

RESEARCH

Open Access



Molecular evolution and expression assessment of *DFRs* in apple

Wen-Fang Li¹, Ju Gao¹, Zong-Huan Ma¹, Ying-Jun Hou¹, Xin Li¹, Juan Mao¹ and Bai-Hong Chen^{1*}

Abstract

Background Anthocyanins are the secondary metabolites of flavonoids in plants. As a key enzyme in the biosynthetic pathway of anthocyanin, dihydroflavonol 4-reductase (DFR) act as an important regulatory point, but *DFR* family genes has not been systematically characterized in apple (*Malus domestica* Borkh.).

Methods The members of *DFR* genes in apple were identified and their gene structure, chromosome distribution, evolutionary relationships, collinearity, cis-component and protein interaction relationships were predicted using bioinformatics methods. The expression patterns of *MdDFRs* in various organs, such as leaves, fruit flushes, fruits, ripe fruit peels, flowers and stems were analyzed using GeneChip expression array analysis. qRT-PCR was employed to analyze the expression levels of *MdDFRs* in different apple varieties with varying levels of fruit skin at maturity.

Results The apple database revealed 96 *DFR* genes, which are distributed on 17 chromosomes and can be divided into 3 subfamilies. These 96 *DFR* genes were mostly composed of α -helix and random coil according to secondary structure prediction, and were mainly expressed in chloroplasts and cytoplasm. MYB binding site involved in flavonoid biosynthetic genes regulation element (MBSI) was identified in the promoter of *MdDFR15/76/81/89/90/91/93/94*. Lignin/flavonoid synthesis-related elements of MYB recognition site and MYB-binding site were identified in the promoters of *MdDFR05/09/13/19/22/24/26/30/31/33/34/46/50/52/54/64/65/69/75/76/79/86*. The internal collinearity analysis of the apple *MdDFR* genome revealed a total of 34 pairs of duplicated gene pairs. Interspecific collinearity analysis showed that there were 66 and 57 homologous gene pairs in apple/tomato and apple/grape, respectively. GeneChip expression array analysis showed that *MdDFR72* and *MdDFR96* were higher expressed in ripe fruit flesh and peel, *MdDFR01/06/67/49/54/91* were higher expressed in flowers, *MdDFR64* was higher expressed in ripe fruit peels and flowers than those of other tissues. Besides, 75 *MdDFR* proteins interacted directly or indirectly with anthocyanidin synthesis related proteins MdANS, MdF3H, MdMYB1, MdMYBPA1 to form a protein interaction network. Interestingly, *MdDFR69* and *MdDFR87* had direct interactions with these four proteins, *MdDFR64* had direct interactions with MdANS and MdF3H. qRT-PCR analysis showed that the expression levels of *MdDFR01/05/31/53/64/69/73/84/87/94/96* were up-regulated with the accumulation of anthocyanins.

Conclusions This study lays a foundation for further research on the function of *DFR* genes in apple.

Keywords Apple, Anthocyanin, Dihydroflavonol 4-reductase (DFR), Molecular evolution, GeneChip expression array, Protein interaction network, Relative expression

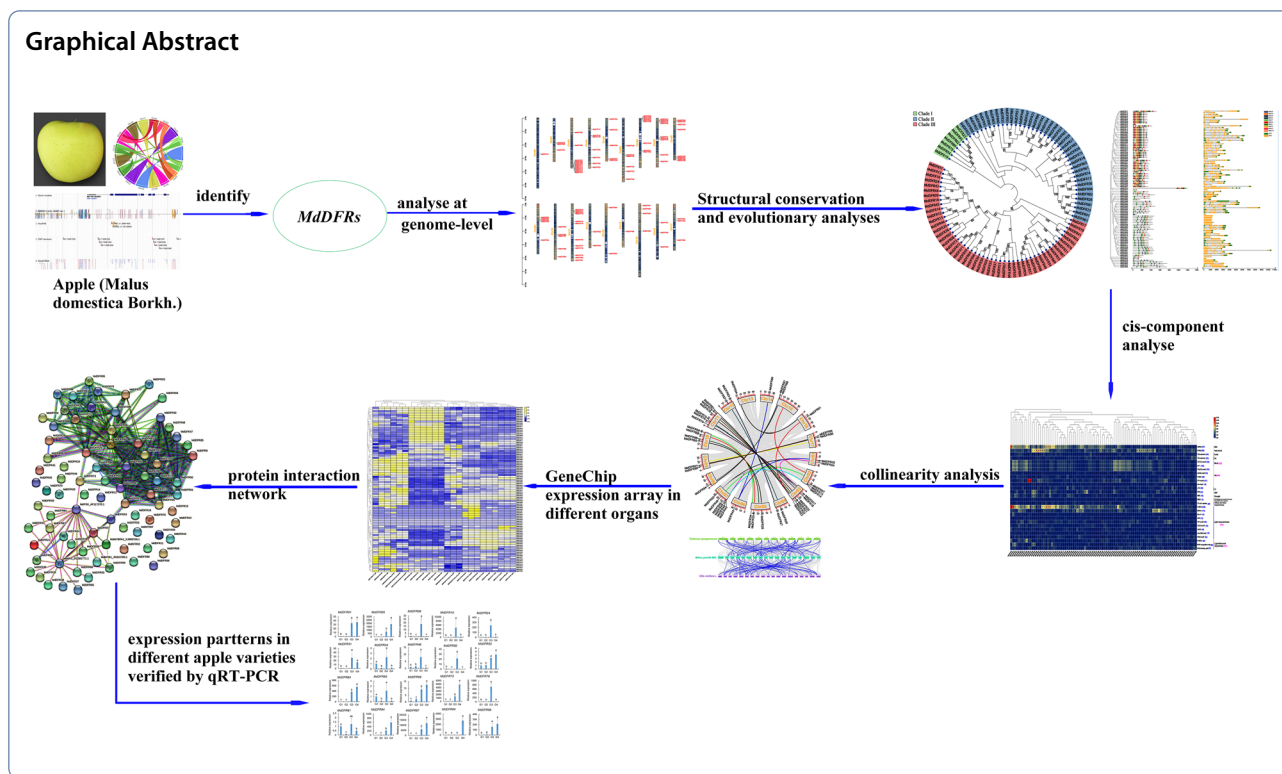
*Correspondence:

Bai-Hong Chen
bhch@gsau.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



Background

Apple (*Malus domestica* Borkh.) is a perennial woody plant of the Rosaceae family and is one of the most popular fruits in China, with a wide cultivation area and high yield throughout the country. Apple fruit is rich in minerals and vitamins, and its nutrient components are soluble, easy to be absorbed by the human body, with the role of calming the nerves, helping sleep, and reducing cholesterol [1]. Apple fruits also contain a certain amount of anthocyanins, which has the function of anti-aging and preventing cardiovascular and cerebrovascular diseases [2]. Anthocyanidin and proanthocyanidin are two end products of flavonoid biosynthesis pathway, which are believed to be synthesized and modified in cytoplasm and endoplasmic reticulum before being sequestered into vacuoles [3]. Anthocyanins are synthesized by phenylalanine via phenylalanine lyase (PAL), cinnamate-4-hydroxylase (C4H), 4-coumaryl CoA ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), and flavonoid 3'5'-hydroxylase (F3'5'H), dihydroflavonol 4-reductase (DFR), and anthocyanin synthase (ANS). DFR catalyzed stereospecific reduction reaction of dihydroflavonol at the C4 position, which enables dihydroflavonol to generate colorless proanthocyanidins, and is one of the most critical enzymes in anthocyanin synthesis [4]. Notably, DFR enzymes exhibit striking

substrate specificities, resulting in accumulation of distinct patterns of anthocyanins, catechins and proanthocyanidins [5].

F3H, as the upstream of DFR, can catalyze the synthesis of dihydroflavonols from flavonoids [6]. DFR further catalyzes catalyze three different colorless dihydroflavonols, including dihydro kaempferol (DHK), dihydro quercetin (DHQ) and dihydromyricetin (DHM), to generate different anthocyanidin precursors, and then synthesize various anthocyanin under the action of ANS [7]. It can be seen that DFR is the first enzyme to convert dihydroflavonol into anthocyanin. DFR can use DHK, DHQ, and DHM as substrates, respectively. The functional expression of *Saussurea medusa* gene *SmDFR* in yeast (*Saccharomyces cerevisiae*) confirms that *SmDFR* has a reducing effect on DHQ and DHK, but not on DHM [8]. The heterologous expression of grape hyacinth (*Muscari* spp.) gene *MaDFR* in tobacco (*Nicotiana tabacum*) resulted in increasing anthocyanin accumulation, leading to a darker flower color [9]. However, due to the different substrate selectivity of different species of DFR, different anthocyanins are synthesized, presenting different plant colors. Structural genes encoding enzymes of the anthocyanin pathways are regulated at the transcriptional level by a MYB-bHLH-WD40 (MBW) ternary complex of R2R3-MYB, basic helix-loop-helix (bHLH), and tryptophan aspartic (WD40)-repeat proteins [10, 11]. Among these

regulatory factors, the R2R3-MYB transcription factors, which play a crucial role in determining the activation of specific downstream genes to control the spatiotemporal accumulation of anthocyanins, have been widely reported in apple, including MdMYBPA1 and MdMYB1 [12, 13]. The regulatory interaction between transcription factors related to anthocyanin synthesis and structural genes has been a focus of research [14–16].

Our previous study reported that the expression of *DFR* was upregulated with the increase of anthocyanin accumulation revealed by comprehensive transcriptome analyses [17]. Due to the substrate specificity of *DFR*, different members of its gene family have different catalytic efficiency for each substrate and different expression levels in different tissues, which causes different tissues or organs of plants show various colors [18]. Given the importance of *DFRs* in the regulation of flavonoid synthesis and the production and accumulation of anthocyanidin, the expression and functional characterization of *DFR* family members in anthocyanidin biosynthesis of apple have not been systematically investigated, especially in different apple organs and varieties, as well as the relationships with other identified anthocyanidin synthesis related genes or transcription factors. Consequently, in light of the discovery of the apple *DFR* gene members, this experiment performed a bioinformatics analysis of the *MdDFR* gene members and predicted the interaction network between the *MdDFRs* and the anthocyanidin synthesis related proteins MdANS, MdF3H, MdMYB1, MdMYBPA1. The role of *DFR* genes in apple anthocyanin synthesis was examined through the analysis of GeneChip expression array of different apple tissues and expression levels of apples with different coloring degrees, which provided theoretical reference for the study of the regulatory network of apple anthocyanin synthesis.

Materials and methods

Plant materials

Apple varieties ‘Golden Delicious’, ‘Ruixue’, ‘Chengji No.1’ and ‘Astar’ with different coloring degrees at maturity were selected as research materials, and named G1, G2, G3 and G4, respectively. Fifteen fruits were collected from each variety, with every five fruits being a duplicate. Fruit skin samples were accurately weighed and quickly frozen in liquid nitrogen for 2 h, and then stored in a refrigerator at -80°C for RNA extraction.

Identification of *DFR* gene members in apple

The FASTA sequence of apple *DFR* gene CDS was obtained by searching gene names in NCBI network database (<https://www.ncbi.nlm.nih.gov/>). InterPro number (IPR001509) of this gene was obtained from the

Pfam website (<https://pfam.xfam.org/>) using the obtained FASTA sequence. Collect all the genes containing InterPro number in Phytozome v13 (<https://phytozome-next.jgi.doe.gov/>). The mRNA and CDS sequences of genes were downloaded from the apple reference genome website (<https://iris.angers.inra.fr/gddh13/>) using the accession number, and the corresponding protein sequences were obtained from the translation website (<https://web.expasy.org/translate/>). Chromosome localization of *DFR* gene members of apple was performed using TBtools (Version 1.108) software.

Bioinformatics analysis of *DFR* gene members in apple

ProtParam tool in ExPasy database (<https://web.expasy.org/protparam/>) was used to analyze the basic physicochemical properties of protein, such as the number, relative molecular weight and instability coefficient. ClustalX, MEGA7.0 and EvolView (<https://www.evolgenius.info/>) were used to construct a phylogenetic tree using the Neighbor-Joining method (NJ) with the bootstrap value of 1000. The gene structure was analyzed by GSDS 2.0 (<http://gsds.cbi.pku.edu.cn/>). The conserved motifs of *MdDFRs* were determined using the MEME online tool (<http://meme-suite.org/tools/meme>) the number of motifs output was 10. The motif analysis and gene structure were visualized using the TBtools software. The subcellular localization prediction of *DFR* gene members was performed online using WoLF PSORT (<http://www.genscript.com/wolf-psort.html>). The secondary structure prediction of *DFR* proteins was carried out using PredictProtein (<https://predictprotein.org/>). The chromosomal location information of *MdDFRs* was obtained from the apple genome website (<https://iris.angers.inra.fr/gddh13/>) and was drawn a picture by MG2C (http://mg2c.iask.in/mg2c_v2.0/).

Promoter cis-acting element analysis

The 2000 bp upstream promoter sequences of *MdDFR* genes were submitted to the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) and PLACE (<https://www.dna.affrc.go.jp/PLACE/>) to predict cis-elements and to subsequently screen cis-elements manually, then visualized at TBtools (Version 1.108).

Intraspecific and interspecific collinearity analysis of *MdDFR* genes

Download the genomes and annotated files of tomato and grape using Phytozome v13. The intraspecific collinearity of *MdDFR* gene pairs as well as the interspecific collinearity of *DFR* members from *Malus domestica*/*Solanum lycopersicum*, *Malus domestica*/*Vitis vinifera* L. was

calculated and visualized by MCSanX method at TBtools (Version 1.108).

Expression pattern and protein interaction analysis of *DFR* genes in apple

The FPKM values of 96 *MdDFRs* in different organs, including leaves, young leaves, fruit flushes, fruits, ripe fruit peels, flowers and stems, were obtained through the apple MDO multidimensional omics database (http://bioinformatics.cau.edu.cn/AppleMDO/gene_detail.php) and standardized with \log_{10} . The expression heat maps were drawn using TBtools. The protein interaction network was predicted by STRING Version 11 (<https://string-db.org/>).

Quantitative RT-PCR (qRT-PCR) analysis

Total RNA was extracted from 100 mg of frozen tissue using an RNA plant (Real Times, Beijing, China) basing on the manufacturer's instructions. The first strand cDNA fragment was synthesized from 1.0 μg of DNase created RNA in a reaction solution (total volume of 20 μL) by using the Evo M-MLV RT kit (Accurate Biology, Changsha, China). Gene specific primers (Additional file 1: Table S1) were designed and synthesized by Sangon Biotech Co., Ltd. (China). Gene expression was detected using qRT-PCR using SYBR[®] Green Premix Pro Taq HS qPCR Kit (Accurate Biology). Using apple *GADPH* (GenBank access NO. CB973647) gene as internal reference. Light Cycler[®] 96 RealTime PCR System (Roche, Basel, Switzerland) was used to determine the expression levels of *MdDFRs* in the peel of four apple cultivars. The reaction system is 20 μL , including 2 μL cDNA, 1 μL upstream and downstream primers, 10 μL SYBR, and 20 μL ddH₂O complement. The PCR reaction was as follows: 95 °C for 30 s, 40 cycles at 95 °C for 5 s and 60 °C for 1 min, 95 °C for 5 s, 55 °C for 30 s and 72 °C for 1 min. The relative expression levels of genes were calculated using the $2^{-\Delta\Delta\text{CT}}$ method [19].

Results

Physicochemical properties and subcellular localization analysis of *DFR* genes in apple

A total of 96 *DFR* genes, named *MdDFR01-MdDFR96*, were obtained through the homologous search on the apple reference genome website, and distributed on 17 apple chromosomes (Fig. 1, Table 1). Among them, chromosome 2 (chr2) and chr11 were the most distributed with 12 genes, followed by chr7 with 9 genes, chr6 and chr8 with 8 genes, chr0 and chr14 with 1 gene each. The number of amino acids encoded by this gene members was ranged from 75 aa (*MdDFR10*) to 1204 aa (*MdDFR70*). The relative molecular weight ranged from 8344.64 D (*MdDFR10*) to 136,293.11 D (*MdDFR70*). The

isoelectric point (PI) ranged from 5.02 (*MdDFR78*) to 9.96 (*MdDFR47*), including 28 basic proteins (PI > 7.5), 44 acidic proteins (PI < 6.5), and 24 neutral proteins (PI 6.5–7.5). The instability index ranged from 18.26 (*MdDFR29*) to 49.93 (*MdDFR05*). There were 23 hydrophilic proteins and 73 unstable hydrophobic proteins. The full length of the mRNA ranged from 353 bp (*MdDFR40*) to 10,524 bp (*MdDFR01*). The full length of the cDNA ranged from 228 bp (*MdDFR10*) to 3615 bp (*MdDFR70*). Among the 96 apple *DFR* genes, 47 were located in the chloroplast, 31 were located in the cytoplasm, and there were 4 genes each in cytoskeleton, extracellular matrix, and mitochondria. There were 3 genes each in nucleus and endoplasmic reticulum, and only 1 member in peroxisome (Table 1).

Secondary structure prediction analysis of *DFR* proteins in apple

The secondary structure of the protein encoded by the apple *DFR* genes was mainly α -helix and random coil, followed by extended strand, and β -corner structure was the least (Table 2). The proportion of α -helix structure was between 29.84% (*MdDFR23*) and 56.06% (*MdDFR70*). The proportion of irregular curly structures ranged from 26.67% (*MdDFR10*) to 48.17% (*MdDFR01*). Extended strand accounted for 10.71% (*MdDFR40*) to 23.00% (*MdDFR77*). β -corner accounted for the smallest proportion of 4.13% (*MdDFR18*) to 12.00% (*MdDFR10*). According to the above analysis, it can be concluded that α -helix and random coil were the main factors in the formation of spiral structure of apple *DFRs*.

Phylogenetic analysis of *DFR* genes in apple

To further understand the evolution and functional characteristics of apple *DFR* genes, the phylogenetic tree of apple *DFR* members was constructed and divided into three subfamilies (Clade I-III, Fig. 2). Among them, the Clade I subfamily includes 7 genes, including *MdDFR19*, *MdDFR22*, *MdDFR54*, *MdDFR62*, *MdDFR79*, *MdDFR89*, and *MdDFR95*. The Clade II subfamily includes 40 genes, and the Clade III subfamily includes 49 apple *DFR* genes (Fig. 2). In each branch, apple *DFR* genes have high homology and close evolutionary relationship, but there are some differences, which preliminarily indicates that their evolution is conservative, and it is speculated that the functions of the lineal homologous genes are similar.

Gene structure and motif sequence analysis of apple *DFRs*

To gain more insight into the evolutionary and structural diversity of *MdDFRs*, the conserved motifs in *MdDFR* proteins were analyzed using the MEME online software. A total of 10 distinct and highly conserved motifs were captured (Fig. 3A). The motif distribution pattern in most *MdDFR* proteins is highly conserved. Clade I

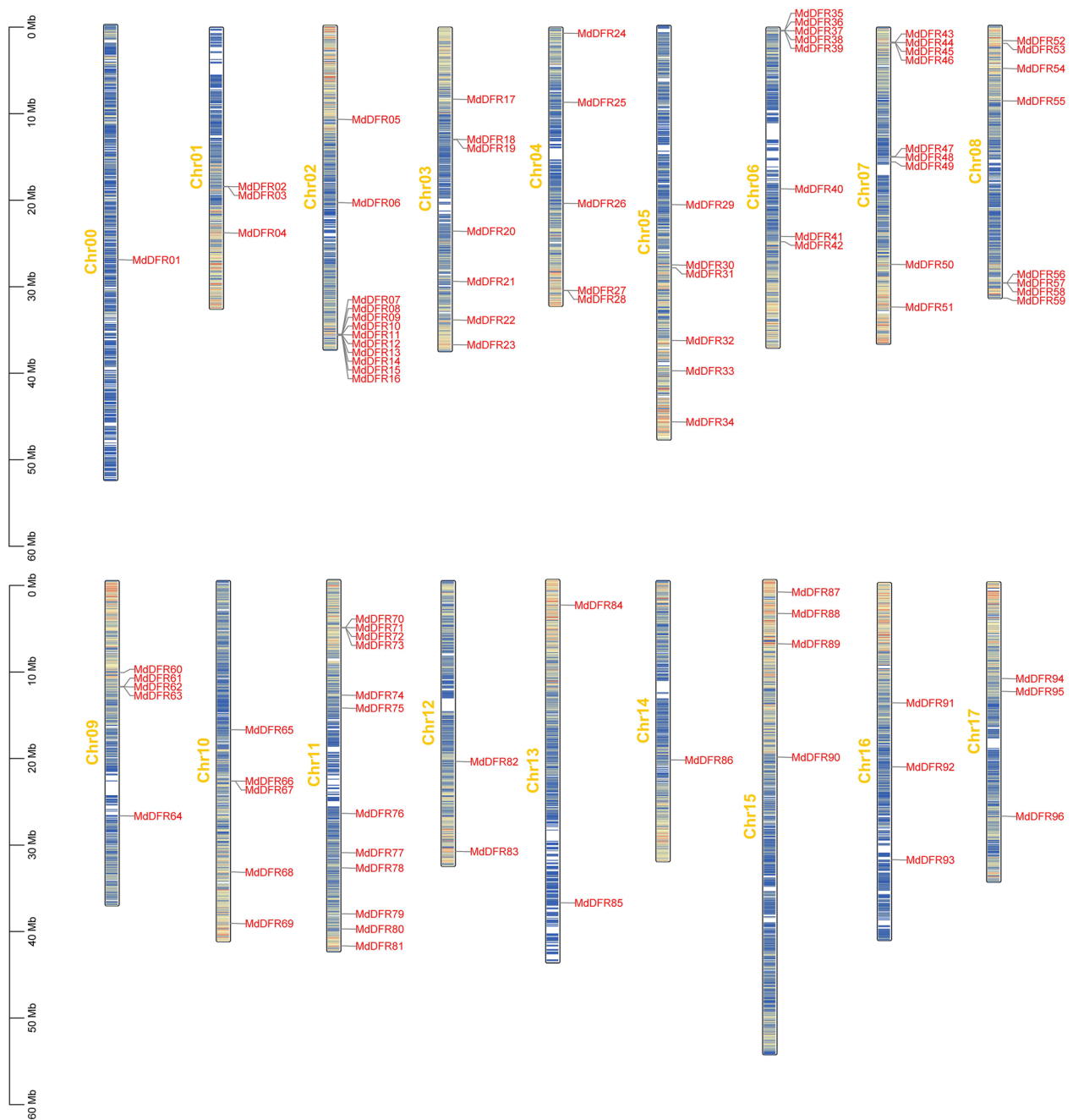


Fig. 1 Chromosome distribution of the *DFR* genes in apple. The left scale indicates the chromosome length (Mb), with *DFR* gene markers on the right side of each chromosome. Different chromosomal colors indicate different gene densities, with red indicating the highest density and blue the lowest density

mainly contain motif 1, motif 2, motif 3, motif 4, motif 7 and motif 8. Clade II and III mainly contains motif 7 and motif 10. To further explore the structural diversity of *MdDFRs*, the intron–exon organization of each gene

was analyzed. As shown in Fig. 3B, the exon number varied from 1 to 12, and most of the *MdDFR* genes have 1–7 exons. In addition, 6 genes contain 9 exons, 3 genes contain 8 exons, 4 genes contain 10 exons, and 4 genes contain 12 exons.

Table 1 Analysis of physicochemical properties and subcellular localization of DFR gene in Apple

Name	ID	Chromosome location	Amino acids	Molecular weight/D	pI	Instability index	Hydrophilic	Full length/bp	cDNA/bp	Subcellular localization
MdDFR01	MD00G1128000	0	465	52,188.03	8.61	32.12	-0.178	10,524	1398	Cytoplasm/Plasma membrane
MdDFR02	MD01G1077400	1	341	37,566.81	5.37	36.2	-0.046	2458	1026	Chloroplast
MdDFR03	MD01G1077900	1	342	37,702.34	6.72	31.8	-0.116	3089	1029	Nucleus
MdDFR04	MD01G1125200	1	342	38,387.13	7.73	32.83	-0.37	3320	1029	Cytoplasm
MdDFR05	MD02G1134100	2	372	41,821.12	6.32	49.93	-0.45	1952	1119	Cytoplasm
MdDFR06	MD02G1205000	2	329	36,336.67	5.8	33.8	-0.064	3377	990	Chloroplast
MdDFR07	MD02G1302800	2	324	35,564.88	6.02	33.99	-0.048	2888	975	Chloroplast
MdDFR08	MD02G1302900	2	322	34,972.3	6.32	30.42	0.08	4285	969	Chloroplast
MdDFR09	MD02G1303400	2	320	34,953.16	5.52	27.65	-0.019	2315	963	Chloroplast
MdDFR10	MD02G1303600	2	75	8344.64	9.31	24.47	-0.409	885	228	Cytoplasm
MdDFR11	MD02G1303800	2	520	56,745.3	5.86	30.74	-0.022	5546	1563	Cytoplasm
MdDFR12	MD02G1303900	2	320	34,689.02	6	22.93	0.09	1748	963	Chloroplast
MdDFR13	MD02G1304200	2	315	34,572.08	6.54	30.3	0.073	3088	948	Chloroplast
MdDFR14	MD02G1304400	2	326	35,846.39	6.26	32.13	0.028	3818	981	Cytoplasm
MdDFR15	MD02G1304900	2	326	35,504.95	6.25	28.63	0	3735	981	Chloroplast
MdDFR16	MD02G1305300	2	320	34,828.07	5.68	33.35	0.016	2152	963	Chloroplast
MdDFR17	MD03G1096800	3	346	39,263.97	6.53	29.46	-0.438	4203	1041	Chloroplast
MdDFR18	MD03G1131500	3	363	40,762.7	5.49	48.89	-0.233	2479	1092	Cytoplasm
MdDFR19	MD03G1131900	3	193	21,672.97	6.82	40.99	-0.018	1199	582	Cytoplasm
MdDFR20	MD03G1172800	3	325	35,447.67	6.02	28.25	-0.046	6365	978	Cytoplasm
MdDFR21	MD03G1214100	3	364	39,373.58	6.78	37.93	0.235	6025	1095	Chloroplast
MdDFR22	MD03G1250400	3	673	75,908.82	6.62	34.16	-0.345	3075	2022	Chloroplast
MdDFR23	MD03G1288200	3	439	49,147.28	8.72	37.33	-0.303	4475	1320	Cytoplasm
MdDFR24	MD04G1006100	4	315	34,687.1	5.82	34.54	0.06	2882	948	Chloroplast
MdDFR25	MD04G1064600	4	350	38,830.36	6.56	24.02	-0.283	3455	1053	Cytoplasm
MdDFR26	MD04G1118200	4	428	47,962.79	8.91	36.4	-0.366	3360	1287	Mitochondria
MdDFR27	MD04G1224100	4	459	51,977.55	7.59	39.65	-0.298	3314	1380	Cytoplasm
MdDFR28	MD04G1224200	4	376	42,590.4	5.99	39.74	-0.455	3467	1131	Cytoskeleton
MdDFR29	MD05G1099800	5	324	35,547.83	6.76	18.26	-0.052	1461	975	Cytoplasm
MdDFR30	MD05G1145800	5	142	15,915.48	5.43	38.67	0.164	617	429	Extracellular matrix
MdDFR31	MD05G1148900	5	327	36,450.14	6.71	34.49	-0.074	1747	984	Chloroplast
MdDFR32	MD05G1230600	5	465	51,041.33	9.58	36.9	-0.254	1398	1398	Cytoplasm/E/R
MdDFR33	MD05G1263800	5	418	45,991.49	8.42	39.05	-0.24	5579	1257	Mitochondria

Table 1 (continued)

Name	ID	Chromosome location	Amino acids	Molecular weight/D	pI	Instability index	Hydrophilic	Full length/bp	cDNA/bp	Subcellular localization
MdDFR34	MD05G1335600	5	339	36,873.53	5.52	35.57	0.008	2865	1020	Chloroplast
MdDFR35	MD06G1003200	6	287	32,201.16	5.59	33.48	-0.139	2409	864	Chloroplast
MdDFR36	MD06G1003300	6	315	35,070.25	6.95	37.45	-0.168	3385	948	ER
MdDFR37	MD06G1003500	6	339	37,957.84	5.89	34.05	-0.086	3376	824	Chloroplast
MdDFR38	MD06G1003700	6	315	35,290.7	7.54	36.28	-0.181	2515	948	Peroxisome
MdDFR39	MD06G1003800	6	310	34,978.2	6.67	39.19	-0.286	2033	933	Cytoplasm
MdDFR40	MD06G1075100	6	84	9752.16	9.26	43.83	-0.502	353	255	Extracellular matrix
MdDFR41	MD06G1103900	6	344	37,444.75	5.17	38.9	-0.01	2404	1035	Chloroplast
MdDFR42	MD06G1108600	6	384	41,987.91	6.66	33.72	-0.123	3709	1155	Chloroplast
MdDFR43	MD07G1020400	7	301	32,386.19	6.17	30.71	0.033	1771	906	Chloroplast
MdDFR44	MD07G1020900	7	321	34,990.43	6.33	30.32	0.055	6884	966	Cytoplasm
MdDFR45	MD07G1021200	7	321	35,146.63	5.75	30.64	0.039	2461	966	Chloroplast
MdDFR46	MD07G1021600	7	424	47,000.12	6.84	26.56	-0.071	5868	1275	Chloroplast
MdDFR47	MD07G1119800	7	438	48,556.83	9.96	41.94	-0.239	1491	1317	ER
MdDFR48	MD07G1120100	7	438	48,401.62	9.88	42.55	-0.216	1317	1317	ER
MdDFR49	MD07G1121700	7	265	29,593.13	6.5	37.64	0.029	1737	798	Chloroplast
MdDFR50	MD07G1194900	7	342	38,432.17	7.73	33.07	-0.363	3056	1029	Cytoplasm
MdDFR51	MD07G1256800	7	407	45,249.17	9.89	38.41	-0.105	1224	1224	Mitochondria
MdDFR52	MD08G1024600	8	343	38,584.95	9.06	46.41	-0.491	2696	1032	Chloroplast
MdDFR53	MD08G1028600	8	347	38,938.78	6.02	31.36	-0.19	2740	1044	Chloroplast
MdDFR54	MD08G1063500	8	130	14,689.74	5.49	33.33	0.014	393	393	Cytoplasm/Cytoskeleton
MdDFR55	MD08G1103100	8	340	37,107.64	8.87	25.57	0.008	1611	1023	Chloroplast
MdDFR56	MD08G1231300	8	267	29,735.22	6.54	33.32	-0.086	1230	804	Cytoplasm
MdDFR57	MD08G1231400	8	331	36,275.57	6.15	28.92	0.001	7363	996	Chloroplast
MdDFR58	MD08G1231500	8	331	36,464.83	5.93	33.05	0.02	3253	996	Chloroplast
MdDFR59	MD08G1250300	8	333	36,345.35	5.79	36.65	-0.097	2531	1002	Cytoplasm
MdDFR60	MD09G1137100	9	492	54,643.35	8.26	41.98	-0.131	3450	1479	Chloroplast
MdDFR61	MD09G1154200	9	349	39,647.33	8.77	23.03	-0.321	1050	1050	Chloroplast/Peroxisome
MdDFR62	MD09G1154300	9	755	85,437.77	8.89	33.75	-0.368	3160	2268	Cytoskeleton
MdDFR63	MD09G1154700	9	439	49,246.31	7.6	35.14	-0.313	2577	1320	Chloroplast
MdDFR64	MD09G1224200	9	339	37,321.48	5.77	34.59	-0.277	6111	1020	Chloroplast
MdDFR65	MD10G1105500	10	324	35,560.83	6.8	19.26	-0.066	1427	975	Chloroplast
MdDFR66	MD10G1145200	10	388	42,974.74	6.82	46.02	-0.058	4041	1167	Nucleus

Table 1 (continued)

Name	ID	Chromosome location	Amino acids	Molecular weight/D	pI	Instability index	Hydrophilic	Full length/bp	cDNA/bp	Subcellular localization
<i>MdDFR67</i>	MD10G1145300	10	357	39,608.62	6.36	39.52	-0.095	1804	1074	Extracellular matrix
<i>MdDFR68</i>	MD10G1241500	10	418	46,024.61	8.61	35.93	-0.236	5481	1257	Mitochondria
<i>MdDFR69</i>	MD10G1311100	10	339	36,781.43	5.29	34.43	0.027	4079	1020	Chloroplast
<i>MdDFR70</i>	MD11G1063800	11	1204	136,293.11	6.6	36.11	-0.196	5680	3615	Nucleus
<i>MdDFR71</i>	MD11G1064000	11	324	35,481.08	6.16	26.84	-0.005	3081	975	Chloroplast
<i>MdDFR72</i>	MD11G1064600	11	324	35,447.98	5.88	25.43	0.004	2792	975	Cytoplasm
<i>MdDFR73</i>	MD11G1064700	11	324	35,522.98	5.58	23.38	-0.003	2775	975	Cytoplasm
<i>MdDFR74</i>	MD11G1143300	11	346	39,263.04	7.74	29.68	-0.449	4429	1041	Chloroplast
<i>MdDFR75</i>	MD11G1154500	11	389	43,741.02	5.71	48.47	-0.295	3724	1170	Cytoplasm
<i>MdDFR76</i>	MD11G1190800	11	325	35,552.86	6.61	26.41	-0.056	5545	978	Chloroplast
<i>MdDFR77</i>	MD11G1215900	11	100	11,167.15	5.53	41.48	0.408	525	303	Extracellular matrix
<i>MdDFR78</i>	MD11G1229100	11	260	28,157.07	5.02	41.73	0.147	2656	783	Cytoplasm
<i>MdDFR79</i>	MD11G1270500	11	461	52,631.28	7.14	37.19	-0.482	1802	1386	Cytoskeleton
<i>MdDFR80</i>	MD11G1285100	11	403	43,716.68	7.16	44.6	-0.172	5323	1212	Chloroplast
<i>MdDFR81</i>	MD11G1309900	11	234	26,339.04	6.66	28.68	-0.492	3635	705	Chloroplast
<i>MdDFR82</i>	MD12G1133600	12	445	49,829.97	8.76	35.09	-0.275	2843	1338	Cytoplasm
<i>MdDFR83</i>	MD12G1240900	12	381	43,184.13	5.99	39.41	-0.149	3415	1146	Cytoskeleton
<i>MdDFR84</i>	MD13G1042500	13	319	35,187.18	6.25	38.69	-0.134	2041	960	Cytoplasm/Chloroplast
<i>MdDFR85</i>	MD13G1274600	13	432	47,898.48	9.63	38.01	-0.263	1299	1299	Cytoplasm
<i>MdDFR86</i>	MD14G1129500	14	347	38,036.26	6.46	28.81	-0.175	3965	1044	Cytoplasm
<i>MdDFR87</i>	MD15G1024100	15	348	39,029.79	5.68	36.2	-0.218	2708	1047	Chloroplast
<i>MdDFR88</i>	MD15G1058400	15	483	52,393.5	8.23	35.08	-0.252	2243	1452	Chloroplast
<i>MdDFR89</i>	MD15G1106300	15	660	74,793.14	6.55	28.31	-0.385	2161	1983	Cytoplasm
<i>MdDFR90</i>	MD15G1247100	15	372	41,806.11	6.36	48.59	-0.457	1139	1119	Cytoplasm
<i>MdDFR91</i>	MD16G1167000	16	437	48,713.5	8.31	34.9	-0.315	3282	1314	Chloroplast
<i>MdDFR92</i>	MD16G1216800	16	157	17,742.56	6.17	27.73	0.145	1099	474	Cytoplasm/Chloroplast
<i>MdDFR93</i>	MD16G1260200	16	432	47,877.41	9.54	39.65	-0.244	1299	1299	Cytoplasm
<i>MdDFR94</i>	MD17G1127400	17	364	40,246.83	8.78	35.54	-0.292	8743	1095	Cytoplasm
<i>MdDFR95</i>	MD17G1140800	17	745	84,110.34	7.27	34.43	-0.265	3721	2238	Chloroplast/Vacuole
<i>MdDFR96</i>	MD17G1222400	17	339	37,176.33	6.02	32.71	-0.24	6210	1020	Chloroplast

Table 2 Secondary structure analysis of DFR gene family proteins in Apple

Gene name	Alpha helix (%)	Random coil (%)	Extended strand (%)	Beta turn (%)	Gene name (%)	Alpha helix (%)	Random coil (%)	Extended strand (%)	Beta turn (%)
<i>MdDFR01</i>	31.18	48.17	15.05	5.59	<i>MdDFR49</i>	43.40	34.72	15.09	6.79
<i>MdDFR02</i>	39.00	37.83	14.96	8.21	<i>MdDFR50</i>	35.67	43.57	14.62	6.14
<i>MdDFR03</i>	42.69	37.72	13.16	6.43	<i>MdDFR51</i>	41.77	36.61	14.74	6.88
<i>MdDFR04</i>	35.96	42.40	15.50	6.14	<i>MdDFR52</i>	38.78	40.52	14.87	5.83
<i>MdDFR05</i>	42.20	38.98	14.25	4.57	<i>MdDFR53</i>	41.21	36.89	14.70	7.20
<i>MdDFR06</i>	40.43	36.47	16.41	6.69	<i>MdDFR54</i>	49.23	26.92	14.62	9.23
<i>MdDFR07</i>	42.28	34.26	16.05	7.41	<i>MdDFR55</i>	42.65	35.59	13.82	7.94
<i>MdDFR08</i>	43.17	33.54	16.15	7.14	<i>MdDFR56</i>	43.82	38.95	11.61	5.62
<i>MdDFR09</i>	42.19	34.06	17.19	6.56	<i>MdDFR57</i>	40.18	36.86	14.80	8.16
<i>MdDFR10</i>	41.33	26.67	20.00	12.00	<i>MdDFR58</i>	39.88	37.46	15.11	7.55
<i>MdDFR11</i>	40.96	34.23	17.69	7.12	<i>MdDFR59</i>	42.64	38.74	12.31	6.31
<i>MdDFR12</i>	42.50	34.69	15.00	7.81	<i>MdDFR60</i>	32.72	40.85	18.90	7.52
<i>MdDFR13</i>	41.90	31.75	18.10	8.25	<i>MdDFR61</i>	37.54	37.54	16.05	8.88
<i>MdDFR14</i>	40.49	34.97	16.87	7.67	<i>MdDFR62</i>	35.63	38.94	17.62	7.81
<i>MdDFR15</i>	41.10	34.66	17.18	7.06	<i>MdDFR63</i>	38.50	41.23	13.44	6.83
<i>MdDFR16</i>	42.50	34.69	15.94	6.88	<i>MdDFR64</i>	42.18	38.64	13.96	5.31
<i>MdDFR17</i>	35.55	40.75	15.61	8.09	<i>MdDFR65</i>	44.75	35.80	12.96	6.48
<i>MdDFR18</i>	33.88	44.63	17.36	4.13	<i>MdDFR66</i>	43.04	34.02	15.72	7.22
<i>MdDFR19</i>	44.56	32.12	13.99	9.33	<i>MdDFR67</i>	41.18	38.10	14.85	5.88
<i>MdDFR20</i>	40.92	34.46	16.31	8.31	<i>MdDFR68</i>	35.17	43.30	14.59	6.94
<i>MdDFR21</i>	41.76	37.09	14.29	6.87	<i>MdDFR69</i>	41.30	37.17	15.04	6.49
<i>MdDFR22</i>	39.23	38.19	16.49	6.09	<i>MdDFR70</i>	56.06	28.65	11.05	4.24
<i>MdDFR23</i>	29.84	46.70	16.86	6.61	<i>MdDFR71</i>	42.28	37.04	13.27	7.41
<i>MdDFR24</i>	40.63	34.29	17.46	7.62	<i>MdDFR72</i>	45.06	33.33	14.20	7.41
<i>MdDFR25</i>	40.00	37.14	15.71	7.14	<i>MdDFR73</i>	45.37	32.72	14.20	7.72
<i>MdDFR26</i>	31.54	48.13	13.79	6.54	<i>MdDFR74</i>	36.99	41.04	14.74	7.23
<i>MdDFR27</i>	39.65	38.56	15.25	6.54	<i>MdDFR75</i>	39.07	41.39	14.40	5.14
<i>MdDFR28</i>	40.16	39.63	14.89	5.32	<i>MdDFR76</i>	40.31	38.15	15.08	6.46
<i>MdDFR29</i>	43.21	35.49	13.27	8.02	<i>MdDFR77</i>	33.00	36.00	23.00	8.00
<i>MdDFR30</i>	44.37	30.28	16.90	8.45	<i>MdDFR78</i>	40.77	42.69	11.92	4.62
<i>MdDFR31</i>	44.95	33.33	14.07	7.65	<i>MdDFR79</i>	39.48	37.31	15.18	8.03
<i>MdDFR32</i>	41.72	39.14	12.47	6.67	<i>MdDFR80</i>	31.51	46.15	17.12	5.21
<i>MdDFR33</i>	37.32	40.91	15.55	6.22	<i>MdDFR81</i>	40.60	37.61	14.10	7.69
<i>MdDFR34</i>	41.30	37.46	14.45	6.78	<i>MdDFR82</i>	34.16	46.52	13.48	5.84
<i>MdDFR35</i>	45.99	35.89	12.20	5.92	<i>MdDFR83</i>	37.27	40.42	16.27	6.04
<i>MdDFR36</i>	46.98	33.97	13.65	5.40	<i>MdDFR84</i>	42.32	34.48	15.36	7.84
<i>MdDFR37</i>	46.31	31.56	15.34	6.78	<i>MdDFR85</i>	43.98	36.81	12.50	6.71
<i>MdDFR38</i>	46.67	33.65	13.65	6.03	<i>MdDFR86</i>	40.35	36.89	16.43	6.34
<i>MdDFR39</i>	48.06	30.65	14.19	7.10	<i>MdDFR87</i>	41.67	36.78	14.66	6.90
<i>MdDFR40</i>	53.57	29.76	10.71	5.95	<i>MdDFR88</i>	33.13	45.55	16.77	4.55
<i>MdDFR41</i>	43.02	35.17	13.95	7.85	<i>MdDFR89</i>	38.94	36.52	16.97	7.58
<i>MdDFR42</i>	37.50	38.54	16.67	7.29	<i>MdDFR90</i>	41.94	37.63	14.78	5.65
<i>MdDFR43</i>	40.86	35.55	16.61	6.98	<i>MdDFR91</i>	31.58	45.77	16.93	5.72
<i>MdDFR44</i>	41.12	34.89	17.76	6.23	<i>MdDFR92</i>	29.94	43.95	19.11	7.01
<i>MdDFR45</i>	41.74	35.51	15.26	7.48	<i>MdDFR93</i>	43.29	38.19	11.81	6.71
<i>MdDFR46</i>	42.69	33.49	16.75	7.08	<i>MdDFR94</i>	35.44	39.56	17.31	7.69
<i>MdDFR47</i>	43.15	35.39	14.84	6.62	<i>MdDFR95</i>	34.90	36.38	19.19	9.53
<i>MdDFR48</i>	42.01	38.13	14.38	5.48	<i>MdDFR96</i>	41.30	40.41	13.57	4.72

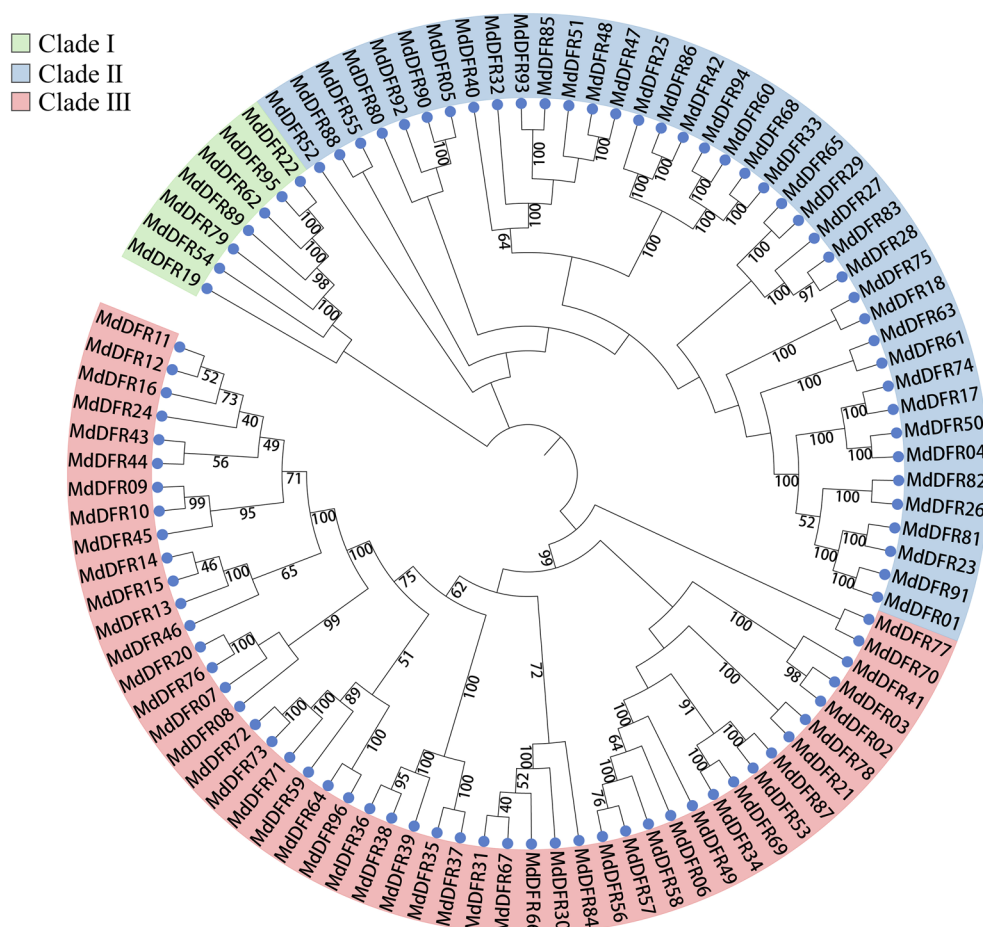


Fig. 2 Phylogenetic analysis of the *MdDFRs*. Phylogenetic trees were constructed using the DFR protein sequences. NJ method was adopted, and the bootstrap value was set to be equal to 1000

Prediction of cis-component analysis of *MdDFRs*

To explore the mechanism of *MdDFR* genes in plant growth and development, the cis-regulatory elements in the 2 kb upstream sequence of *MdDFR* genes were predicted by PlantCARE (Fig. 4). In the *MdDFR* genes' promoter, a variety of hormones (ABA, Auxin, SA, MeJA, GA, ERF)-, stress (heat shock, drought, low temperature, defense and stress)-, MYB binding site involved in flavonoid biosynthetic genes regulation-, light responsiveness- and lignin/flavonoid biosynthesis-related cis-acting elements were identified. Among them, light responsive related cis-acting elements were the most at 751, followed by GA related at 542, MeJA related at 322, and lighting/flavonoid biosynthesis related at 211. Abscisic acid (ABA) response element (ABRE), auxin response element (TGA-element), salicylic acid (SA) response element (TCA-element), MeJA response element (CGTCA-motif and as-1), gibberellin (GA) response element (TGACG-motif, GARE-motif, P-BOX, GT1-motif, GA-motif), low temperature response element (LTR),

ERF response element (ERE), drought response element (MBS), MYB binding site involved in flavonoid biosynthetic genes regulation element (MBSI), defense and stress responsiveness element (TC-rich repeats), light responsiveness element (G-box, Box4, Box4 II, ACE TCT-motif, TCCC-motif, I-box, chs-CMA1a/2a, and WUN-motif), and lignin/flavonoid biosynthesis-related cis-acting elements (W-BOX, MYB recognition site and MYB-binding site) were present in almost all *MdDFR* promoter sequences. MBSI element was identified in the promoter of *MdDFR15/76/81/89/90/91/93/94*. Lignin/flavonoid synthesis-related elements of MYB recognition site and MYB-binding site were both identified in the promoters of *MdDFR05/09/13/19/22/24/26/30/31/33/34/46/50/52/54/64/65/69/75/76/79/86*.

Intraspecific and interspecific colinearity analysis of *DFRs*

The internal colinearity analysis of the apple *MdDFR* genome revealed a total of 34 pairs of duplicated gene pairs (Fig. 5A), which were located on chromosomes

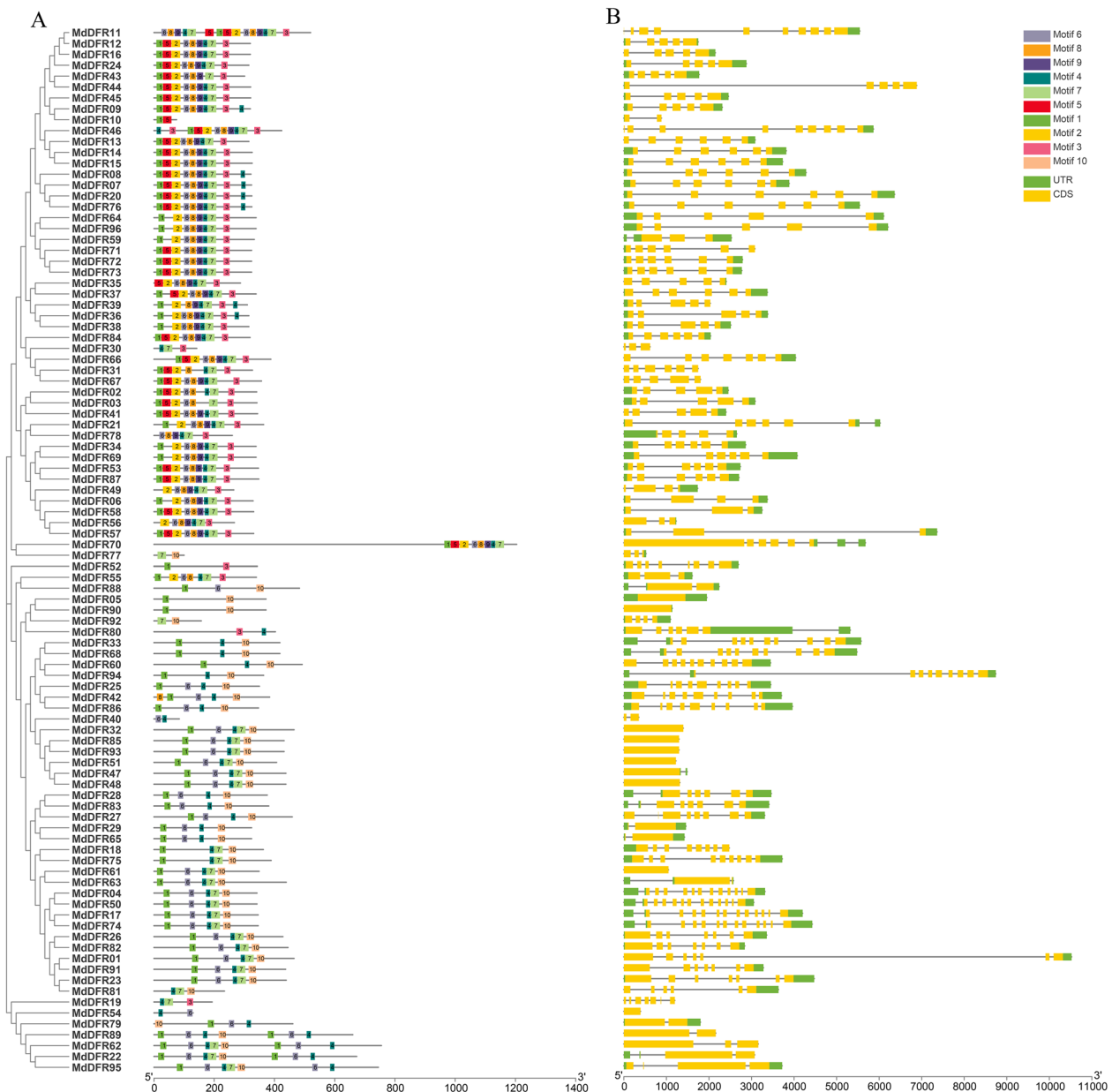


Fig. 3 Motif and gene structure analysis of *MdDFRs*. **A** Analysis of conserved motif of *DFRs* in apple. **B** The exon–intron structure of *MdDFRs*

chr1, chr2, chr3, chr4, chr5, chr6, chr7, ch8, chr9 and chr13. Most of these duplicated gene pairs were located on chromosomes 2 and 3 with 8 and 9 pairs, respectively. This indicates that some *MdDFR* genes may be generated through gene duplication, which may have similar functions. However, no linked genes were found on chromosomes 10, 11, 12, 14, 15 and 16. To further elucidate the phylogenetic mechanism and homology relationship of the *DFR* genes in apple, synteny analysis was conducted between apple and two representative species, grape (*Vitis vinifera* L.) and tomato (*Solanum lycopersicum*).

The results showed that there were 66 pairs of homologous genes in apple and tomato, mainly concentrated in apple chr3, chr7 and chr11, with 8, 10 and 8 pairs of homologous genes, respectively (Fig. 5B, Dataset1). There were 57 pairs of homologous genes in apple and grape, which were mainly concentrated in apple chr3, chr7 and chr11, with 7, 7 and 9 pairs of homologous genes, respectively. Therefore, it is speculated that the *DFR* gene is relatively conservative in evolution. In addition, some *DFR* genes were identified to be associated with 2–3 pairs of orthologs, whereas some had no homologous pairs,

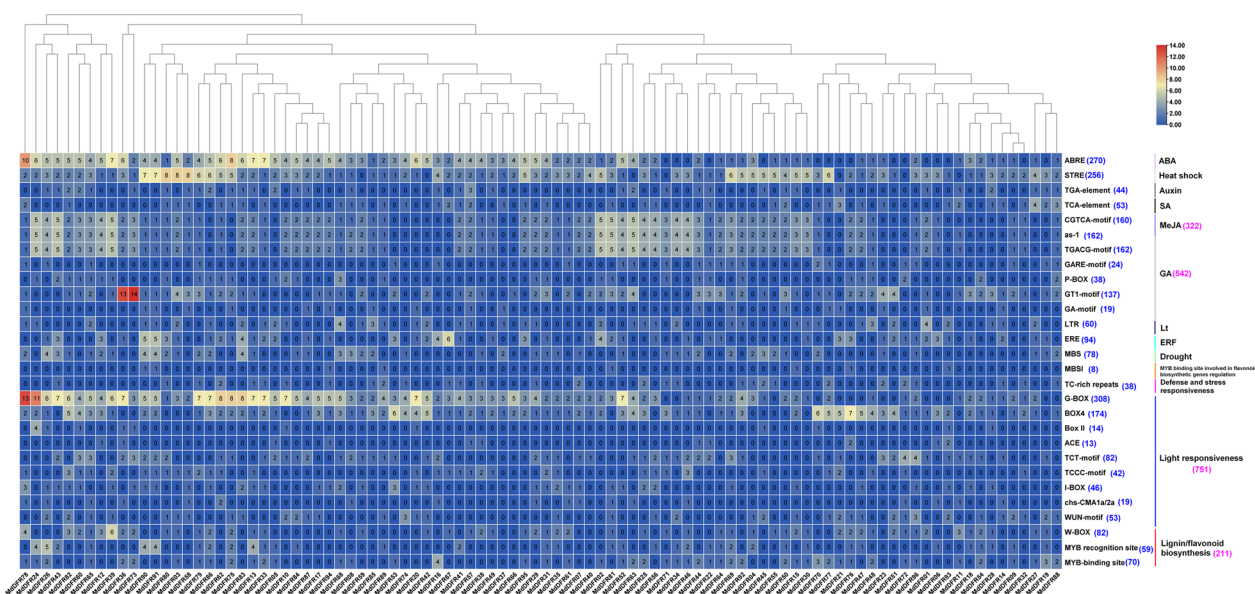


Fig. 4 Cis-regulatory element analysis of the *MdDFRs*. Different numbers represent the number of cis-regulatory elements, and the color from blue to red represents the number of elements from fewer to more

which may be due to differences in evolutionary speed and genetic specificity between species.

Expression pattern analysis and protein interaction prediction

The expression patterns of *MdDFR* gene members in different developmental organs were analyzed, including leaves, young leaves, fruits, fruit fleshes, fruit peels, flowers and stems. It was found that the expression levels in the same subfamily were similar (Fig. 6). The expression level of *MdDFR72*, *MdDFR96* was higher in ripe fruit fleshes and peels, whereas lower in other tissues. The expression level of *MdDFR01*, *MdDFR06*, *MdDFR67*, *MdDFR49*, *MdDFR54* and *MdDFR91* was higher in flowers, whereas lower in other tissues. The expression level of *MdDFR64* was higher in ripe fruit peels and flowers, whereas lower in other tissues.

The interactions among 96 *MdDFR* proteins were predicted by the STRING online website. The results showed that 75 out of 96 *MdDFR* proteins form a complex network structure (Fig. 7). The three-dimensional structure of the *MdDFR05*, *MdDFR20*, *MdDFR28*, *MdDFR53*, *MdDFR62*, *MdDFR64*, *MdDFR69*, *MdDFR72*, *MdDFR76*, *MdDFR79*, *MdDFR83*, *MdDFR87*, *MdDFR88*, *MdDFR90*, *MdDFR95* and *MdDFR96* proteins is known. Besides, these 75 *MdDFR* proteins interacted directly or indirectly with anthocyanidin synthesis related proteins *MdANS* (DQ381771.1), *MdF3H* (AF117270.1), *MdMYB1* (JN315783.1), *MdMYBPA1* (KJ909759.1) to form a protein interaction network. Among these interacting

proteins, *MdDFR21*, *MdDFR31*, *MdDFR35*, *MdDFR36*, *MdDFR37*, *MdDFR38*, *MdDFR39*, *MdDFR55*, *MdDFR58*, *MdDFR59*, *MdDFR64*, *MdDFR67*, *MdDFR69*, *MdDFR78*, *MdDFR84*, *MdDFR87* and *MdDFR96* were directly interacted with *MdANS*. *MdDFR21*, *MdDFR31*, *MdDFR53*, *MdDFR58*, *MdDFR59*, *MdDFR64*, *MdDFR66*, *MdDFR67*, *MdDFR69*, *MdDFR73*, *MdDFR78*, *MdDFR84*, *MdDFR87* and *MdDFR96* were directly interacted with *MdF3H*. *MdDFR53*, *MdDFR58*, *MdDFR69* and *MdDFR87* were directly interacted with *MdMYBPA1*. *MdDFR87* and *MdDFR69* were directly interacted with *MdMYB1*. Interestingly, *MdDFR69* and *MdDFR87* interact directly with these four anthocyanidin synthesis related proteins, *MdANS*, *MdF3H*, *MdMYBPA1* and *MdMYB1*.

Expression analysis of DFR gene members in different apple cultivars

Based on the above results, 20 aforementioned *MdDFRs* were selected for expression level analysis in the fruit skin of four different colored apple varieties. Our previous research revealed that the degree of pigmentation and anthocyanin content in the fruit skin from G1 to G4 apple varieties were gradually increased [17]. The results of qRT-PCR analysis in this research showed that the expression levels of *MdDFR01*, *MdDFR05*, *MdDFR31*, *MdDFR53*, *MdDFR64*, *MdDFR69*, *MdDFR73*, *MdDFR84*, *MdDFR87*, *MdDFR94* and *MdDFR96* were up-regulated with the increase of fruit skin pigmentation and the accumulation of anthocyanins (Fig. 8). Among these 20 genes, the expression level of *MdDFR87* was the highest.

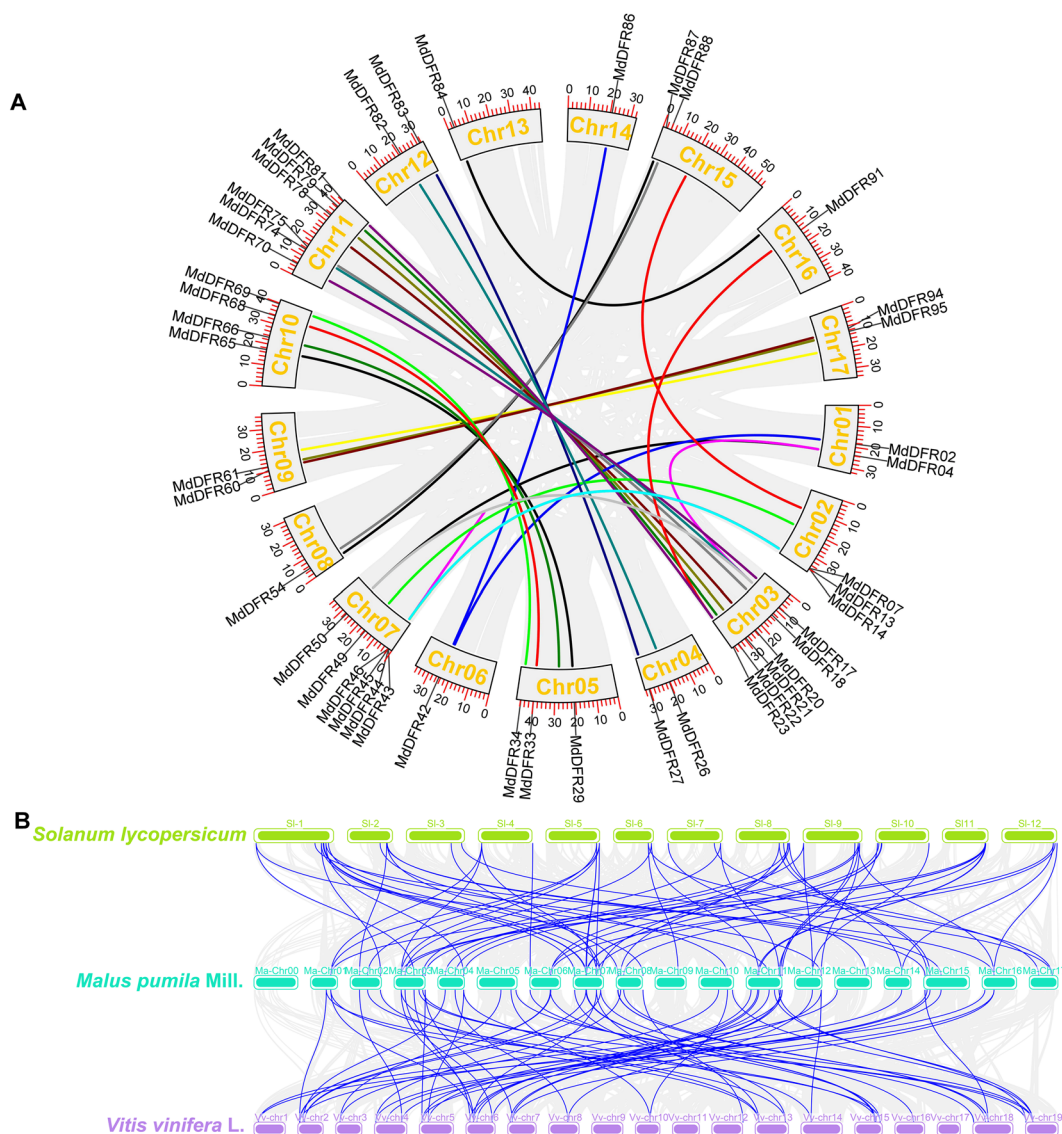


Fig. 5 Collinearity analysis of *DFRs*. **A** Collinearity analysis of *MdDFRs*. **B** Collinearity analysis of *DFRs* in apple and two representative plants tomato (*Solanum lycopersicum*) and grape (*Vitis vinifera* L.). The gray lines in the background show collinearity between the apple and tomato/grape genomes. The blue lines show collinearity between the *MdDFRs* and tomato/grape genome

Besides, the expression levels of *MdDFR06*, *MdDFR15*, *MdDFR24*, *MdDFR46*, *MdDFR50* and *MdDFR78* were significantly higher in G3 than those of other three apple varieties. However, the expression levels of *MdDFR34*, *MdDFR65*, and *MdDFR81* were all lower, and the expression levels in G1 and G3 were higher than those in G2 and G4.

Discussion

After being extracted from maize and snapdragon using the transposon tagging approach [20], the *DFR* gene was later discovered in other plants [21–23]. The

96 gene members of *DFR* were categorized into three subclasses (Fig. 2, 3), and there were 66 and 57 pairs of homologous genes in apple/tomato and apple/grape (Fig. 5b), respectively, exhibiting significant sequence conservation in apple. *DFR* is a structural gene that regulates the phenylalanine metabolic pathway and has a significant impact on plant coloring [24]. In transgenic tobacco lines, overexpression of *Hosta ventricose* gene *HvDFR* enhanced anthocyanin accumulation and elevated critical genes, such as *F3H*, *F3'H*, *ANS*, and *UFGT* [25]. Heterologous expression of *Ophiorrhiza japonica* gene *OjDFR1* in transgenic tobacco contributed to

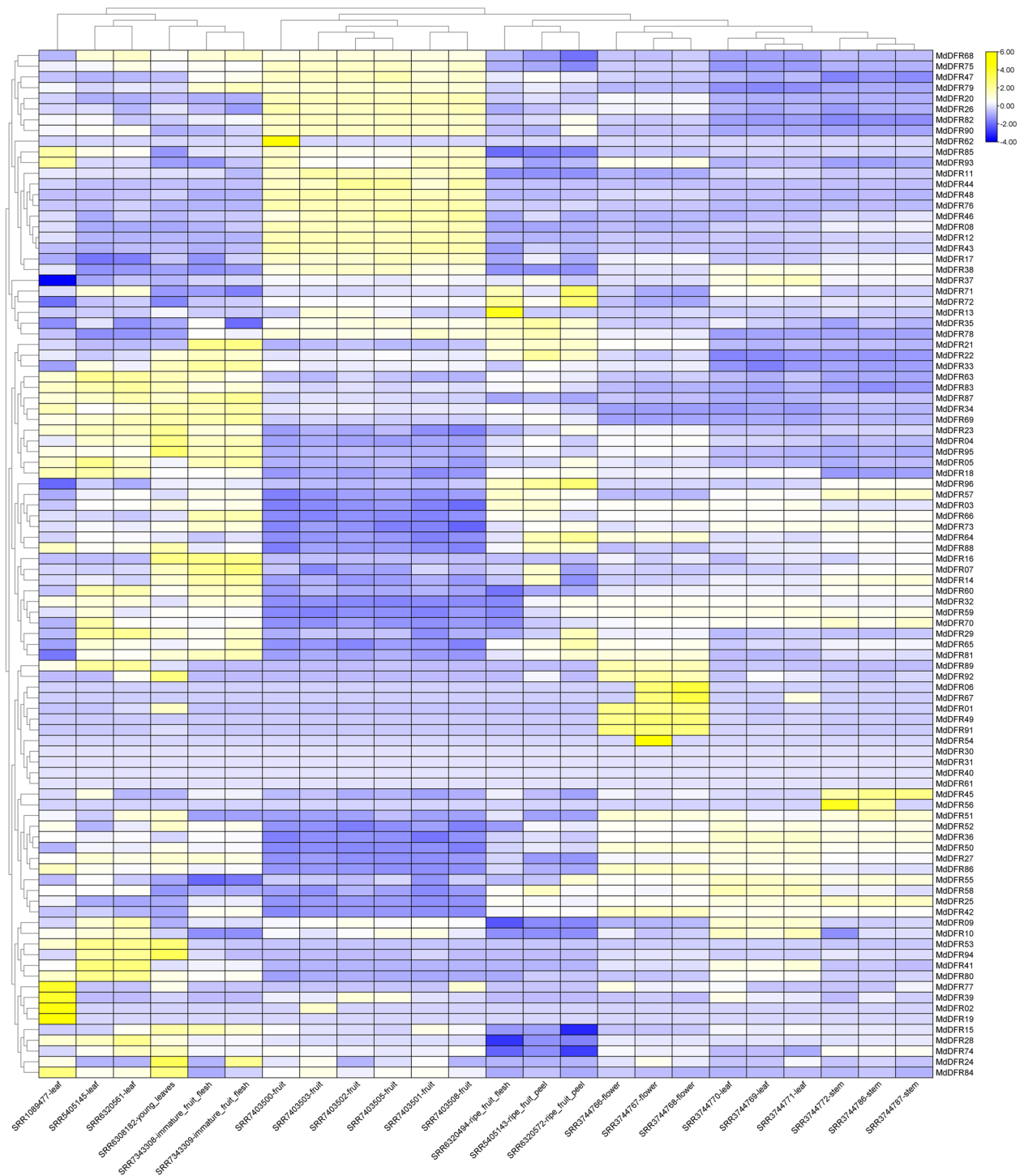


Fig. 6 Expression of *MdDFRs* in different tissues of apple. Yellow and blue shading represented the up-regulated or down-regulated expression level, respectively. The scale denoted the relative expression level

darker flower color via up-regulating the expressions of endogenous *NtANS* and *NtUFGT* [26]. Physicochemical property analysis showed that most *MdDFR*

genes were acidic proteins, followed by basic proteins and neutral proteins (Table 1). Subcellular predictions found that 47 *MdDFRs* were located in the chloroplast,

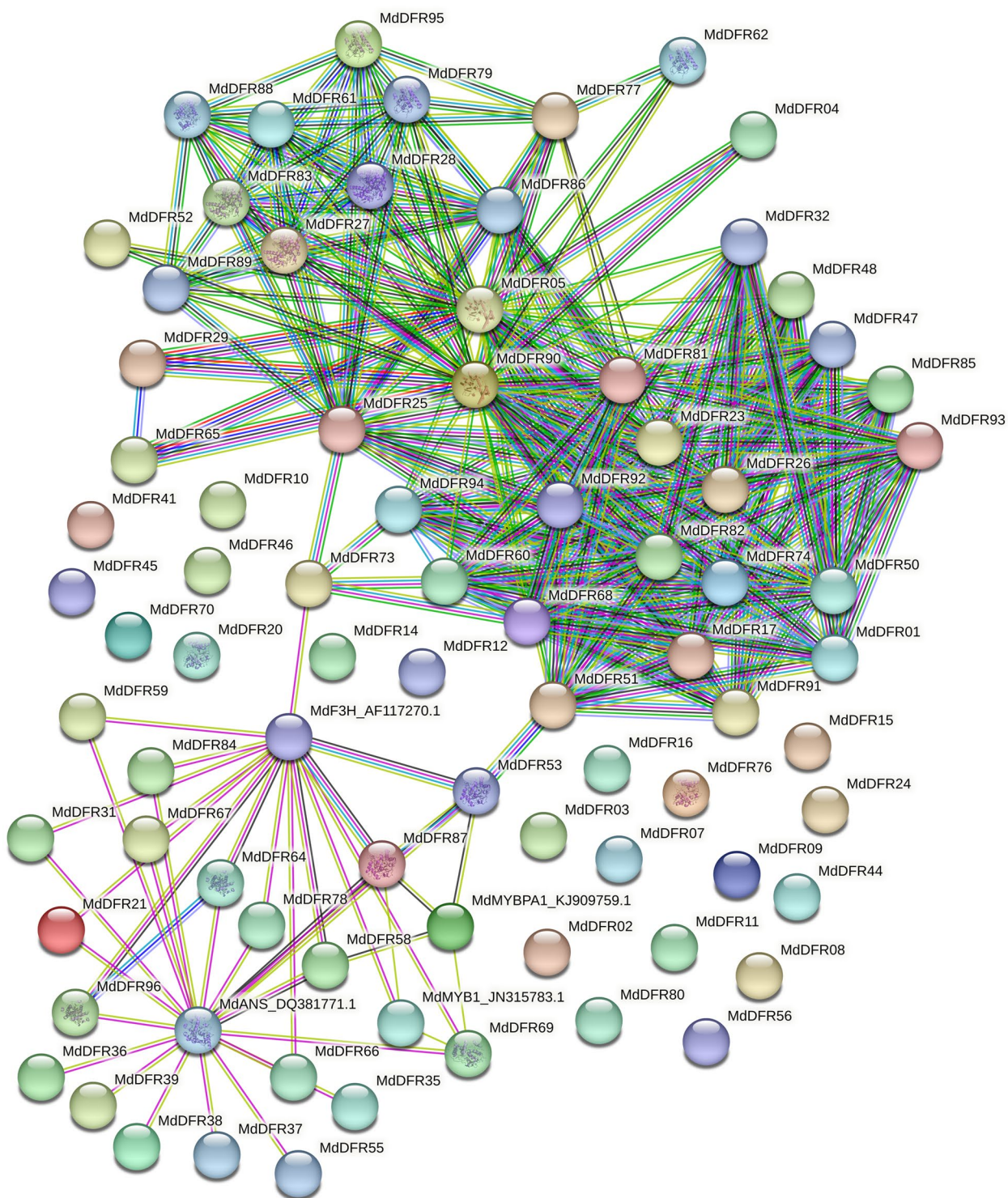


Fig. 7 Analysis of protein interaction of 96 DFRs, ANS (DQ381771.1), F3H (AF117270.1), MYB1 (JN315783.1) and MYBPA1 (KJ909759.1) in apple. Nodes indicate proteins. Empty nodes indicate the protein of unknown 3D structures, and filled nodes indicate that 3D structures are known or predicted. The connection between nodes indicates the interaction between proteins, and different colors correspond to different types of interaction

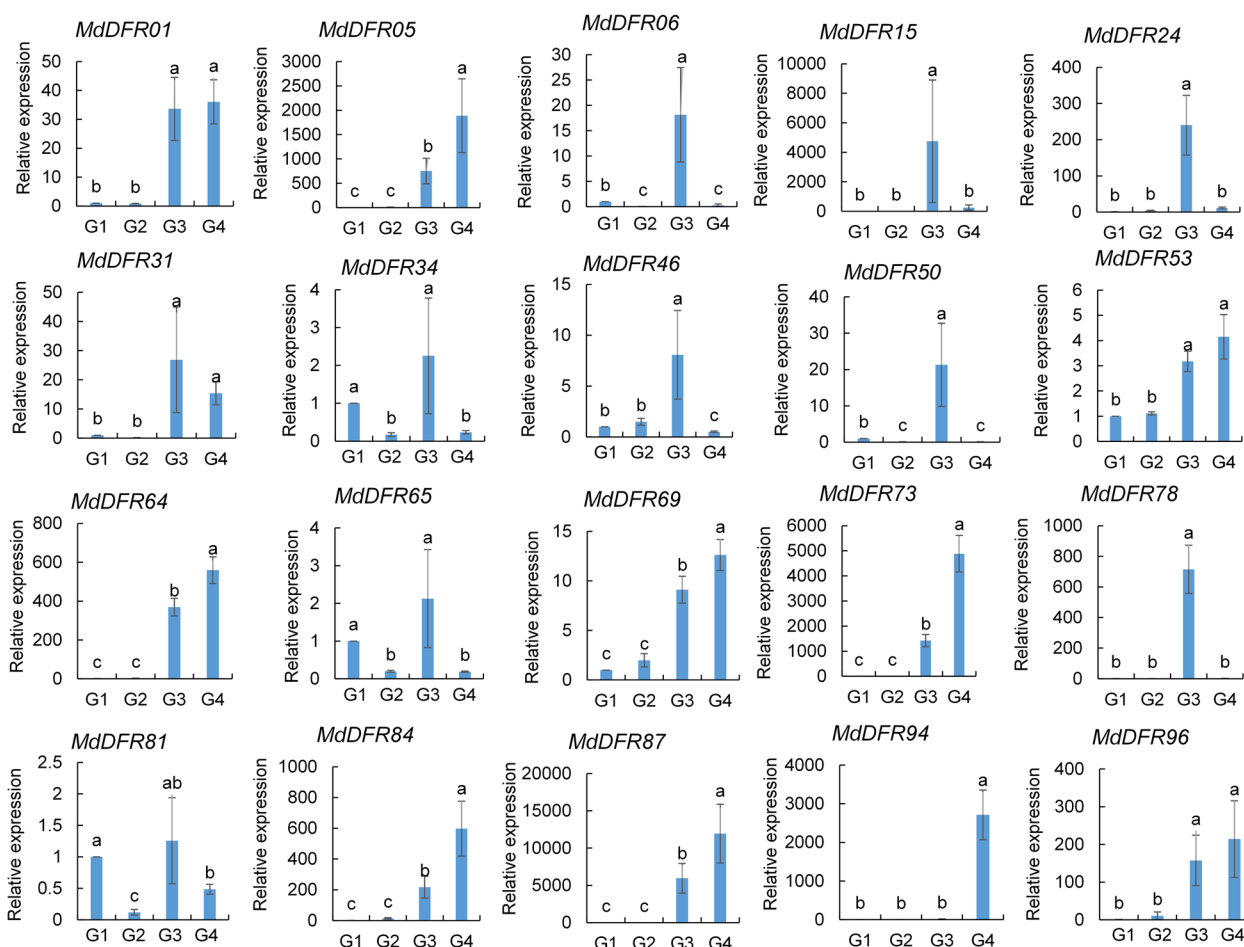


Fig. 8 Relative expression levels of *MdDFRs* in ‘Golden Delicious’ (G1), ‘RuiXue’ (G2), ‘Chengji No.1’ (G3) and ‘Astar’ (G4). G1 was used as the control. Error bars represent the mean \pm SE from three biological repeats. Different letters denote significant differences ($P < 0.05$)

31 *MdDFRs* were located in the cytoplasm (Table 1), which is consistent with the conclusion that anthocyanins are synthesized mainly in the cytoplasm [27]. Therefore, we speculate that *DFR* genes are involved in the synthesis of apple anthocyanins.

Gene duplication plays an important role in the evolution of organisms, including tandem duplication, local duplication and whole genome duplication [28]. Duplicate genes may be maintained through subfunctionalization and neofunctionalization at the level of expression or coding sequence [29]. Although 96 *DFR* genes were distributed on 17 apple chromosomes (Fig. 1, Table 1), chr2 and chr11 each had 12 genes, followed by chr7 with 9 genes, chr6 and chr8 with 8 genes each, which was speculated to be formed by tandem repeats. Internal collinearity analysis revealed a total of 34 *DFR* duplicate gene pairs, mostly located on chromosomes 2 and 3, with 8 and 9 pairs, respectively (Fig. 5a), suggesting that these genes may have similar functions. This suggests

that tandem duplication may be the main reason for the amplification of *DFR* family in apple.

The different colors of flowers and fruits of higher plants are closely related to plant pigments, which are anthocyanins and carotenoids. Anthocyanin transport and accumulation are regulated by many factors, mainly internal factors, but also by environmental factors. *DFR* is also regulated by environmental factors, and many plants also synthesize anthocyanins to protect the organism under environmental stress, which is accompanied by the increase of *DFR* expression in vivo [22]. Previous studies have shown that in most plants, anthocyanin synthesis requires light induction [30–32]. Under the effect of strong light, the expression of structural genes and regulatory genes will be enhanced [33]. Natural variations of *DFR* were analyzed using an expanded population of 282 accessions belonging to the spiny *Solanum* group, and the single-nucleotide polymorphisms were found at the MYB recognition site in the promoter region, which

causes differential expression of *DFR* and affects anthocyanin accumulation in fruits of the detected accessions [27]. In the present research, lignin/flavonoid biosynthesis-related cis-acting elements were present in almost all *MdDFR* promoter sequences (Fig. 4). The past decades have seen remarkable inroads made into our understanding of MYB proteins and their binding sites more specifically. The MBSI (aaaAaaC(G/C)GTTA) and MBSII (aaaAGTTAGTTA) were the binding sites of *Ph3* in *Petunia hybrid* and involved in flavonoid biosynthetic genes regulation [34]. MBSI element was identified in the promoter of *MdDFR15/76/81/89/90/91/93/94* in this study. MYB recognition site and MYB-binding site were both identified in the promoters of *MdDFR 05/09/13/19/22/24/26/30/31/33/34/46/50/52/54/64/65/69/75/76/79/86*. What's more, the expression level of *MdDFR54* and *MdDFR91* was higher in apple flowers, and *MdDFR64* was higher in ripe apple fruit peels and flowers, whereas lower in other tissues (Fig. 6). The expression levels of *MdDFR05/31/64/69/94* were up-regulated with the increase of fruit skin pigmentation verified by qRT-PCR (Fig. 8). Therefore, the position and number of cis-acting elements related to anthocyanin synthesis in *DFR* promoter region are the causes for the changes in apple anthocyanin accumulation.

The biosynthetic pathway of anthocyanins has been defined, but its regulatory network still needs to be perfected. MYB plays an important role in promoting plant growth and development, secondary metabolism regulation, hormone regulation, and stress response [35–38]. These transcription factors can recognize and bind to structural gene promoters to enhance or inhibit anthocyanin synthesis [39]. In tree peony (*Paeonia suffruticosa* Andr.), *PsMYB44* negatively regulates anthocyanin biosynthesis by directly binding to the *PsDFR* promoter and subsequently inhibiting blotch formation, which helps to elucidate the molecular regulatory network of anthocyanin-mediated blotch formation in plants [40]. The ectopic combined expression of transcription factors MYB, bHLH, and WDR in *Lotus japonicus* indicated that the *DFR2* promoter was activated [41]. *MdMYB1* appears to coordinately regulate structural genes in the anthocyanin pathway and segregates with red-skinned apples [42]. *MdbHLH33* can directly binds to the *MdMYBPA1* promoter and interacts with it to promote anthocyanin biosynthesis in response to low temperatures red-fleshed apples [12]. Recent study shows that mdm-miR858 targets transcription factor genes *MdMYB9* and *MdMYBPA1* to participate in anthocyanin accumulation in apple [43]. In addition, *F3H* and *ANS*, as upstream and downstream structural genes of *DFR*, respectively, are crucial in the phenylalanine metabolic pathway [6, 7]. In the present research, 75 *MdDFR* proteins interacted

directly or indirectly with anthocyanidin synthesis related proteins MdANS, MdF3H, MdMYB1 and MdMYBPA1 to form a protein interaction network (Fig. 7). Besides, *MdDFR69* and *MdDFR87* had direct interactions with these four proteins, *MdDFR64* had direct interactions with MdANS and MdF3H (Fig. 7), and all of these three genes were up-regulated with the increase of fruit skin pigmentation (Fig. 8), suggesting that the expression and regulation of these genes and transcription factors were coordinated with the production of anthocyanin. Therefore, further research on the transcription factors regulated by *DFR* upstream and downstream is of great significance to improve the regulatory network of anthocyanin synthesis.

Conclusions

In this study, a total of 96 members of the *DFR* genes were identified from the apple database and were mainly expressed in chloroplasts and cytoplasm. Lignin/flavonoid synthesis-related elements of MBSI, MYB recognition site and MYB-binding site were identified in the promoters of *MdDFR05/09/13/15/19/22/24/26/30/31/33/34/46/50/52/54/64/65/69/75/76/79/81/86/89/90/91/93/94*. There were 34 pairs of duplicated *MdDFR* gene pairs in apple, 66 and 57 homologous *DFR* gene pairs in apple/tomato and apple/grape, respectively. *MdDFR64* were directly interacted with MdANS and MdF3H, and the expression level of *MdDFR64* was higher in ripe apple fruit peels and flowers other than other tissues. Besides, *MdDFR69* and *MdDFR87* had direct interactions with anthocyanidin synthesis related proteins MdANS, MdF3H, MdMYB1, MdMYBPA1. The expression levels of *MdDFR64*, *MdDFR69* and *MdDFR87* were up-regulated with the accumulation of anthocyanins in apple verified by qRT-PCR.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40538-023-00470-z>.

Additional file 1: Table S1: Sequence of primers used for qRT-PCR analysis.

Acknowledgements

Not applicable.

Author contributions

BHC and WFL designed the experiments; WFL, JG, and ZHM performed the experiments; YJH, XL, MYC and JM performed data analysis; WFL wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by the 2023 Gansu Province University Teacher Innovation Fund Project (2023A-055), Double First-Class Major Scientific Research Project of Gansu Provincial Department of Education

(GSSYLXM-02), and the Science and Technology Major Project of Gansu Province (22ZD6NA045).

Availability of data and materials

The data that support the finding of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Horticulture, Gansu Agricultural University, Lanzhou 730070, China.

Received: 12 July 2023 Accepted: 4 September 2023

Published online: 25 September 2023

References

- Chai SC, Houshmand S, Saadat RL, Payton ME, Brummel-Smith K, Arjmandi BH. Daily apple versus dried plum: impact on cardiovascular disease risk factors in postmenopausal women. *J Acad Nutr Diet*. 2012;112(8):1158–68.
- Hair R, Sakaki JR, Chun OK. Anthocyanins, microbiome and health benefits in aging. *Molecules*. 2021;26(3):537.
- Wang R, Lu N, Liu C, Dixon RA, Wu Q, Mao Y, Yang Y, et al. *MtGSTF7*, a TT19-like GST gene, is essential for accumulation of anthocyanins, but not proanthocyanins in *Medicago truncatula*. *J Exp Bot*. 2022;73(12):4129–46.
- Martens S, Teeri T, Forkmann G. Heterologous expression of dihydroflavonol 4-reductases from various plants. *FEBS Lett*. 2002;531(3):453–8.
- Zhou H, Lin-Wang K, Wang F, Espley RV, Ren F, Zhao J, Ogutu C, et al. Activator-type R2R3-MYB genes induce a repressor-type R2R3-MYB gene to balance anthocyanin and proanthocyanidin accumulation. *New Phytol*. 2018;221(4):1919–34.
- Li WF, Ning GX, Mao J, Guo ZG, Zhou Q, Chen BH. Whole-genome DNA methylation patterns and complex associations with gene expression associated with anthocyanin biosynthesis in apple fruit skin. *Planta*. 2019;250:1833–47.
- Lim SH, You MK, Kim DH, Kim JK, Lee JY, Ha SH. RNAi-mediated suppression of dihydroflavonol 4-reductase in tobacco allows fine-tuning of flower color and flux through the flavonoid biosynthetic pathway. *Plant Physiol Bioch*. 2016;109:482–90.
- Li H, Qiu J, Chen F, Lv X, Fu C, Zhao D, Hua X, et al. Molecular characterization and expression analysis of dihydroflavonol 4-reductase (DFR) gene in *Saussurea medusa*. *Mol Biol Rep*. 2012;39:2991–9.
- Liu H, Lou Q, Ma J, Su B, Gao Z, Liu Y. Cloning and functional characterization of dihydroflavonol 4-reductase gene involved in anthocyanidin biosynthesis of grape hyacinth. *Int J Mol Sci*. 2019;20(19):4743.
- Albert NW, Davies KM, Lewis DH, Zhang H, Montefiori M, Brendolise C, Boase MR, et al. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell*. 2014;26:962–80.
- Yan S, Chen N, Huang Z, Li D, Zhi J, Yu B, Liu X, et al. *Anthocyanin Fruit* encodes an R2R3-MYB transcription factor, SIAN2-like, activating the transcription of *SIMYBATV* to fine-tune anthocyanin content in tomato fruit. *New Phytol*. 2019;225(5):2048–63.
- Wang N, Qu C, Jiang S, Chen Z, Xu H, Fang H, Su M, et al. The proanthocyanidin-specific transcription factor MdMYBPA1 initiates anthocyanin synthesis under low-temperature conditions in red-fleshed apples. *Plant J*. 2018;96(1):39–55.
- Hu Y, Cheng H, Zhang Y, Zhang J, Niu S, Wang X, Li W, et al. The MdMYB16/MdMYB1-miR7125-MdCCR module regulates the homeostasis between anthocyanin and lignin biosynthesis during light induction in apple. *New Phytol*. 2021;231:1105–22.
- Nakatsuka T, Saito M, Yamada E, Fujita K, Kakizaki Y, Nishihara M. Isolation and characterization of *GtMYBP3* and *GtMYBP4*, orthologues of R2R3-MYB transcription factors that regulate early flavonoid biosynthesis, in gentian flowers. *J Exp Bot*. 2012;63(18):6505–17.
- Peng Y, Thrimawithana AH, Cooney JM, Jensen DJ, Espley RV, Allan AC. The proanthocyanin-related transcription factors MYBC1 and WRKY44 regulate branch points in the kiwifruit anthocyanin pathway. *Sci Rep*. 2020;10:14161.
- Wang S, Zhang Z, Li LX, Wang HB, Zhou H, Chen XS, Feng SQ. Apple MdMYB306-like inhibits anthocyanin synthesis by directly interacting with MdMYB17 and MdbHLH33. *Plant J*. 2022;110(4):1021–34.
- Li WF, Mao J, Yang SJ, Guo ZG, Ma ZH, Dawuda MM, Zuo CW, et al. Anthocyanin accumulation correlates with hormones in the fruit skin of 'Red Delicious' and its four generation bud sport mutants. *BMC Plant Biol*. 2018;18:363.
- Tian J, Chen MC, Zhang J, Li KT, Song TT, Zhang X, Yao C. Characteristics of dihydroflavonol 4-reductase gene promoters from different leaf colored *Malus crabapple* cultivars. *Hortic Res*. 2017;4:17070.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*. 2001;25:402–8.
- O'Reilly C, Shepherd NS, Pereira A, et al. Molecular cloning of the *a1* locus of *Zea mays* using the transposable elements *En* and *Mu1*. *EMBO J*. 1985;4(4):877–82.
- Zhang P, Wen PF, Wen SB, Wang W, Pan QH, Zhan JC, et al. Molecular cloning of dihydroflavonol-4-reductase gene from grape berry and preparation of an anti-DFR polyclonal antibody. *Vitis*. 2008;3:47.
- Ahmed NU, Park JI, Jung HJ, Yang TJ, Hur Y, Nou IS. Characterization of dihydroflavonol 4-reductase (DFR) genes and their association with cold and freezing stress in *Brassica rapa*. *Gene*. 2014;550(1):46–55.
- Zhao A, Ding R, Wang C, Chen C, Wang D. Insights into the catalytic and regulatory mechanisms of dihydroflavonol 4-reductase, a key enzyme of anthocyanin synthesis in *Zanthoxylum bungeanum*. *Tree Physiol*. 2023;43(1):169–84.
- Shin DH, Choi MG, Kang CS, Park CS, Choi SB, Park YI. Overexpressing the wheat dihydroflavonol 4-reductase gene *TaDFR* increases anthocyanin accumulation in an *Arabidopsis dfr* mutant. *Genes Genomics*. 2016;38(4):333–40.
- Qin S, Liu Y, Cui B, Cheng J, Liu S, Liu H. Isolation and functional diversification of dihydroflavonol 4-reductase gene *HvDFR* from *Hosta ventricosa* indicate its role in driving anthocyanin accumulation. *Plant Signal Behav*. 2022;17(1):2010389.
- Sun W, Zhou N, Feng C, Sun S, Tang M, Tang X, Ju Z, et al. Functional analysis of a dihydroflavonol 4-reductase gene in *Ophiorrhiza japonica* (*OjDFR1*) reveals its role in the regulation of anthocyanin. *PeerJ*. 2021;20(9): e12323.
- Wang X, Chen X, Luo S, Ma W, Li N, Zhang W, Tikunov Y, et al. Discovery of a DFR gene that controls anthocyanin accumulation in the spiny Solanum group: roles of a natural promoter variant and alternative splicing. *Plant J*. 2022;111(4):1096–109.
- Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon-intron structure. *Proc Natl Acad Sci USA*. 2012;109:1187–92.
- Duarte JM, Cui L, Wall PK, Zhang Q, Zhang X, Leebens-Mack J, et al. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of *Arabidopsis*. *Mol Biol Evol*. 2006;23(2):469–78.
- Yang T, Ma H, Li Y, Zhang Y, Zhang J, Wu T, Song T, et al. Apple MPK4 mediates phosphorylation of MYB1 to enhance light-induced anthocyanin accumulation. *Plant J*. 2021;106(6):1728–45.
- Xing Y, Sun W, Sun Y, Li J, Zhang J, Wu T, Song T, et al. MPK6-mediated HYS phosphorylation regulates light-induced anthocyanin accumulation in apple fruit. *Plant Biotechnol J*. 2023;21(2):283–301.
- Zhao Y, Sun J, Cherono S, An JP, Allan AC, Han Y. Colorful hues: insight into the mechanisms of anthocyanin pigmentation in fruit. *Plant Physiol*. 2023;192(3):1718–32.

33. Cominelli E, Gusmaroli G, Allegra D, et al. Expression analysis of anthocyanin regulatory genes in response to different light qualities in *Arabidopsis thaliana*. *J Plant Physiol.* 2008;165(8):886–94.
34. Solano R, Nieto C, Paz-Ares JMYB. Ph3 transcription factor from *Petunia hybrida* induces similar DNA-bending/distortions on its two types of binding site. *Plant J.* 1995;8:673–82.
35. Abe H. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell.* 2003;15:63–78.
36. Song SS, Qi TC, Huang H, Ren QC, Wu DW, Chang CQ, Peng W, et al. The Jasmonate-ZIM domain proteins interact with the R2R3-MYB transcription factors MYB21 and MYB24 to affect Jasmonate-regulated stamen development in Arabidopsis. *Plant Cell.* 2011;23(3):1000–13.
37. Xie XB, Li S, Zhang RF, Zhao J, Chen YC, Zhao Q, Yao YX, et al. The bHLH transcription factor *MdbHLL3* promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. *Plant Cell Environ.* 2012;35(11):1884–97.
38. Vimolmangkang S, Han YP, Wei GC, Korban SS. An apple MYB transcription factor, MdMYB3, is involved in regulation of anthocyanin biosynthesis and flower development. *BMC Plant Biol.* 2013;13(1):176.
39. Yang YN, Yao GF, Zheng D, Zhang S, Wang C, Zhang M, Wu J. Expression differences of anthocyanin biosynthesis genes reveal regulation patterns for red pear coloration. *Plant Cell Rep.* 2014;34(2):189–98.
40. Luan Y, Chen Z, Tang Y, Sun J, Meng J, Tao J, Zhao D. Tree peony *PsMYB44* negatively regulates petal blotch distribution by inhibiting dihydroflavonol-4-reductase gene expression. *Ann Bot.* 2023;131(2):323–34.
41. Yoshida K, Iwasaka R, Shimada N, Ayabe S, Aoki T, Sakuta M. Transcriptional control of the dihydroflavonol 4-reductase multigene family in *Lotus japonicus*. *J Plant Res.* 2010;123:801–5.
42. Takos AM, Jaffé FW, Jacob SR, Bogs J, Robinson SP, Walker AR. Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol.* 2006;142:1216–32.
43. Li Z, Liu W, Chen Q, Zhang S, Mei Z, Yu L, Wang C, et al. Mdm-miR858 targets MdMYB9 and MdMYBPA1 to participate anthocyanin biosynthesis in red-fleshed apple. *Plant J.* 2023;113(6):1295–309.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
