutant shows a significant reduction in flavonol content and increased anthocyanin and
134 proanthocyanidin levels (49, 50, 53). The *f3h* mutant has a reduced flavonoid content, including
135 flavonols, anthocyanins, and proanthocyanidins, due to a generally reduced flux in the
136 flavonoid pathway (21, 54).
137 To unambiguously annotate the *in planta* multifunctionalities of F3H, FLS and ANS (which does
138 not necessarily reflect the *in vitro* function), multiple mutants are required, like it was done in
139 the case of the *fls1*/*ans* double mutant, which demonstrated the *in planta* FLS side activity of
140 ANS (40, 49).
141 In this study, we generated novel *A. thaliana f3h*/*fls1* and *f3h*/*ans* double mutants, as well as
142 an *f3h*/*fls1*/*ans* triple mutant to examine the phenotypic changes resulting from the loss of
143 activity of these flavonoid 2-ODDs. The *f3h*/*fls1*/*ans* mutant was characterised in detail, using
144 genomic, transcriptomic and metabolomic methods. Comprehensive phenotyping of the

145 flavonoid composition of the established flavonoid 2-ODD mutant set (composed of single,
146 double, and triple mutants) provided new insights into the *in planta* multifunctionalities of
147 *A. thaliana* F3H, FLS1 and ANS and the resulting metabolic fluxes. We found a flux into a
148 family of eriodictyol derivatives, which is most prominent in the *f3h*/*fls1*/*ans* triple mutant. The
149 knowledge gained in this study will be valuable for targeted engineering of flavonoid
150 composition and content in plants, especially crops.
151

152 # Results

153 ## Genomic characterisation of the *A. thaliana f3h*/*fls1*/*ans* mutant

154 To generate an *A. thaliana f3h*/*fls1*/*ans* mutant set, we used previously published insertional
155 alleles of the respective genes, showing the known loss-of-function phenotypes (*f3h* (*tt6-2*,
156 GK-292E08, Col-0 background), *fls1* (*fls1-2,* RIKEN_PST16145, Nö-0 background), *ans*
157 (*tds4-4*, SALK_073183, Col-0 background)). An *fls1*/*ans* double mutant was crossed with the
158 *f3h* single mutant. By screening the F2 progeny, homozygous *f3h*/*fls1*/*ans*, *f3h*/*fls1* and *f3h*/*ans*
159 mutants were identified by PCR-based genotyping. A *de novo* long-read assembly of the
160 *f3h*/*fls1*/*ans* mutant genome was generated to analyse the respective mutated loci at the
161 genomic level. The 160.57 Mbp assembly comprises 162 contigs with an N50 of 8.17 Mbp
162 (Supplementary Table S1). The assembly size slightly exceeds the size of the Col-0
163 Community-Consensus reference genome sequence assembly (133,336,906 bp) likely due to
164 the increased heterozygosity derived from the Nö-0 background of the *fls1* mutant. In total,
165 95.5 % complete BUSCOs were identified, indicating a high level of assembly completeness.
166 The assembly was used to validate the position of the T-DNA insertions at the *f3h* and *ans*
167 loci, and the position of the immobilised Ds transposon at the *fls1* locus (Figure 2). Two T-DNAs
168 comprising 11,393 bp sequence are inserted head-to-head into the 5'UTR of *f3h*, precisely
169 between 19,025,263 - 19,025,278 bp on chromosome three. The 5,417 bp Ds transposon is
170 inserted between 2,804,241 - 2,804,242 bp on chromosome five into the *fls1* locus in the first
171 exon. A large T-DNA array comprised of several T-DNAs and the corresponding binary vector

backbone (BVB) sequences was detected between 12,005,371 - 12,005,393 bp on chromosome four, inserted into the front part of the second exon of *ans*. Due to the high amount of repetitive sequences derived from this complex T-DNA array, the array is split across two contigs in the assembly and therefore the exact size (spans >119 kbp) could not be determined. A T-DNA left border (LB) was located at both ends of this insertion. Two additional T-DNA insertion events have been identified (Supplementary Figure S1). However, these insertions do not disrupt any additional annotated genes.

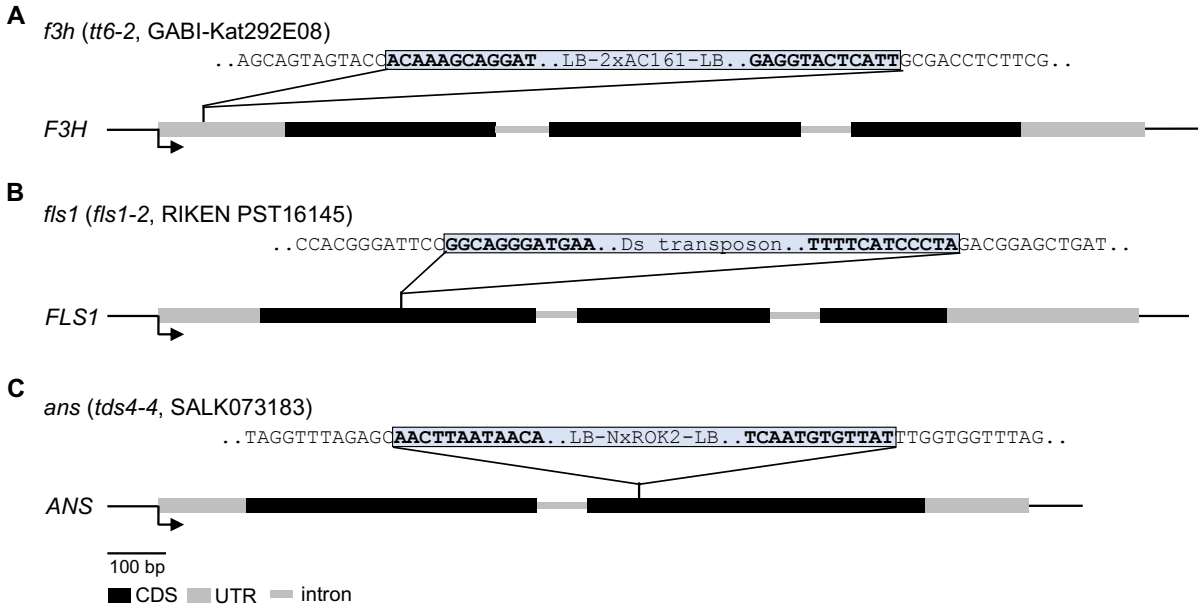

Figure 2: Genomic characterisation of the *f3h*/*fls1*/*ans* mutant. Schematic drawing of the (A) *F3H* (*AT3G51240*), (B) *FLS1* (*AT5G08640*) and (C) *ANS* (*AT4G22880*) genes illustrating the mutations in *f3h* (*tt6-2*), *fls1* (*fls1-2*) and *ans* (*tds4-4*), respectively. Black boxes mark the CDSs and grey boxes indicate UTRs and grey thin lines indicate introns, bent arrows indicate the transcription start. Dashed lines mark the respective T-DNA or transposon insertion. Sequences belonging to the T-DNA or the transposon insertion are shown in a blue box in bold and are flanked by the corresponding genomic sequences. LB, left border.

Transcriptome analysis of the *f3h*/*fls1*/*ans* mutant

To validate the insertion mutant alleles at the transcriptional level and to examine the transcriptional changes that occur in the *f3h*/*fls1*/*ans* mutant, RNA-Seq data from six-days-old seedlings were analysed. Col-0 and Nö-0 seedlings were used as controls. First, the expression profiles of the respective knockout alleles were analysed in the *f3h*/*fls1*/*ans* mutant (Supplementary Figure S2). For the *ans* and *fls1* mutant alleles, only marginal read coverage

195 was observed in the upstream regions of the T-DNA and transposon insertion positions,
196 respectively. These transcripts, which are greatly truncated compared to their normal lengths,
197 most likely result in non-functional proteins. Surprisingly, the *f3h* mutant allele showed a similar
198 expression profile as the wildtype allele, indicating that the *F3H* locus is still transcribed in the
199 *f3h*/*fls1*/*ans* mutant. However, mapping of the *f3h*/*fls1*/*ans* RNA-Seq reads to the *f3h*/*fls1*/*ans*
200 genome assembly revealed chimeric RNA-Seq reads containing the T-DNA LB sequence and
201 the sequence of the 5'UTR of *f3h*, emphasising that part of the T-DNA sequence is also
202 transcribed (Supplementary Figure S2). Several alternative translation start codons were
203 identified upstream of the original start site of *f3h* (Supplementary Figure S3), which could be
204 used for translation from the mutant *f3h* primary transcript and ultimately lead to a frameshift
205 and/or premature stop codon, resulting in a non-functional protein (refer to the discussion for
206 details).
207 In total, 1,643 and 1,203 differentially expressed genes (DEGs) were identified when
208 comparing *f3h*/*fls1*/*ans* with Col-0 and Nö-0, respectively (Figure 3, Supplementary File S1).
209 Of the DEGs derived from the comparison between *f3h*/*fls1*/*ans* and Col-0, 768 were
210 up-regulated and 875 were down-regulated, while 818 genes were up-regulated and 385
211 genes were down-regulated when *f3h*/*fls1*/*ans* was compared to Nö-0. To investigate the
212 transcriptional changes occurring in the *f3h*/*fls1*/*ans* mutant, 86 genes which are differentially
213 expressed between *f3h*/*fls1*/*ans* and Col-0 and between *f3h*/*fls1*/*ans* and Nö-0, but not
214 additionally between Col-0 and Nö-0, were further analysed (Figure 3). After a manual
215 screening (see Methods for details), 35 DEGs remained, which are listed in Table 1. Of these,
216 28 were up-regulated in the *f3h*/*fls1*/*ans* mutant. The *FLS1* gene was the only known gene
217 related to flavonoid biosynthesis in the list of DEGs.

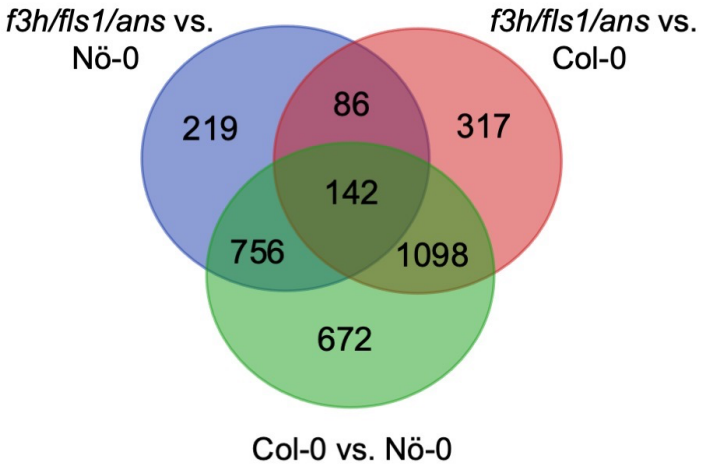

Figure 3: Venn diagram summarising the number of differentially expressed genes between seedlings of *f3h*/*fls1*/*ans* and the wildtypes Col-0 and Nö-0.

Table 1: List of differentially expressed genes in *f3h*/*fls1*/*ans* vs. Col-0 and Nö-0. Listed are a subset of differentially expressed genes, that were identified as DEGs in both, the *f3h*/*fls1*/*ans* vs. Col-0 and *f3h*/*fls1*/*ans* vs. Nö-0 comparisons, but notably, they were not differentially expressed between Col-0 and Nö-0. AGI, Arabidopsis gene identifier; l2fc, log_2 fold change; n.a., no annotation available.

AGI	l2fc vs. Col-0	l2fc vs. Nö-0	name / annotation
AT2G46765	7,25	5,74	hypothetical protein
AT5G08640	-5,54	-5,71	FLS1, flavonol synthase 1
AT2G20160	-4,71	-5,31	MEO, E3 ubiquitin ligase SCF complex subunit SKP1/ASK1 family protein
AT3G59884	3,35	1,84	MIR827A
AT2G14610	3,31	3,64	PR1 / CAPE9, pathogenesis-related protein 1
AT2G08820	2,3	1,37	n.a.
AT5G43570	2,22	1,38	PR-6 proteinase inhibitor family.
AT4G13493	2,21	1,64	MIR850A
AT4G06075	2,15	2,20	MIR5026 precursor
AT2G08825	2,11	1,40	n.a.
AT5G08855	2,11	1,96	n.a.
AT3G48560	2,08	2,14	AHAS / ALS, acetohydroxy acid synthase
AT3G46900	-1,98	-1,69	COPT2, Copper transporter 2, copper ion transmembrane transporter
AT1G52070	-1,93	-1,67	mannose-binding lectin superfamily protein
AT2G03850	1,93	1,07	LEA, late embryogenesis abundant protein family protein
AT2G11810	1,93	1,36	MGD3 / MGDC monogalactosyl diacylglycerol synthase 3
AT2G20723	1,84	1,79	snoRNA
AT2G20722	1,7	1,65	snoRNA

AT3G54730	1,62	3,16	transcription repressor
AT2G20670	1,52	1,62	DUF506, putative sugar phosphate exchanger
AT3G04210	1,36	1,59	TN13, disease resistance protein
AT2G44470	1,35	1,49	BGLU29, beta glucosidase 29
AT2G02850	1,34	1,16	ARPN, plantacyanin, copper ion binding
AT1G07135	1,29	1,23	glycine-rich protein
AT4G39675	1,28	1,32	hypothetical protein
AT1G60095	-1,27	-1,36	mannose-binding lectin superfamily protein
AT3G61930	1,26	1,17	hypothetical protein
AT2G07698	-1,23	-1,91	ATPase, F1 complex, alpha subunit protein
AT2G29330	1,23	1,54	TRI, tropinone reductase
AT1G36060	1,18	1,03	TG / WIND3, translucent green, Integrase-type DNA-binding superfamily protein
AT4G18980	1,18	1,58	AtS40-3
AT3G44260	1,17	1,13	CAF1A, Polynucleotidyl transferase, ribonuclease H-like superfamily protein
AT4G28420	1,14	1,62	tyrosine transaminase family protein
AT3G19430	-1,12	-1,90	LEA protein-like protein, late embryogenesis abundant protein-related
AT4G24570	1,01	1,10	DIC2, dicarboxylate carrier 2

228

229 **Phenotypic characterisation of the 2-ODD mutant collection including the novel**
230 ***f3h*/*fls1*, *f3h*/*ans*, and *f3h*/*fls1*/*ans* mutants**

231 To phenotypically characterise the flavonoid composition of the *f3h*/*fls1*/*ans* mutant compared
232 to the corresponding wildtypes, single and double mutants, the accumulation of flavanones,
233 flavonols, anthocyanins and proanthocyanidins (PA) was analysed (Figure 4, Figure 5). First,
234 the accumulation of flavonols was analysed in norflurazon-bleached seedlings, stained with
235 diphenylboric acid 2-aminoethyl ester (DPBA) (Figure 4A). Both wildtypes, Col-0 and Nö-0,
236 showed an orange colouration indicating the presence of quercetin glycosides. The *chs*
237 mutant, included as a control, is deficient in flavonol glycosides and thus appeared blue due
238 to sinapate derivatives. However, the *ans* mutant displayed an even more intense orange
239 fluorescence than the wildtypes. While the *f3h* mutant showed a red fluorescence, the *fls1*
240 mutant appeared green to light yellow. The *fls1*/*ans* mutant showed a pale yellow colour. The
241 novel *f3h*/*fls1*, *f3h*/*ans*, and *f3h*/*fls1*/*ans* mutants appeared red like the *f3h* mutant.
242 HPTLC and targeted UPLC-MS/MS analysis (including MRM quantification) of the flavonoid
243 composition in siliques revealed a wide range of flavonol glycosides in both wildtypes (Figure

244 4C, Supplementary Table S3). Consistent with DPBA staining, the *chs* mutant was deficient in
245 flavonoids. Accordingly, the level of *p*-coumaric acid was on average 186 µg per g dry weight
246 in *chs* siliques compared to up to 17 µg in all other lines, showing that this flavonoid precursor
247 accumulates in the *chs* mutant as the flavonoid biosynthesis pathway is blocked
248 (Supplementary Table S3). The *f3h*, *ans* and *f3h*/*ans* mutants showed a similar pattern of
249 flavonol glycosides as the wildtypes (Figure 4C). Moreover, the *ans* mutant accumulated the
250 highest amounts of flavonol glycosides with up to 3,549 µg quercetin-3-*O*-rhamnoside (Q-3R)
251 per g dry silique weight, whereas it was <273 µg in the other mutant lines and <2,656 µg in the
252 wildtypes (Supplementary Table S3). Accordingly, the *ans* knockout re-routes flavonoid
253 production to the flavonol biosynthesis branch while the other knockouts cause reduced
254 flavonol levels. The *fls1* mutant accumulated dihydroflavonols and small amounts of Q-3R,
255 whereas Q-3R levels decreased further to an average of 31.3 µg per g dry weight in *fls1*/*ans*
256 siliques while dihydroflavonol levels increased (Figure 4C and F, Supplementary Table S3).
257 The *fls1*/*ans* mutant accumulated the highest dihydroflavonol amounts compared to all lines
258 analysed, with an average dihydroquercetin level of 1,845 µg per g dry silique weight, showing
259 that dihydroflavonols accumulate when a further conversion to flavonols or anthocyanins is not
260 possible. Surprisingly, the *f3h*/*fls1* and *f3h*/*fls1*/*ans* mutants showed marginal and comparable
261 dihydroflavonols levels of up to 55 µg dihydrokaempferol per g dry weight, indicating the
262 presence of an (unknown) 2-ODD enzyme with F3H activity (Supplementary Table S3).
263 Interestingly, all mutant lines carrying the *f3h* allele accumulated a 'red (fluorescent)
264 metabolite', seen between Q-3R and K-3R-7R in HPTLC, indicating that a disruptive *f3h* allele
265 is essential for its production (Figure 4C). These results were in line with the red colouration of
266 *in situ* DPBA-stained seedlings carrying the *f3h* allele (Figure 4A). In contrast to flavonol
267 glycosides and dihydroflavonols, the accumulation of the 'red metabolite' was only visible late
268 after DPBA staining. The accumulation of the 'red metabolite' was further investigated in
269 DPBA-stained siliques under white and UV light (Supplementary Figure S4). The 'red
270 metabolite' visibly accumulated in immature seeds (Figure 4B) of all lines carrying the *f3h*
271 knockout allele and was also observed under white light. The *f3h*/*ans* and *f3h*/*fls1*/*ans* mutants

272 showed the strongest intensity and the ‘red metabolite’ appeared to accumulate during seed
273 maturation (Supplementary Figure S4). In addition, it was detected in young stamens of
274 mutants harbouring the *f3h* allele (Supplementary Figure S5). Interestingly, the eriodictyol
275 content of all analysed lines measured by UPLC-MS/MS were similar to the accumulation
276 pattern of the ‘red metabolite’ observed by HPTLC (Figure 4C and E). In detail, the average
277 level of eriodictyol was <1.4 µg per g dry silique weight in lines with a functional *F3H* allele,
278 whereas the average eriodictyol content of *f3h* knockout lines increased to 170 µg in *f3h*,
279 300 µg in *f3h*/*fls1,* 457 µg in *f3h*/*ans* and 442 µg in *f3h*/*fls1*/*ans*, indicating that eriodictyol
280 accumulates as a consequence of the *f3h* knockout (Figure 4E). To further characterise the
281 ‘red metabolite’, HPTLC was performed with available eriodictyol glycoside standards together
282 with methanolic extracts of Col-0 and *f3h*/*fls1*/*ans* siliques (Figure 4D). The ‘red metabolite’
283 had the same retention factor (R_F) as eriodictyol-7-*O*-glucoside, while the R_F values of the
284 eriodictyol disaccharides and the eriodictyol aglycone differed from that of the ‘red metabolite’.
285 Taken together, these results suggest that the metabolite underlying the red colouration of
286 mutants carrying the *f3h* allele is likely to be an eriodictyol hexoside, e.g.,
287 eriodictyol-7-*O*-glucoside.
288 Further chromatographic analysis with ion mobility separation in a high resolution mass
289 spectrometer (UPLC-TWIMS-QToF) on the *f3h*/*fls1*/*ans* flavonoid profile in siliques confirmed
290 the presence of eriodictyol-7-*O*-glucoside compared to a standard, the results show a mass
291 error below 3 ppm and an average of delta collision cross section (CCS) delta of 0.06 % in
292 ESI+ and 2.23 % in ESI- mode (Supplementary Table S4). The average content of
293 eriodictyol-7-*O*-glucoside determined by UPLC-MS/MS (high resolution mass spectrometry) in
294 the *f3h*/*fls1*/*ans* triple mutant is 10.38 µg per g dry weight (Supplementary Table S5).
295 Interestingly, the profile in *f3h*/*fls1*/*ans* is more complex (Supplementary Figures S6-S8). The
296 annotation of the eriodictyol derivates profile are based on $HDMS^E$ data, fragmentation
297 patterns by UPLC-MS/MS and isotopic distribution. The UPLC-TWIMS-QToF $HDMS^E$ profile
298 obtained in ESI+ and ESI- mode and subsequently UPLC-MS/MS analysis revealed the
299 presence of another eriodictyol hexoside, four di-hexosides, two tri-hexosides and two

300 rutinosides or neohesperidosides, respectively (Supplementary Tables S4 and S6). In ESI-
301 mode the main loss by UPLC-MS/MS is 162 Da (Supplementary Figure S8), corresponding to
302 the loss of a hexoside, the fragmentation has the presence of *m/z* 287, *m/z* 151 and *m/z*135
303 (Supplementary Figure S8), indicating a core with molecular formula $C_{15}H_{12}O_6$. Eriodictyol and
304 dihydrokaempferol have the same precursor, naringenin (Figure 1), and the same molecular
305 formula. Therefore, a comparison of fragmentation patterns, under the same methodology, of
306 eriodictyol and dihydrokaempferol standards by UPLC-MS/MS and UPLC-HDMS^E low and
307 high energy is required to improve the metabolite annotation. As is shown in Supplementary
308 Figures S9-S11, the absence of *m/z* 177 in all compounds of interest, including the
309 eriodictyol-7-*O*-glucoside, indicates an eriodictyol core. Two compounds, tentatively annotated
310 as eriodictyol rutinoside or neohesperidoside have two losses of 162 Da and 308 Da in
311 ESI- and 308 Da in ESI+ mode (Supplementary Figure S8, Supplementary Tables S4 and S6).
312 The presence of eriodictyol-7-*O*-rutinoside is excluded by the mismatch in retention time and
313 CCS values in the triple mutant by HDMS^e (Supplementary Table S4) and is in agreement with
314 the results shown in Figure 4D. Moreover, the dihydroquercetin profile of the *f3h*/*fls1*/*ans* triple
315 mutant compared to Col-0 is interesting, as it contains three different hexosides
316 (Supplementary Table S4). These results shed light on the accumulation of the eriodictyol and
317 dihydroquercetin aglycons and their subsequent modifications in the *f3h*/*fls1*/*ans* mutant.
318

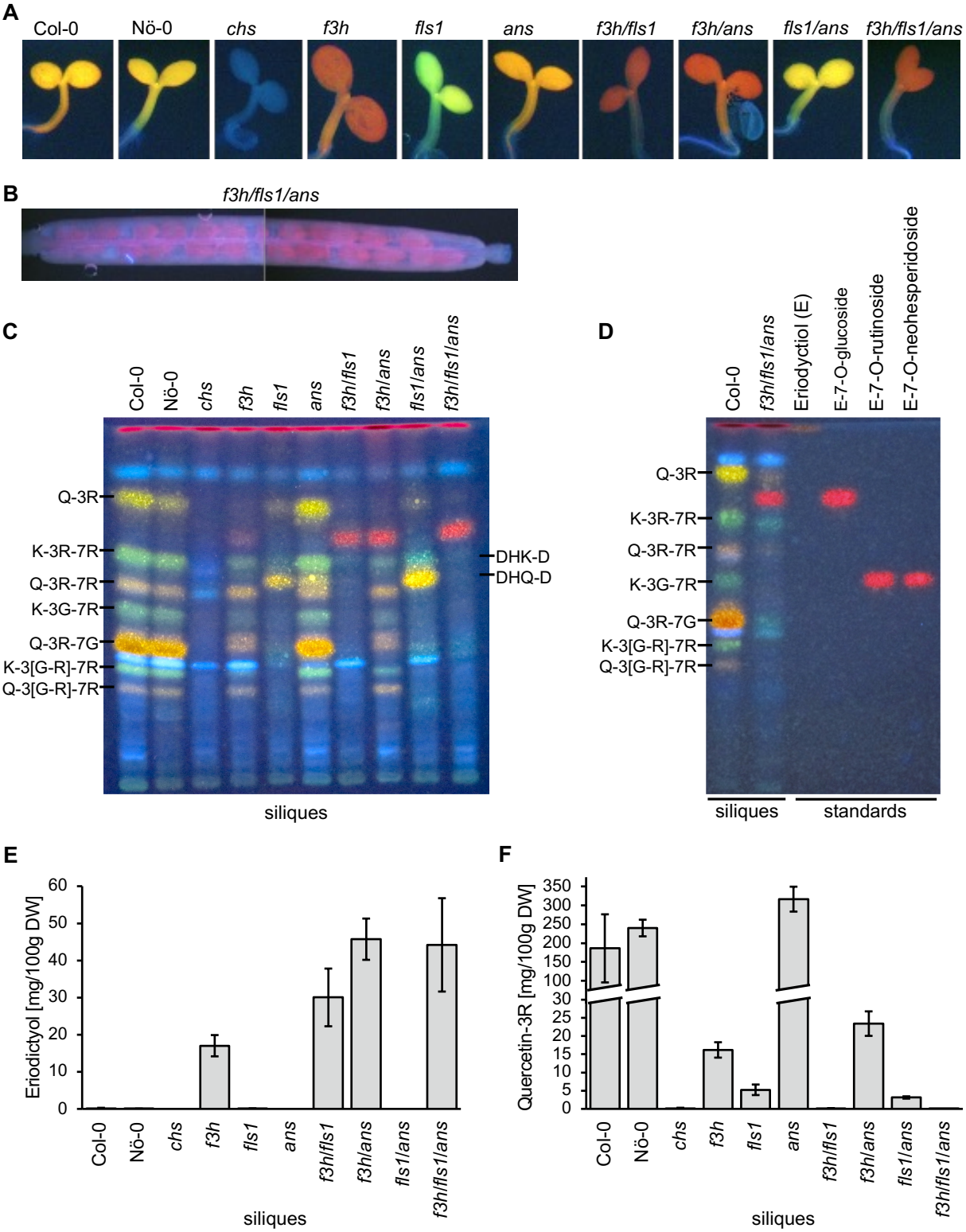

Figure 4: Accumulation of an eriodictyol hexoside in *A. thaliana* 2-ODD mutants carrying the *f3h* allele. (A) Seedlings stained *in situ* with diphenylboric acid 2-aminoethyl ester (DPBA) and visualised under UV light. Quercetin derivatives appear orange, kaempferol derivatives green, and blue corresponds to sinapate acid derivatives. (B) *In situ* DPBA-stained silique of the *f3h*/*fls1*/*ans* mutant visualised under UV light. (C) Accumulation of flavonol- and dihydroflavonol glycosides in siliques analysed by HPTLC and visualised under UV light after

326 treatment with DPBA. Kaempferol derivatives are green, quercetin derivatives are orange and
327 sinapate acid derivatives are blue. G, glucose; K, kaempferol; Q, quercetin; R, rhamnose;
328 DHK-D, dihydrokaempferol derivative; and DHQ-D, dihydroquercetin derivative. (D) Eriodictyol
329 standards analysed by HPTLC and visualised under UV light after treatment with DPBA. E,
330 eriodictyol. (E) Eriodictyol and (F) quercetin-3-*O*-rhamnoside contents in siliques analysed by
331 UPLC-MS/MS MRM. Error bars indicate the standard deviation of four independent
332 measurements.

333

334 Besides flavanones and flavonols, other downstream (end-)products of the flavonoid
335 biosynthesis were analysed including anthocyanins and PAs. The accumulation of
336 anthocyanins was analysed in sucrose-induced, norflurazon-bleached seedlings (Figure 5A).
337 While both wildtypes, *f3h*, *fls1*, and *f3h*/*fls1* mutants show anthocyanin accumulation, *chs*, *ans*,
338 *f3h*/*ans*, *fls1*/*ans*, and *f3h*/*fls1*/*ans* mutants do not. These results are consistent with the
339 analysis in sucrose-induced, native seedlings and were validated by photometric quantification
340 of anthocyanins in seedlings (Figure 5A and B).
341 Seed pigmentation is influenced by the levels of PA in the seed coat, with a darker seed colour
342 corresponding to higher PAs levels. Analysis of the native seed colour of the wildtypes revealed
343 characteristic dark brown colour due to the presence of large amounts of PAs, while the *chs*
344 mutant showed a pale yellow seed indicative of the well-known PA deficiency and transparent
345 testa phenotype (Figure 5A). Most of the *2-ODD* mutants showed an intermediate seed colour
346 phenotype. However, the *fls1* mutant showed a similar brown seed colour as the wildtypes. In
347 accordance, UPLC-MS/MS analysis revealed that epicatechin levels of Col-0, Nö-0, and *fls1*
348 averaged between 799 µg and 1,108 µg per g dry silique weight, while all other lines showed
349 average levels <7 µg (Supplementary Table S3). Similar observations were made for
350 procyanidins (Figure 5D, Supplementary Table S3). Thus, PAs and their precursor epicatechin
351 are only produced in considerable amounts when *CHS, F3H*, and *ANS* are intact.

352

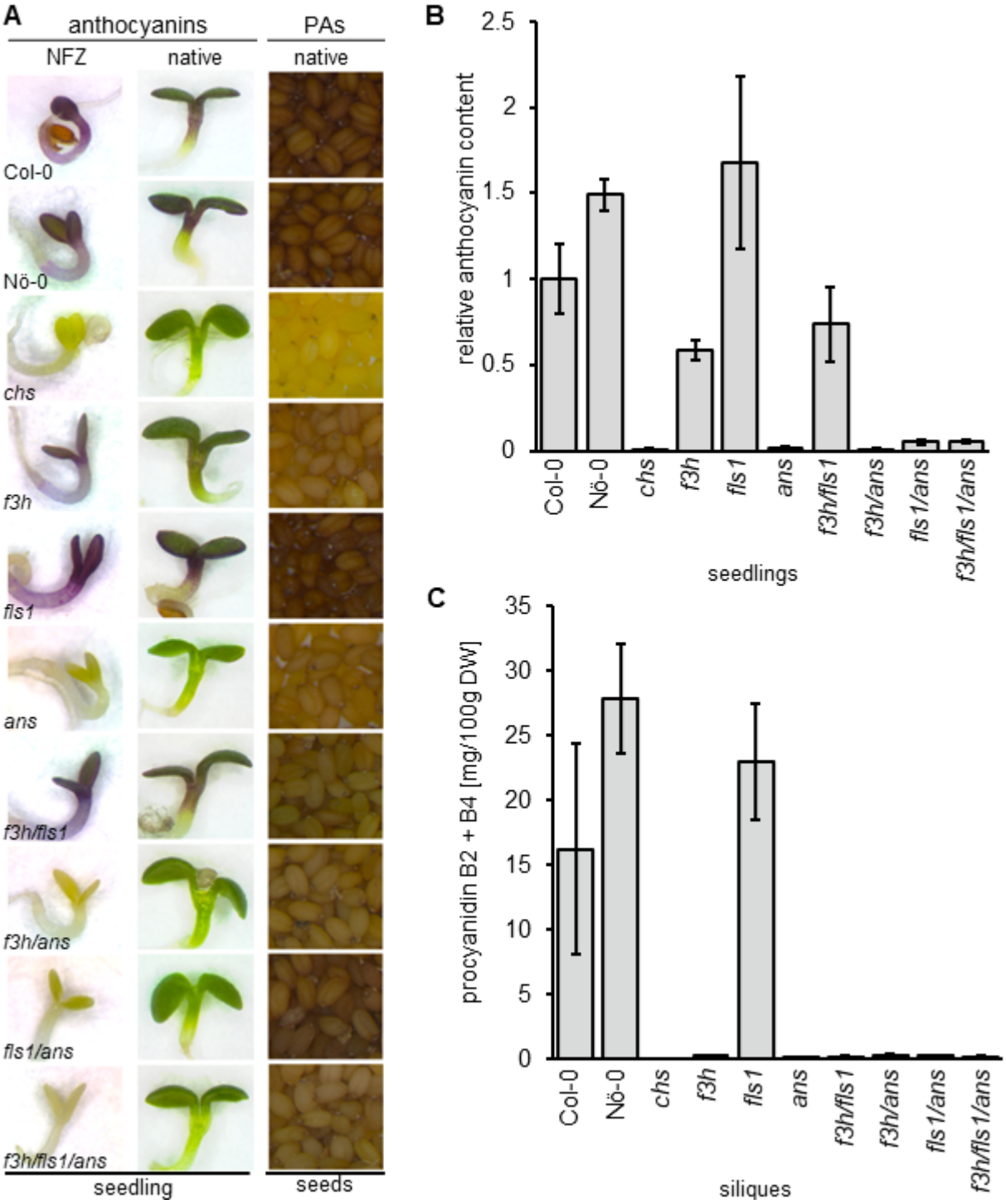

Figure 5: Anthocyanin and proanthocyanidin accumulation in *A. thaliana* 2-ODD mutants. (A) First column: norflurazon (NFZ)-bleached seedlings grown on 4% sucrose for anthocyanin induction. Second column: anthocyanin accumulation in seedlings grown on 0.5 MS plates with 4% sucrose. Third column: native seed colour. (B) Sucrose-induced anthocyanin accumulation in 6-days-old *A. thaliana* seedlings. Three independent biological replicates were measured for each sample. Error bars indicate standard deviation of the mean anthocyanin levels. (C) Procyanidin B2 + B4 content in siliques analysed by UPLC-MS/MS. Error bars indicate the standard deviation of four independent measurements.

363 # Discussion

364 The three major flavonoid 2-ODD enzymes, F3H, FLS and ANS, are closely related and
365 therefore often have overlapping catalytic activities. We have generated an *A. thaliana*
366 *f3h*/*fls1*/*ans* mutant and analysed genomic, transcriptomic and phenotypic changes in order to
367 increase the knowledge about flavonoid biosynthesis related 2-ODDs in *A. thaliana.*

368 ## Genome sequence of the *f3h*/*fls1*/*ans* triple mutant

369 To characterise the *f3h*/*fls1*/*ans* triple mutant at the genomic level, a *de novo* Oxford Nanopore
370 Technology (ONT) genome assembly was generated. In-depth analysis of the respective
371 *2-ODD* mutant loci revealed the expected insertion positions per mutant allele (49, 51, 55). To
372 our knowledge, these mutant alleles have never been characterised using long-read
373 technology. Thus, in this study, the insertion of repeated T-DNAs and their BVB sequences at
374 the *ans* (*tds4-4*) locus were identified. Moreover, the ONT assembly revealed that two T-DNAs
375 inserted head-to-head in the 5'UTR of *f3h*. Such complex T-DNA insertion events, involving
376 arrays of multiple T-DNAs as well as BVB sequences, have been reported before to be more
377 the rule than the exception for *A. thaliana* T-DNA insertion mutants (56-59).

378 ## Transcriptome analysis of the *f3h*/*fls1*/*ans* triple mutant

379 RNA-Seq analysis in seedlings revealed that *FLS1* is the only differentially expressed 2-ODD
380 between the *f3h*/*fls1*/*ans* mutant and the wildtypes and not additionally differentially expressed
381 between the wildtypes Col-0 and Nö-0. *ANS* is downregulated in the triple mutant compared
382 to the two wildtypes, but additionally also downregulated in Nö-0 compared to Col-0. *F3H* was
383 not identified as differentially expressed gene (DEGs) due to remaining RNA-Seq read
384 coverage at the *f3h* loci downstream of the T-DNA insertion site (Supplementary Figure S2).
385 For *ANS*, which comprises two exons, read coverage of the most of the first exon was observed
386 upstream of the complex T-DNA insertion array. The resulting transcript is severely truncated
387 and, as previously confirmed by the phenotype of the homozygous mutant (51, 52), results in
388 a non-functional protein.

389

390 Interestingly, the expression level of the *f3h* locus was similar to that of the wildtype alleles and
391 covered the entire *F3H* exons. Comparison of the RNA-Seq data to the *f3h*/*fls1*/*ans* genome
392 assembly identified several chimeric RNA-Seq reads spanning part of the T-DNA and the
393 5'UTR of *f3h*, emphasising that part of the T-DNA::genome junction is transcribed.
394 Transcription is likely driven by the 2'promoter of the T-DNA, located 723 bp upstream of the
395 5'UTR of the *f3h* locus (Supplementary Figure S3). A similar observation was made in a
396 *phytochrome C* mutant by *Monte et al.* (60). They analysed a T-DNA insertion line that showed
397 *PHYC* transcript levels similar to the wildtype. However, protein analysis revealed, that the
398 plants did not accumulate the corresponding protein. Similar limitations are conceivable in our
399 study. The 11,393 bp T-DNA insertion could either lead to an extended and, in consequence,
400 non-functional protein or one of the start codons located on the T-DNA could be used as an
401 alternative start codon, ultimately resulting in a frameshift and/or premature stop codon
402 (Supplementary Figure S3). Taken together, the omics data and the very distinct phenotype of
403 the *f3h*/*fls1*/*ans* mutant strongly suggest a non-functional *f3h* allele, in line with previous results
404 (17, 55). Furthermore, genomic and transcriptomic data of the *ans* and *fls1* loci showed that
405 these loci are *null*-alleles.
406 Apart from the *FLS1* gene, no other genes encoding enzymes, regulators or transporters of
407 flavonoid biosynthesis were identified as DEGs between the *f3h*/*fls1*/*ans* mutant and the
408 wildtypes. Nevertheless, several identified DEGs are noteworthy. For example, *miR827A* is
409 upregulated in *f3h*/*fls1*/*ans*, compared to the wildtypes. *miR827* expression is known to be
410 sucrose-induced and it targets ethylene response factors (ERFs) to regulate the expression of
411 *SUCROSE SYNTHASE* which has been proposed as a regulator of anthocyanin biosynthetic
412 genes (61). Accordingly, increased *miR827A* expression could be a means of the *f3h*/*fls1*/*ans*
413 mutant to counteract anthocyanin deficiency. Of course, this requires an unknown feedback
414 regulation circuit. *TROPINONE REDUCTASE* (*TRI*) transcript abundance is increased in
415 *f3h*/*fls1*/*ans* seedlings. Moreover, based on publicly available RNA-Seq datasets deposited in
416 of the European Bioinformatics Institute (EBI) Expression Atlas, *TRI* is expressed in Col-0
417 siliques (62, 63). TRI has previously been reported to reduce the flavanone backbone (64).

418 Since the flavonol- and anthocyanin branches of flavonoid biosynthesis are blocked in the
419 *f3h*/*fls1*/*ans* mutant, flavanones like eriodictyol accumulate in *f3h*/*fls1*/*ans* and could be
420 reduced by TRI activity. *BETA GLUCOSIDASE29* (*BGLU29*) is upregulated in *f3h*/*fls1*/*ans*
421 compared to the wildtypes. BGLU29 belongs to the family of beta-glucosidase enzymes (65).
422 Although some BGLUs have been shown to be functional glucosyltransferases involved in the
423 glycosylation of flavonoids (BGLU10 (66) and BGLU6 (67)), it is very unlikely that BGLU29 has
424 glycosyltransferase activity, as it is most closely related to BGLU28 - BGLU32, all of which
425 having glycoside hydrolase activity (68). BGLU29 may be involved in the hydrolysis of flavonoid
426 glycosides. In addition, the expression of *ACETOHYDROXY ACID SYNTHASE* (*AHAS*), which
427 is involved in the biosynthesis of valine, leucine and isoleucine from pyruvate (69), is increased
428 in the *f3h*/*fls1*/*ans* mutant. Since flavonoids are derived from phenylalanine, increased *AHAS*
429 expression could hint to an accumulation of phenylalanine in *f3h*/*fls1*/*ans* and an effort by the
430 plant to produce other amino acids instead.

431 **Phenotype**

432 ***In planta* multifunctionality of flavonoid 2-ODDs**

433 Previous studies revealed a “leaky” phenotype of the *2-ODD* single mutants and the *fls1*/*ans*
434 double mutant, which was attributed to partial compensation of other 2-ODD dioxygenases,
435 e.g., the bi- and multifunctionality of FLS and ANS, respectively (21, 27, 49). Owens *et al.* (27)
436 showed, that FLS1 not only accepts dihydroflavonols as substrate, but can also convert
437 naringenin to dihydrokaempferol *in planta*, thus exhibiting F3H side activity. ANS has been
438 shown to produce dihydroquercetin and quercetin from naringenin *in vitro*, indicating its F3H
439 and FLS side activity (40, 70, 71). Furthermore, the accumulation of flavonol glycosides in the
440 *fls1* mutant and further flavonol reduction in the *fls1*/*ans* mutant showed that ANS also has an
441 FLS activity *in planta* (49). This knowledge is consistent with our analyses and the novel
442 *f3h*/*ans*, *f3h*/*fls1* and *f3h*/*fls1*/*ans* mutants allowed an in-depth analysis of the *in planta*
443 multifunctionality of F3H, FLS, and ANS. The flavonol glycoside patterns of the *f3h*, *ans* and

444 the novel *f3h*/*ans* mutants were like the ones of the wildtypes confirming the significant *in*
445 *planta* F3H activity of FLS1.
446 The *fls1*/*ans* mutant still accumulates flavonol glycosides, suggesting that F3H has additional
447 *in planta* FLS activity. Alternatively, FLS3 may have residual *in planta* FLS activity, as FLS *in*
448 *vivo* activity of FLS3 has been shown in yeast (72). As all mutant lines with a non-functional
449 ANS contain no anthocyanins, F3H and FLS do not seem to have any ANS activity. Finally,
450 the marginal and comparable amounts of dihydroflavonol of up to 55 µg dihydrokaempferol per
451 g dry silique weight identified in the *f3h*/*fls1* and *f3h*/*fls1*/*ans* mutants indicate that ANS does
452 not appear to have significant F3H activity *in planta*. However, the identification of residual
453 dihydroflavonols in the *f3h*/*fls1*/*ans* mutant was rather surprising since the two major flavonoid
454 2-ODD responsible for dihydroflavonol accumulation, F3H and FLS1, were knocked-out in this
455 mutant.
456 Although the biological significance of these trace amounts of dihydroflavonols and the
457 enzyme(s) leading to their production is questionable, it raises the question of which enzyme(s)
458 could be involved in their biosynthesis. The dihydroflavonols could be produced by an
459 (unknown) 2-ODD with F3H activity or F3H side activities of one of the FLS paralogs. While
460 *AthFLS2*, *AthFLS4*, *AthFLS5* and *AthFLS6* were identified as pseudogenes or encoding
461 non-functional proteins, minor FLS activity of *Ath*FLS3 was reported under prolonged assay
462 conditions in yeast (27, 72). Although FLS3 showed no F3H or FLS activity in *Escherichia coli*,
463 *in planta* F3H activity was identified for FLS3 homologs from banana and rapeseed, the latter
464 sharing 67 % amino acid sequence identity with *Ath*FLS3 (42, 46). To investigate a potential
465 *in planta* F3H and FLS activity of *Ath*FLS3, an *fls3* knockout allele could be introduced in the
466 *f3h*/*fls1*/*ans* mutant to generate a *fls3*/*f3h*/*fls1*/*ans* mutant. A reduction of dihydroflavonols and
467 flavonols in the *fls3*/*f3h*/*fls1*/*ans* mutant compared to the *f3h*/*fls1*/*ans* mutant would provide
468 evidence for an F3H and FLS activity of FLS3 *in planta*.
469 Alternatively, the production of residual dihydroflavonols in the *f3h*/*fls1*/*ans* mutant may involve
470 FNSI, another flavonoid 2-ODD closely related to F3H. Bifunctional enzymes with FNSI and
471 F3H activity have been identified in gymnosperms and mosses (73). FNSI enzymes were

472 thought to be restricted to the Apiaceae family (74) and thus absent from *A. thaliana,* until the
473 identification of a DOWNY MILDEW RESISTANT6 (DMR6/FNSI) homolog with FNSI activity,
474 alongside reports of functional FNSIs from liverworts, *Oryza sativa*, and *Zea mays* (26, 73, 75,
475 76). However, *DMR6*/*FNSI* is primarily expressed in leaves and not in siliques (26), where the
476 residual amounts of dihydroflavonols were detected. Moreover, the endogenous activity of
477 DMR6/FNSI is not sufficient to compensate the dihydroflavonol deficient phenotype of the
478 *f3h*/*fls1* or *f3h*/*fls1*/*ans* mutants, indicating that DMR6/FNSI does not have substantial F3H
479 activity *in planta*. Taken together, these findings suggest that if DMR6/FNSI harbours F3H side
480 activity, it is likely to play a minor role, if any, in the production of the residual dihydroflavonols
481 detected in *f3h*/*fls1*/*ans* siliques.
482 In summary, the established 2-ODD mutant collection could be used in the future to analyse
483 not only potential multifunctionality of DMR6/FNSI, but rather 2-ODDs in general. This is of
484 particular importance for the FNSI family, as it is no longer confined to the Apiaceae family,
485 and knowledge on FNSI multifunctionality is rare, especially in angiosperms (73, 77). As
486 mentioned above, flavonoid 2-ODDs show broad substrate specificities *in vitro,* which do not
487 necessarily correspond to product accumulation *in vivo* (21). This emphasises the importance
488 of the established mutant collection for the *in planta* characterisation of flavonoid 2-ODD
489 multifunctionality, especially for studies characterising their functionalities solely *in vitro* (78,
490 79). For example, *in planta* complementation experiments using the novel mutant collection
491 could validate the FLS side activity found *in vitro* and the absence of F3H side activity of
492 *Vitis vinifera*'s ANS (79).

493 **Flavonoid accumulation in the absence of the three major flavonoid 2-ODDs**

494 The mutant collection was harnessed to investigate changes in the metabolic flux associated
495 with the knockout of the three major flavonoid 2-ODDs F3H, FLS1, and ANS. Interestingly, all
496 mutant lines carrying the *f3h* allele accumulate a metabolite in seedlings, immature seeds,
497 young stamens and siliques, which appears 'red' after DPBA-staining in visible light and under
498 UV light. This 'red (fluorescent) metabolite' is absent in the wildtypes and is most clearly seen
499 in an HPTLC of methanolic extracts from siliques (Figure 4B). The characteristic red

500 fluorescence of the *f3h* seedlings after DPBA staining has previously been observed and
501 suggested to be derived from an orthodihydroxy- or orthotrihydroxy flavonoid (80). In addition,
502 3'hydroxylation of the B ring was suggested to be required for the red fluorescence upon
503 staining, as DPBA-stained *f3h*/*f3'h* and *chs*/*f3h* mutants showed a light blue to pale
504 fluorescence (80). Owens *et al.* (21) suggested that F3'H might convert the accumulating
505 naringenin to eriodictyol which could be further reduced to a 3-desoxy flavonoid by the FNR
506 activity of native DFR (21). Wollenweber (81) postulated that the metabolite underlying the red
507 spot on DPBA-stained TLC plates is an eriodictyol derivative. We show that high levels of
508 eriodictyol accumulate in siliques of lines carrying an *f3h* knockout allele. This is in accordance
509 with the high expression of *F3'H* in siliques compared to flowers, leaves, stems, and roots (22).
510 In addition, we also observed that *F3'H* is expressed in *f3h*/*fls1*/*ans* seedlings (Supplementary
511 Figure S12), which showed the characteristic red colouration after *in situ* DPBA-staining. While
512 HPTLC-based flavonoid glycoside analysis showed that the 'red metabolite' is most likely an
513 eriodictyol hexoside, the results by UPLC-MS/MS and UPLC-TWIMS-QToF analyses revealed
514 a profile with various metabolites annotated as eriodictyol hexosides, including
515 eriodictyol-7-*O*-glucoside, and additional possible eriodictyols with neohesperidoside or
516 rutinoside derivatisation.
517 We therefore postulate that F3'H converts naringenin to eriodictyol, which may then be further
518 modified by glycosylation via glycosyltransferases. For example, the
519 flavonoid-7-*O*-glycosyltransferase UGT73B1 has been shown to use eriodictyol as a preferred
520 substrate for the catalysis to eriodictyol-7-*O*-glycoside (35). Considering the previously
521 presented DEGs, the tropinone reductase (TRI) may be involved in eriodictyol reduction, as a
522 previous study showed that TRI can reduce flavanones *in vitro* (82). In the future, knockout
523 alleles of the identified candidate genes could be introduced in the *f3h*/*fls1*/*ans* mutant to
524 analyse their role in the eriodictyol hexoside biosynthesis and degradation.
525
526

527 # Conclusions

528 We present a detailed characterisation of a novel *A. thaliana f3h*/*fls1*/*ans* mutant and the
529 corresponding single and double mutants. We have used genomic, transcriptomic, and
530 metabolite data to extend the knowledge about flavonoid 2-ODDs in *A. thaliana.* The set of *f3h,*
531 *fls1* and *ans* single, double and triple mutants will enable further work on flavonoid 2-ODDs
532 and their multifunctionalities in *A. thaliana*, as these are often limited to *in vitro* studies or
533 enzyme assays in *E. coli* or yeast (45, 78, 83). The *f3h*/*fls1*/*ans* mutant will also facilitate the
534 characterisation of multifunctional 2-ODDs from other species *in planta*. Finally, we have
535 shown that eriodictyol and several eriodictyol derivatives accumulate in the absence of a
536 functional F3H *in planta* with eriodictyol-7-*O*-glucoside being one of the predominant eriodictyol
537 derivatives. (Homo-)eriodictyols are bitter-masking flavanones with no additional strong taste
538 or flavour and are thus promising flavour modifiers for food applications and pharmaceuticals
539 (84).
540

541 # Methods

542 ## Plant material

543 Columbia-0 (Col-0, NASC ID N1092) and Nössen-0 (Nö-0, NASC ID N3081) accessions were
544 used as wildtype controls. A *chs* mutant (*tt4-11*, SALK_020583, Col-0 background, T-DNA
545 from pROK2 plasmid (18, 85)) was used as a flavonoid deficient control. The *A. thaliana*
546 mutants *f3h* (*tt6-2*, GK-292E08, Col-0 background, T-DNA from pAC161 plasmid (17)), *fls1*
547 (*fls1-2,* RIKEN_PST16145, Nö-0 background, Ds transposon (49,16)), *ans* (*tds4-4*,
548 SALK_073183, Col-0 background (51, 52)) and an *fls1*/*ans* double mutant (49) were used for
549 crosses and as controls.

550 ## Crossing and genotyping of mutants

551 To generate the novel *f3h*/*fls1*, *f3h*/*ans*, and *f3h*/*fls1*/*ans* mutants, the *fls1*/*ans* double mutant
552 was crossed with the *f3h* mutant. Individuals of the F2 progeny were analysed by PCR
553 genotyping using GoTaq2 DNA polymerase (Promega) according to standard protocols.

554 Sequences of primers used for genotyping are provided in Supplementary Table S2. Plants
555 homozygous for the respective mutant alleles were used for further analyses.

556 DNA extraction, library construction, and sequencing

557 ONT was used to generate long-reads of the *f3h*/*fls1*/*ans* mutant. Genomic DNA was isolated
558 from rosette leaves using the NucleoBond HMW DNA Kit (Macherey-Nagel, Düren, Germany)
559 according to the manufacturer's instructions. DNA quality was assessed via PicoGreen
560 (Quant-iT) measurement on a FLUOstar OPTIMA plate reader (BMG labtech), and pulsed-field
561 gel electrophoresis. The Short Read Eliminator (SRE) Kit (Circulomics, Baltimore, MD, USA)
562 was used to enrich long DNA fragments and results were validated by PicoGreen
563 measurements. 1.2 µg of high molecular weight DNA was used for library preparation
564 according to the SQK-LSK109 protocol (Oxford Nanopore Technologies (ONT), Oxford, UK).
565 Sequencing was performed on a R9.4.1 flow cell on a GridION. The flow cell was treated with
566 nuclease flush to load another library twice. Basecalling was performed using Guppy v.4.2.3
567 (https://community.nanoporetech.com).

568 *De novo* ONT genome assembly and read mapping

569 A *de novo* ONT genome assembly of the *f3h*/*fls1*/*ans* mutant was constructed with Canu v2.1.1
570 (86). No polishing of the assembly was performed, as the assembly was only used to analyse
571 the respective mutant loci. Contigs <100 kb were removed, followed by the calculation of
572 assembly statistics as previously described (87). Assembly quality was assessed via BUSCO
573 v3.0.2 using the embryophyta_odb9 lineage dataset (88). For manual curation of loci of interest
574 reads were mapped to the *A. thaliana* TAIR10 reference genome sequence (89) using
575 minimap2 (90). Minimap2 was run in ONT mode (-ax map-ont) and no secondary alignments
576 were allowed (--secondary=no). BAM files were sorted and indexed using samtools (91).
577 Manual curation of regions of interest was performed by using Integrative Genomics Viewer
578 (92) and BLAST (93). The assembly data was screened for the presence of additional T-DNAs
579 using loreta (57).

580 **RNA-Seq analysis: Ribonucleic acid extraction, library construction, and**
581 **sequencing**

582 RNA was isolated from 6-days-old seedlings using the NucleoSpin® RNA Plant Kit
583 (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. Four
584 biological replicates were used for each genotype (Col-0, Nö-0, and *f3h*/*fls1*/*ans*). The RNA
585 quality was validated using a 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA).
586 Sequencing libraries were constructed from 1 µg of total RNA, following the TrueSeq v2
587 protocol. Single end sequencing of 84 nt was performed on an Illumina NextSeq500 at the
588 Sequencing Core Facility of the Center for Biotechnology (CeBiTec) at Bielefeld University.

589 **RNA-Seq mapping and count table generation**

590 After assessing read quality using FastQC (94), reads were mapped to the *A. thaliana* Col-0
591 TAIR10 reference genome sequence (95) using STAR v.2.7.1a (96). STAR was run in Basic
592 mode allowing a maximum of 5 % mismatches per read length and using a minimum of 90 %
593 matches per read length. Mapping statistics were calculated based on STAR.log files using a
594 customised python script (available online via
595 https://github.com/hschilbert/3xODD/blob/main/parse_STAR_log_file_create_mapping_statis
596 tic.py). We used featureCounts v2.0.1 (97) with the -O --fraction parameters to generate count
597 tables based on the *A. thaliana* araport11 gff3 file (98) in order to assign a fractional count of
598 1/y from a reads, where y is the total number of meta-features/features overlapping with the
599 reads.

600 **Analysis of differentially expressed genes**

601 To identify differentially expressed genes (DEGs) between Col-0, Nö-0 and the *f3h*/*fls1*/*ans*
602 mutant in young seedlings, DESeq2 (99) was used to identify DEGs with a $\log_2$ fold change >1
603 and adjusted p-value <0.05 were selected for downstream analysis. Selected DEGs (86 DEGs
604 derived from the comparison of *f3h*/*fls1*/*ans* mutant to both wildtypes) were manually curated
605 by in-depth analysis of DEG's RNA-Seq read coverage against the TAIR10 reference genome
606 sequence to exclude DEGs arising from mapping artefacts, e.g., genes with partial read

607 coverage. The Venn diagram was generated using
608 https://bioinformatics.psb.ugent.be/webtools/Venn/.

609 **Flavonoid staining in *A. thaliana***

610 Visualisation of DPBA/Triton X-100-stained flavonoids in norflurazon-bleached *A. thaliana*
611 seedlings was performed according to Stracke *et al.* (100) using epifluorescence microscopy.
612 Visualisation of flavonoids in ethanol-bleached siliques of adult *A. thaliana* plants was
613 performed according to Stracke *et al.* (101).

614 **High performance thin layer chromatography (HPTLC) of flavonols and**
615 **dihydroflavonols**

616 Plants for flavonols extraction were grown in a greenhouse with 16 h of light illumination per
617 day at 22 °C. Extraction and analysis of dihydroflavonol and flavonol glycoside compounds
618 were performed as described in Stracke *et al.* (49). Siliques were homogenised in 80 %
619 methanol, incubated at 70 °C for 15 min and centrifuged at 14,000 x g for 10 min. After
620 vacuum-drying of the resulting supernatants, the dried pellets were redissolved in 80 %
621 methanol and analysed by HPTLC on a Silica Gel 60 plate (Merck, Darmstadt, Germany). The
622 mobile phase used was ethyl acetate, formic acid, acetic acid and water (100:26:12:12).
623 Dihydroflavonols and flavonol glycosides were detected by DPBA staining and UV light.

624 **(Poly)phenol extraction and targeted Ultra Performance Liquid Chromatography**
625 **– Multiple Reaction Monitoring Mass Spectrometry (UPLC-MS/MS MRM) analysis**

626 Samples of plant material were taken directly in the laboratory, lyophilized, and stored at room
627 temperature until used for further analysis. 100 mg of ground tissue were mixed with 1 mL of
628 80 % methanol for extraction of (poly)phenolic compounds. Samples and solvent were mixed
629 on an orbital shaker for 1 h at room temperature, sonicated for 30 min and the extraction
630 continued overnight at 4 °C in the dark. Each extract was filtered through a MILLEX 13-0.22 µm
631 PTFE membrane into glass vials and analysed as follows.
632 Sample extracts were injected immediately after filtration. Targeted UPLC-MS/MS with MRM
633 quantification analysis was performed on a Waters Acquity UPLC system (Milford, MA, USA)

634 consisting of a binary pump, an online vacuum degasser, an autosampler and a column
635 compartment. Separation of the phenolic compounds was conducted on a Waters Acquity HSS
636 T3 column 1.8 µm, 100 mm × 2.1 mm, kept at 40 °C. The analysis of phenolic compounds was
637 performed according to the method described by Vrhovsek *et al.* (102) using water with 0.1 %
638 formic acid and acetonitrile with 0.1 % formic acid as mobile phases for the gradient.
639 Mass spectrometry detection was performed on a Waters Xevo TQMS instrument equipped
640 with an electrospray ionisation (ESI) source. The capillary voltage was 3.5 kV in positive mode
641 and −2.5 kV in negative mode; the source was maintained at 150 °C; the desolvation
642 temperature was 500 °C; cone gas flow was 50 L/h; and desolvation gas flow was 800 L/h.
643 Unit resolution was applied to each quadrupole. Prior to UPLC-MS/MS MRM analysis, infusion
644 of each individual standard was performed to optimize the MRM conditions for quantification.
645 Data processing was performed using the Mass Lynx TargetLynx Application Manager
646 (Waters).

647 **Ultra-Performance Liquid Chromatography - Traveling Wave Ion Mobility-Mass**
648 **Spectrometry - Quadrupole Time of Flight (UPLC-TWIMS-QToF) Mass**
649 **Spectrometry and tandem mass spectrometry (UPLC-MS/MS) analysis**

650 The metabolite profile of Col-0 and *f3h*/*fls1*/*ans* biological replicate sample extracts were
651 acquired in an UPLC Acquity I system (Waters Corporation) coupled to a high-resolution mass
652 spectrometer HDMS Synapt XS (Waters Corporation). The chromatographic methodology was
653 performed according to Arapitsas *et al.* (103) with some modifications. The chromatographic
654 separation was carried out in an Acquity Premier HSS T3 1.8 µm, 2.1 x 100 mm column
655 (Waters) at 40 °C. The mobile phases A and B consisted of water with 0.1 % of formic acid
656 and methanol with 0.1 % formic acid, respectively, pumped with a flow rate of 0.3 mL/min. High
657 definition MS^e ($HDMS^e$) data acquisition with a scan time of 0.2 s was performed in continuum
658 mode with a mass range of 50-2000 Da in resolution mode (≥ 21500 FWHM at *m/z* 785), for
659 both ionisation modes (ESI+ and ESI-). The source temperature was settled at 120 °C, the
660 desolvation temperature at 450 °C, the nebuliser gas volume flow was 50 L/h and the
661 desolvation gas volume flow rate was 800 L/h. For ESI- acquisition mode, the capillary voltage

662 was 2.0 kV, and sampling cone was 30 V, whereas for ESI+ mode, the capillary voltage was
663 3.0 kV, and the sampling cone was 40 V. A collision energy ramp of 20-40 V was applied in
664 the transfer cell for high energy. For the ion mobility separation, the helium cell flow rate was
665 180 mL/min and for nitrogen it was 90 mL/min. The wave velocity was set to 650 m/s and the
666 transfer wave velocity to 141 m/s. The calibration was performed using the Major Mix IMS/ToF
667 Calibration kit solution (Waters Corporation). A leucine enkephalin solution (200 pg/µL) was
668 continuously infused as lock mass at flow rate of 10 µL/min during the whole analysis.
669 MassLynx v.4.2 and UNIFI software (Waters Corporation) were used for data acquisition and
670 processing. The collision cross sections (CCS) obtained were compared with an *in-house*
671 standard library.
672 A target UPLC-MSMS analysis was performed with the same source conditions in resolution
673 mode in a range of 50-2000 Da, with a scan time of 1 s. The collision energy was applied as a
674 ramp in the trap from 10-50 V for ESI- and 10-40 V for ESI+.

675 **Determination of anthocyanin accumulation**

676 Anthocyanin accumulation was visualised in 6-days-old seedlings grown on filter paper soaked
677 with 3 ppm of the bleaching herbicide norflurazon supplemented with 4 % sucrose to induce
678 anthocyanin accumulation. In addition, anthocyanin content was quantified photometrically in
679 6-days-old seedlings grown on 0.5 MS plates with 4 % sucrose as described by Mehrtens *et*
680 *al.* (104). Three independent biological replicates were measured for each sample.
681

682 # Declarations

683 **Ethics approval and consent to participate**

684 Not applicable.

685 **Consent for publication**

686 Not applicable.

Availability of data and materials

All sequence data generated in this study are available under the ENA/NCBI Bioproject ID PRJEB60353. In detail, the RNA-Seq data of *A. thaliana* Col-0, Nö-0, and the *f3h*/*fls1*/*ans* mutant can be found with under the ENA/NCBI IDs: ERS14716218 - ERS14716221, ERS14716222 - ERS14716225 and ERS14716226 - ERS14716229, respectively. The genomic ONT data can be accessed using the ENA/NCBI IDs: ERS14716230 (FASTQ) and ERS14716231 (FAST5). The *A. thaliana f3h*/*fls1*/*ans* mutant ONT assembly can be accessed at the DOI: 10.4119/unibi/2979807.

Cometing interests

The authors declare that they have no competing interest.

Funding

This research was funded by the BMBF project RaPEQ, grant numbers "FKZ 031B0888A" (RaPEQ-2) and "FKZ 031B1305B" (RaPEQ-3), as well as by core funding of the Chair of Genetics and Genomics of Plants provided by Bielefeld University through Faculty of Biology. We acknowledge support for the publication costs by the Open Access Publication Fund of Bielefeld University and the Deutsche Forschungsgemeinschaft (DFG).

Authors‘ contributions

MB, HMS, RS and BW conceived and designed the research. MB, HMS, RS and SM carried out experiments and analysed the data. VS performed high resolution mass spectrometry experiments and interpreted the data. AA performed the metabolite quantification methodology. RS and BW supervised the project. MB and HMS drafted the manuscript. RS, SM and BW revised the manuscript.

Acknowledgments

We thank Andrea Voigt, Melanie Kuhlmann and Nicoletta Schmidt for excellent technical assistance and Martin Sagasser for fruitful discussions. Moreover, we thank Nils Kleinbölting

712 for his support regarding *loreta*. We thank the Center for Biotechnology (CeBiTec) at Bielefeld
713 University for providing an environment to perform the computational analyses.
714

715 **Supplementary Material**

716 Supplementary Table S1: *A. thaliana f3h/fls1/ans* mutant *de novo* ONT genome assembly
717 statistics.
718
719 Supplementary Table S2: Oligonucleotide primers used in this work.
720
721 Supplementary Table S3: Polyphenol quantification in siliques of *A. thaliana* wildtypes and
722 2-ODD mutants.
723
724 Supplementary Table S4: Annotation of the eriodictyol and dihydroquercetin profile of siliques
725 of *A. thaliana* wildtypes and *f3h/fls1/ans* mutant.
726
727 Supplementary Table S5: Polyphenol quantification in siliques of *A. thaliana* wildtype and
728 *f3h/fls1/ans* mutant.
729
730 Supplementary Table S6: Dihydroquercetin and eriodictyol profile of siliques of *A. thaliana*
731 *f3h/fls1/ans* mutant.
732
733 Supplementary File S1: Differentially expressed genes between seedlings of different
734 genotypes.
735
736 Supplementary Figure S1: Additional T-DNA insertions in the *f3h/fls1/ans* mutant identified by
737 loreta.
738
739 Supplementary Figure S2: Genomic and transcriptomic characterisation of the mutant loci of
740 the *A. thaliana f3h/fls1/ans* mutant.
741
742 Supplementary Figure S3: The *f3h* locus of the *A. thaliana f3h/fls1/ans* mutant.
743
744 Supplementary Figure S4: *In situ* flavonoid staining of siliques from *A. thaliana* wildtype
745 and 2-ODD mutant plants.
746
747 Supplementary Figure S5: *In situ* flavonoid staining of inflorecences from *A. thaliana* wildtype
748 and 2-ODD mutant plants.
749
750 Supplementary Figure S6: Comparison of the eriodictyol profile between the *A. thaliana*
751 *f3h/fls1/ans* mutant and Col-0 in UPLC-HDMSe ESI+.
752
753 Supplementary Figure S7: Comparison of the eriodictyol profile between the *A. thaliana*
754 *f3h/fls1/ans* mutant and Col-0 in UPLC-HDMSe ESI-.
755
756 Supplementary Figure S8: UPLC-MS/MS profile of one biological replicate from the *A. thaliana*
757 *f3h/fls1/ans* mutant in ESI- mode.
758

759 Supplementary Figure S9: ESI- fragmentation of dihydrokaempferol standard by UPLC-
760 MS/MS and UPLC-TWIM-QToF (HDMSe) high resolution mass spectrometry.
761
762 Supplementary Figure S10: Fragmentation of eriodictyol standard by UPLC-MS/MS and
763 UPLC-TWIM-QToF (HDMSe) high resolution mass spectrometry.
764
765 Supplementary Figure S11: ESI- UPLC-TWIM-QToF (HDMSe) high resolution mass
766 spectrometry of compound 1 and 8 (Supplementary Table S4) in the *f3h/fls1/ans* triple mutant
767 biological replicate.
768
769 Supplementary Figure S12: The *F3'H* locus of the *A. thaliana f3h/fls1/ans* mutant.
770

771 # References

772

773 1. Wen W, Alseekh S, Fernie AR. Conservation and diversification of flavonoid
774 metabolism in the plant kingdom. Current Opinion in Plant Biology. 2020;55:100-8.

775 2. Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry,
776 cell biology, and biotechnology. Plant Physiology. 2001;126:485-93.

777 3. Ishikura N, Yoshitama K. Anthocyanin-Flavonol Co-pigmentation in Blue Seed Coats
778 of Ophiopogon jaburan. Journal of Plant Physiology. 1984;115(2):171-5.

779 4. Gronquist M, Bezzerides A, Attygalle A, Meinwald J, Eisner M, Eisner T. Attractive and
780 defensive functions of the ultraviolet pigments of a flower *(Hypericum calycinum)*. Proceedings
781 of the National Academie of Sciences of the United States of America. 2001;98(24):13745-50.

782 5. Grotewold E. The Science of Flavonoids. Grotewold E, editor. Columbus, Ohio, USA:
783 Springer New York, NY; 2006. 274 p.

784 6. Peer WA, Murphy AS. Flavonoids and auxin transport: modulators or regulators?
785 Trends in Plant Science. 2007;12(12):556-63.

786 7. Mo Y, Nagel C, Taylor LP. Biochemical complementation of chalcone synthase mutants
787 defines a role for flavonols in functional pollen. Proceedings of the National Academy of
788 Sciences of the United States of America. 1992;89:7213-7.

789 8. Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL. Arabidopsis Flavonoid Mutants Are
790 Hypersensitive to UV-B Irradiation. The Plant Cell. 1993;5(2):171-9.

791 9. Xie P, Shi JY, Tang SY, Chen CX, Khan A, Zhang FX, et al. Control of Bird Feeding
792 Behavior by Tannin1 through Modulating the Biosynthesis of Polyphenols and Fatty Acid-
793 Derived Volatiles in Sorghum. Molecular Plant. 2019;12(10):1315-24.

794 10. Dauer A, Hensel A, Lhoste E, Knasmüller S, Mersch-Sundermann V. Genotoxic and
795 antigenotoxic effects of catechin and tannins from the bark of Hamamelis virginiana L. in
796 metabolically competent, human hepatoma cells (Hep G2) using single cell gel
797 electrophoresis. Phytochemistry. 2003;63(2):199-207.

798 11. Kandil FE, Smith MA, Rogers RB, Pepin MF, Song LL, Pezzuto JM, et al. Composition
799 of a chemopreventive proanthocyanidin-rich fraction from cranberry fruits responsible for the

800 inhibition of 12-O-tetradecanoyl phorbol-13-acetate (TPA)-induced ornithine decarboxylase
801 (ODC) activity. Journal of Agricultural and Food Chemistry. 2002;50(5):1063-9.

802 12. Schijlen E, Ric de Vos CH, Jonker H, van den Broeck H, Molthoff J, van Tunen A, et al.
803 Pathway engineering for healthy phytochemicals leading to the production of novel flavonoids
804 in tomato fruit. Plant Biotechnology Journal. 2006;4(4):433-44.

805 13. Reddy AM, Reddy VS, Scheffler BE, Wienand U, Reddy AR. Novel transgenic rice
806 overexpressing anthocyanidin synthase accumulates a mixture of flavonoids leading to an
807 increased antioxidant potential. Metabolic Engineering. 2007;9(1):95-111.

808 14. Scharbert S, Hofmann T. Molecular definition of black tea taste by means of quantitative
809 studies, taste reconstitution, and omission experiments. Journal of Agricultural and Food
810 Chemistry. 2005;53(13):5377-84.

811 15. Koornneef M. Mutations affecting the testa colour in Arabidopsis. Arabidopsis
812 Information Services. 1990;27:1-4.

813 16. Ito T, Motohashi R, Kuromori T, Mizukado S, Sakurai T, Kanahara H, et al. A new
814 resource of locally transposed Dissociation elements for screening gene-knockout lines in
815 silico on the Arabidopsis genome. Plant Physiology. 2002;129(4):1695-9.

816 17. Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B. An *Arabidopsis thaliana*
817 T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics.
818 Plant Molecular Biology. 2003;53(1-2):247-59.

819 18. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, et al. Genome-wide
820 insertional mutagenesis of Arabidopsis thaliana. Science. 2003;301(5633):653-7.

821 19. Ferrer J-L, Jez JN, Bowman ME, Dixon RA, Noel JP. Structure of chalcone synthase
822 and the molecular basis of plant polyketide biosynthesis. Nature Structural Biology.
823 1999;6(8):775-84.

824 20. Jez JM, Bowman ME, Dixon RA, Noel JP. Structure and mechanism of the
825 evolutionarily unique plant enzyme chalcone isomerase. Nature Structural Biology.
826 2000;7(9):786-91.

827 21. Owens DK, Crosby KC, Runac J, Howard BA, Winkel BS. Biochemical and genetic
828 characterization of Arabidopsis flavanone 3beta-hydroxylase. Plant Physiology and
829 Biochemistry. 2008;46(10):833-43.

830 22. Schoenbohm C, Martens S, Eder C, Forkmann G, Weisshaar B. Identification of the
831 *Arabidopsis thaliana* flavonoid 3'-hydroxylase gene and functional expression of the encoded
832 P450 enzyme. Biological Chemistry. 2000;381:749-53.

833 23. Martens S, Forkmann G. Cloning and expression of flavone synthase II from Gerbera
834 hybrids. The Plant Journal. 1999;20(5):611-8.

835 24. Gebhardt YH, Witte S, Steuber H, Matern U, Martens S. Evolution of flavone synthase
836 I from parsley flavanone 3beta-hydroxylase by site-directed mutagenesis. Plant Physiology.
837 2007;144(3):1442-54.

838 25. Ali MB, McNear DHJ. Induced transcriptional profiling of phenylpropanoid pathway
839 genes increased flavonoid and lignin content in Arabidopsis leaves in response to microbial
840 products. BMC Plant Biology. 2014;14(84).

841 26. Falcone Ferreyra ML, Emiliani J, Rodriguez EJ, Campos-Bermudez VA, Grotewold E,
842 Casati P. The Identification of Maize and Arabidopsis Type I FLAVONE SYNTHASEs Links
843 Flavones with Hormones and Biotic Interactions. Plant Physiology. 2015;169(2):1090-107.

844 27. Owens DK, Alerding AB, Crosby KC, Bandara AB, Westwood JH, Winkel BS.
845 Functional analysis of a predicted flavonol synthase gene family in Arabidopsis. Plant
846 Physiology. 2008;147(3):1046-61.

847 28. Shirley BW, Hanley S, Goodman HM. Effects of ionizing radiation on a plant genome:
848 analysis of two Arabidopsis transparent testa mutations. The Plant Cell. 1992;4:333-47.

849 29. Davies KM, Schwinn KE, Deroles SC, Manson DG, Lewis DH, Bloor SJ, et al.
850 Enhancing anthocyanin production by altering competition for substrate between flavonol
851 synthase and dihydroflavonol 4-reductase. Euphytica. 2003;131(3):259-68.

852 30. Wilmouth RC, Turnbull JJ, Welford RW, Clifton IJ, Prescott AG, Schofield CJ. Structure
853 and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*. Structure.
854 2002;10(1):93-103.

855 31. Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. Role of anthocyanidin reductase,
856 encoded by BANYULS in plant flavonoid biosynthesis. Science. 2003;299(5605):396-9.

857 32. Zhang Y, Butelli E, Martin C. Engineering anthocyanin biosynthesis in plants. Current
858 Opnion in Plant Biology. 2014;19:81-90.

859 33. Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, et al. Functional
860 genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-
861 expressing an MYB transcription factor. The Plant Journal. 2005;42(2):218-35.

862 34. Jones P, Messner B, Nakajima J, Schaffner AR, Saito K. UGT73C6 and UGT78D1,
863 glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana*.
864 Journal of Biological Chemistry. 2003;278(45):43910-8.

865 35. Kim JH, Kim BG, Park Y, Ko JH, Lim CE, Lim J, et al. Characterization of flavonoid 7-
866 O-glucosyltransferase from Arabidopsis thaliana. Bioscience, Biotechnology, and
867 Biochemistry. 2006;70(6):1471-7.

868 36. Cheng AX, Han XJ, Wu YF, Lou HX. The function and catalysis of 2-oxoglutarate-
869 dependent oxygenases involved in plant flavonoid biosynthesis. International Journal of
870 Molecular Sciences. 2014;15(1):1080-95.

871 37. Kawai Y, Ono E, Mizutani M. Evolution and diversity of the 2-oxoglutarate-dependent
872 dioxygenase superfamily in plants. The Plant Journal. 2014;78(2):328-43.

873 38. Prescott AG, Lloyd MD. The iron(II) and 2-oxoacid-dependent dioxygenases and their
874 role in metabolism. Natural Product Reports. 2000;17(4):367-83.

875 39. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B. Oxidative
876 demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature.
877 2002;419(6903):174-8.

878 40. Turnbull JJ, Nakajima J, Welford RW, Yamazaki M, Saito K, Schofield CJ. Mechanistic
879 studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis:
880 anthocyanidin synthase, flavonol synthase, and flavanone 3-beta-hydroxylase. Journal of
881 Biological Chemistry. 2004;279(2):1206-16.

41. Martens S, Preuß A, Matern U. Multifunctional flavonoid dioxygenases: Flavonol and anthocyanin biosynthesis in Arabidopsis thaliana L. Phytochemistry. 2010;71(10):1040-9.

42. Busche M, Acatay C, Martens S, Weisshaar B, Stracke R. Functional Characterisation of Banana (Musa spp.) 2-Oxoglutarate-Dependent Dioxygenases Involved in Flavonoid Biosynthesis. Frontiers in Plant Science. 2021;12:701780.

43. Turnbull JT, Sobey WJ, Aplin RT, Hassan A, Firmin JL, Schofield CJ, et al. Are anthocyanidins the immediate products of anthocyanidin synthase? Chemical Communications. 2000:2473-4.

44. Yan Y, Chemler J, Huang L, Martens S, Koffas MA. Metabolic engineering of anthocyanin biosynthesis in Escherichia coli. Applied and Environmental Microbiology. 2005;71(7):3617-23.

45. Park S, Kim DH, Park BR, Lee JY, Lim SH. Molecular and Functional Characterization of Oryza sativa Flavonol Synthase (OsFLS), a Bifunctional Dioxygenase. Journal of Agricultural and Food Chemistry. 2019;67(26):7399-409.

46. Schilbert HM, Schöne M, Baier T, Busche M, Viehöver P, Weisshaar B, et al. Characterization of the *Brassica napus* flavonol synthase gene family reveals bifunctional flavonol synthases. Frontiers in Plant Science. 2021;12:733762.

47. Yonekura-Sakakibara K, Higashi Y, Nakabayashi R. The Origin and Evolution of Plant Flavonoid Metabolism. Frontiers in Plant Science. 2019;10:943.

48. Prescott AG, Stamford NPJ, Wheeler G, Firmin JL. In vitro properties of a recombinant flavonol synthase from *Arabidopsis thaliana*. Phytochemistry. 2002;60(6):589-93.

49. Stracke R, De Vos RCH, Bartelniewoehner L, Ishihara H, Sagasser M, Martens S, et al. Metabolomic and genetic analyses of flavonol synthesis in *Arabidopsis thaliana* support the *in vivo* involvement of leucoanthocyanidin dioxygenase. Planta. 2009;229(2):427-45.

50. Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, Hahlbrock K, et al. Knock-out mutants from an En-1 mutagenized *Arabidopsis thaliana* population generate new phenylpropanoid biosynthesis phenotypes. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(21):12432-7.

51. Abrahams S, Lee E, Walker AR, Tanner GJ, Larkin PJ, Ashton AR. The Arabidopsis TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. The Plant Journal. 2003;35(5):624-36.

52. Bowerman PA, Ramirez MV, Price MB, Helm RF, Winkel BS. Analysis of T-DNA alleles of flavonoid biosynthesis genes in Arabidopsis ecotype Columbia. BMC Research Notes. 2012;5.

53. Routaboul JM, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, Einhorn J, et al. Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. Planta. 2006;224(1):96-107

54. Pelletier MK, Burbulis IE, Shirley BW. Disruption of specific flavonoid genes enhances the accumulation of flavonoid enzymes and endproducts in Arabidopsis seedlings. Plant Molecular Biology. 1999;40(1):45-54.

922 55. Appelhagen I, Thiedig K, Nordholt N, Schmidt N, Huep G, Sagasser M, et al. Update
923 on *transparent testa* mutants from *Arabidopsis thaliana*: characterisation of new alleles from
924 an isogenic collection. Planta. 2014;240(5):955-70.

925 56. Jupe F, Rivkin AC, Michael TP, Zander M, Motley ST, Sandoval JP, et al. The complex
926 architecture and epigenomic impact of plant T-DNA insertions. PLoS Genetics.
927 2019;15(1):e1007819.

928 57. Pucker B, Kleinbolting N, Weisshaar B. Large scale genomic rearrangements in
929 selected Arabidopsis thaliana T-DNA lines are caused by T-DNA insertion mutagenesis. BMC
930 Genomics. 2021;22(1):599.

931 58. Rajapriya V, Kannan P, Sridevi G, Veluthambi K. A rare transgenic event of rice with
932 Agrobacterium binary vector backbone integration at the right T-DNA border junction. J Plant
933 Biochem Biot. 2021;30(3):421-8.

934 59. Wu H, Sparks CA, Jones HD. Characterisation of T-DNA loci and vector backbone
935 sequences in transgenic wheat produced by Agrobacterium-mediated transformation.
936 Molecular Breeding. 2006;18(3):195-208.

937 60. Monte E, Alonso JM, Ecker JR, Zhang YL, Li X, Young J, et al. Isolation and
938 characterization of phyC mutants in Arabidopsis reveals complex crosstalk between
939 phytochrome signaling pathways. The Plant Cell. 2003;15(9):1962-80.

940 61. Gao J, Peng H, Chen FB, Luo M, Li WB. Genome-wide analysis of transcription factors
941 related to anthocyanin biosynthesis in carmine radish (Raphanus sativus L.) fleshy roots.
942 PeerJ. 2019;7:e8041.

943 62. Liu J, Deng S, Wang H, Ye J, Wu HW, Sun HX, et al. CURLY LEAF Regulates Gene
944 Sets Coordinating Seed Size and Lipid Biosynthesis. Plant Physiology. 2016;171(1):424-36.

945 63. Mergner J, Frejno M, List M, Papacek M, Chen X, Chaudhary A, et al. Mass-
946 spectrometry-based draft of the Arabidopsis proteome. Nature. 2020;579(7799):409-14.

947 64. Reinhardt N. Untersuchung der katalytischen Eigenschaften Tropinonreduktase-
948 ähnlicher kurzkettiger Dehydrogenasen/Reduktasen aus Brassicaceen. Halle (Saale): Martin-
949 Luther-Universität Halle-Wittenberg; 2014.

950 65. Xu Z, Escamilla-Trevino L, Zeng L, Lalgondar M, Bevan D, Winkel B, et al. Functional
951 genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. Plant Molecular
952 Biology. 2004;55(3):343-67.

953 66. Ishihara H, Tohge T, Viehöver P, Fernie AR, Weisshaar B, Stracke R. Natural variation
954 in flavonol accumulation in Arabidopsis is determined by the flavonol glucosyltransferase
955 BGLU6. Journal of Experimental Botany. 2016;67(5):1505-17.

956 67. Miyahara T, Sakiyama R, Ozeki Y, Sasaki N. Acyl-glucose-dependent
957 glucosyltransferase catalyzes the final step of anthocyanin formation in Arabidopsis. Journal
958 of Plant Physiology. 2013;170(6):619-24.

959 68. Sugiyama R, Li R, Kuwahara A, Nakabayashi R, Sotta N, Mori T, et al. Retrograde
960 sulfur flow from glucosinolates to cysteine in Arabidopsis thaliana. Proceedings of the National
961 Academy of Sciences of the United States of America. 2021;118(22):e2017890118.

962 69. Duggleby RG, McCourt JA, Guddat LW. Structure and mechanism of inhibition of plant
963 acetohydroxyacid synthase. Plant Physiology and Biochemistry. 2008;46(3):309-24.

964 70. Turnbull JJ, Nagle MJ, Seibel JF, Welford RW, Grant GH, Schofield CJ. The C-4
965 stereochemistry of leucocyanidin substrates for anthocyanidin synthase affects product
966 selectivity. Bioorganic & Medicinal Chemistry Letters. 2003;13(21):3853-7.

967 71. Welford RWD, Turnbull JJ, Claridge TDW, Prescott AG, Schofield CJ. Evidence for
968 oxidation at C-3 of the flavonoid C-ring during anthocyanin biosynthesis. Chemical
969 Communications. 2001(18):1828-9.

970 72. Preuss A, Stracke R, Weisshaar B, Hillebrecht A, Matern U, Martens S. *Arabidopsis*
971 *thaliana* expresses a second functional flavonol synthase. FEBS Letters. 2009;583(12):1981–
972 6.

973 73. Li DD, Ni R, Wang PP, Zhang XS, Wang PY, Zhu TT, et al. Molecular Basis for
974 Chemical Evolution of Flavones to Flavonols and Anthocyanins in Land Plants. Plant
975 Physiology. 2020;184(4):1731-43.

976 74. Pucker B, Iorizzo M. Apiaceae FNS I originated from F3H through tandem gene
977 duplication. PLoS One. 2023;18(1):e0280155.

978 75. Han XJ, Wu YF, Gao S, Yu HN, Xu RX, Lou HX, et al. Functional characterization of a
979 *Plagiochasma appendiculatum* flavone synthase I showing flavanone 2-hydroxylase activity.
980 FEBS Letters. 2014;588(14):2307-14.

981 76. Lee YJ, Kim JH, Kim BG, Lim Y, Ahn JH. Characterization of flavone synthase I from
982 rice. BMB Reports. 2008;41(1):68-71.

983 77. Wang Y, Shi Y, Li K, Yang D, Liu N, Zhang L, et al. Roles of the 2-Oxoglutarate-
984 Dependent Dioxygenase Superfamily in the Flavonoid Pathway: A Review of the Functional
985 Diversity of F3H, FNS I, FLS, and LDOX/ANS. Molecules. 2021;26(21):6745.

986 78. Lukacin R, Wellmann F, Britsch L, Martens S, Matern U. Flavonol synthase from *Citrus*
987 *unshiu* is a bifunctional dioxygenase. Phytochemistry. 2003;62(3):287-92.

988 79. Zhang JR, Trossat-Magnin C, Bathany K, Negroni L, Delrot S, Chaudiere J. Oxidative
989 Transformation of Dihydroflavonols and Flavan-3-ols by Anthocyanidin Synthase from Vitis
990 vinifera. Molecules. 2022;27(3):1047.

991 80. Sheahan JJ, Cheong H, Rechnitz GA. The colorless flavonoids of *Arabidopsis thaliana*
992 (Brassicaceae). I. A model system to study the orthodihydroxy structure. American Journal of
993 Botany. 1998;85(4):467-75.

994 81. E. W. Unusual Flavanones from a Rare American Fern. Zeitschrift für Naturforschung
995 C. 1981;36:604-6.

996 82. Reinhardt N, Fischer J, Coppi R, Blum E, Brandt W, Drager B. Substrate flexibility and
997 reaction specificity of tropinone reductase-like short-chain dehydrogenases. Bioorganic
998 Chemistry. 2014;53:37-49.

999 83. Xu F, Li L, Zhang W, Cheng H, Sun N, Cheng S, et al. Isolation, characterization, and
1000 function analysis of a flavonol synthase gene from Ginkgo biloba. Molecular Biology Reports.
1001 2012;39(3):2285-96.

1002 84. Ley JP, Krammer G, Reinders G, Gatfield IL, Bertram HJ. Evaluation of bitter masking
1003 flavanones from Herba Santa (Eriodictyon californicum (H. and A.) Torr., Hydrophyllaceae).
1004 Journal of Agricultural and Food Chemistry. 2005;53(15):6061-6.

1005 85. Buer CS, Sukumar P, Muday GK. Ethylene modulates flavonoid accumulation and
1006 gravitropic responses in roots of Arabidopsis. Plant Physiology. 2006;140(4):1384-96.

1007 86. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable
1008 and accurate long-read assembly via adaptive k-mer weighting and repeat separation.
1009 Genome Research. 2017;27(5):722-36.

1010 87. Siadjeu C, Pucker B, Viehöver P, Albach DC, Weisshaar B. High Contiguity De Novo
1011 Genome Sequence Assembly of Trifoliate Yam (Dioscorea dumetorum) Using Long Read
1012 Sequencing. Genes. 2020;11(3):E274.

1013 88. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO:
1014 assessing genome assembly and annotation completeness with single-copy orthologs.
1015 Bioinformatics. 2015;31(19):3210-2.

1016 89. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The
1017 Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic
1018 Acids Research. 2012;40(Database issue):D1202-D10.

1019 90. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics.
1020 2018;34(18):3094-100.

1021 91. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
1022 Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

1023 92. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al.
1024 Integrative genomics viewer. Nature Biotechnology. 2011;29(1):24-6.

1025 93. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
1026 tool. Journal of Molecular Biology. 1990;215(3):403-10.

1027 94. Andrews S. FastQC, a quality control tool for high throughput sequence data. 2018
1028 [Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

1029 95. Lee H, Chawla HS, Obermeier C, Dreyer F, Abbadi A, Snowdon R. Chromosome-Scale
1030 Assembly of Winter Oilseed Rape *Brassica napus*. Frontiers in Plant Science. 2020;11:496.

1031 96. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast
1032 universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

1033 97. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for
1034 assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30.

1035 98. Cheng CY, Krishnakumar V, Chan A, Thibaud-Nissen F, Schobel S, Town CD.
1036 Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. The Plant
1037 Journal. 2017;89(4):789-804.

1038 99. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
1039 RNA-seq data with DESeq2. Genome Biology. 2014;15(12):550.

1040 100. Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, et al. Differential
1041 regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in
1042 different parts of the *Arabidopsis thaliana* seedling. The Plant Journal. 2007;50(4):660-77.

1043 101. Stracke R, Jahns O, Keck M, Tohge T, Niehaus K, Fernie AR, et al. Analysis of
1044 PRODUCTION OF FLAVONOL GLYCOSIDES-dependent flavonol glycoside accumulation in

1045 *Arabidopsis thaliana* plants reveals MYB11-, MYB12- and MYB111-independent flavonol
1046 glycoside accumulation. New Phytologist. 2010;188(4):985-1000.

1047 102. Vrhovsek U, Masuero D, Gasperotti M, Franceschi P, Caputi L, Viola R, et al. A versatile
1048 targeted metabolomics method for the rapid quantification of multiple classes of phenolics in
1049 fruits and beverages. Journal of Agricultural and Food Chemistry. 2012;60(36):8831-40.

1050 103. Arapitsas P, Mattivi F. LC-MS Untargeted Protocol for the Analysis of Wine. Methods
1051 in Molecular Biology. 2018;1738:225-35.

1052 104. Mehrtens F, Kranz H, Bednarek P, Weisshaar B. The Arabidopsis transcription factor
1053 MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. Plant Physiology.
1054 2005;138(2):1083-96.
1055