Apiaceae *FNS I* originated from *F3H* through tandem gene duplication

Boas Pucker[1,2] and Massimo Iorizzo[3,4]

[1] Institute of Plant Biology, TU Braunschweig, Braunschweig, Germany

[2] BRICS, TU Braunschweig, Braunschweig, Germany

[3] Plants for Human Health Institute, NC State University, Kannapolis, North Carolina, USA

[4] Department of Horticultural Science, NC State University, Raleigh, North Carolina, USA

Corresponding authors: b.pucker@tu-braunschweig.de; miorizz@ncsu.edu

Abstract

Background: Flavonoids are specialized metabolites with numerous biological functions in stress response and reproduction of plants. Flavones are one subgroup that is produced by the flavone synthase (FNS). Two distinct enzyme families evolved that can catalyze the biosynthesis of flavones. While the membrane-bound FNS II is widely distributed in seed plants, one lineage of soluble FNS I appeared to be unique to Apiaceae species.

Results: We show through phylogenetic and comparative genomic analyses that Apiaceae *FNS I* evolved through tandem gene duplication of flavanone 3-hydroxylase (*F3H)* followed by neofunctionalization. Currently available datasets suggest that this event happened within the Apiaceae in a common ancestor of *Daucus carota* and *Apium graveolens*. The results also support previous findings that *FNS I* in the Apiaceae evolved independent of *FNS I* in other plant species.

Conclusion: We validated a long standing hypothesis about the evolution of Apiaceae FNS I and predicted the phylogenetic position of this event. Our results explain how an Apiaceae-specific *FNS I* lineage evolved and confirm independence from other *FNS I* lineages reported in non-Apiaceae species.

Keywords

Flavone synthase, flavanone 3-hydroxylase, flavonoid biosynthesis, carrot, comparative genomics, neofunctionalization, evolution, tandem gene duplication

Introduction

A plethora of specialized metabolites including flavonoids is produced by plants. These compounds provide an evolutionary advantage under certain environmental conditions. Flavonoids are produce in response to stresses like ultra violet (UV) radiation, cold, or drought [1, 2]. Especially visible is the pigmentation of flowers and fruits by anthocyanins which are one subclass of the flavonoids [3, 4]. Other subclasses include the proanthocyanidins which contribute to the pigmentation of seed coats [5] or flavonols which are produced in response to UV stress [6]. These branches of the flavonoid biosynthesis are well studied and conserved in many plant species and represent a model system for the investigation of the specialized metabolism in plants. A less conserved branch of the flavonoid biosynthesis leads to flavones (Fig. 1), which are important in signaling and defense against pathogens [7]. Flavones are derivatives of phenylalanine which is channeled through the general phenylpropanoid pathway to the chalcone synthase (CHS). This enzyme is the first committed step of the flavonoid biosynthesis. Chalcone isomerase (CHI) and flavone synthase (FNS) are the following steps involved in the formation of the flavone apigenin. F3'H can convert naringenin into eriodictyol which serves as substrate for the formation of the flavone luteolin. FNS activity evolved independently in different phylogenetic lineages [8]. Distributed over a wide phylogenetic range is FNS II, a membrane-bound cytochrome P450-dependent monooxygenase [9]. An independent lineage of FNS I, a soluble Fe^{2+}/2-oxoglutarate-dependent dioxygenase (2-ODD), was identified in the Apiaceae and appeared to be restricted to that family [10]. However, other studies report FNS I functionality in other plant species like OsFNSI in *Oryza sativa* [11], EaFNSI in *Equisetum arvense* [12], PaFNSI in *Plagiochasma appendiculatum* [13], AtDMR6 in *Arabidopsis thaliana* [14], and ZmFNSI-1 in *Zea mays* [14]. These lineages were presented as independent evolutionary events and are not orthologs of Apiaceae *FNS I* [8, 15]. Recently, a study revealed that FNS I is widely distributed in liverworts and is the most likely origin of seed plant flavanone 3-hydroxylase (F3H) [8]. Reports of enzymes with multiple functions like F3H/FNS I [8, 10] or F3H/FLS [16, 17] indicate that the 2-ODD family has a high potential for the acquisition of new functionalities.

Apiaceae FNS I shows high sequence similarity to F3H thus both were previously classified as DOXC28 in a systematic investigation of the 2-ODD family [18]. It was also hypothesized that Apiaceae *FNS I* evolved from *F3H* of seed plants by duplication and subsequent divergence [19, 20]. F3H and FNS I accept the same substrate (Fig. 1) which suggests that competition takes place if both enzymes are present in the same cell. The specific activity of both enzymes is defined by a small number of diagnostic amino acid residues [8, 10]. Substitution of these amino acids in F3H results in FNS I activity [10]. Exchange of three residues can already confer a partial function of the other enzyme. The authors hypothesize that a substitution of seven amino acid residues substantially modifies the pocket of the active site hence changing the orientation of the substrate. This is expected to cause a syn-elimination of hydrogen from carbon-2 (FNS activity) instead of hydroxylation of carbon-3 (F3H activity) [10].

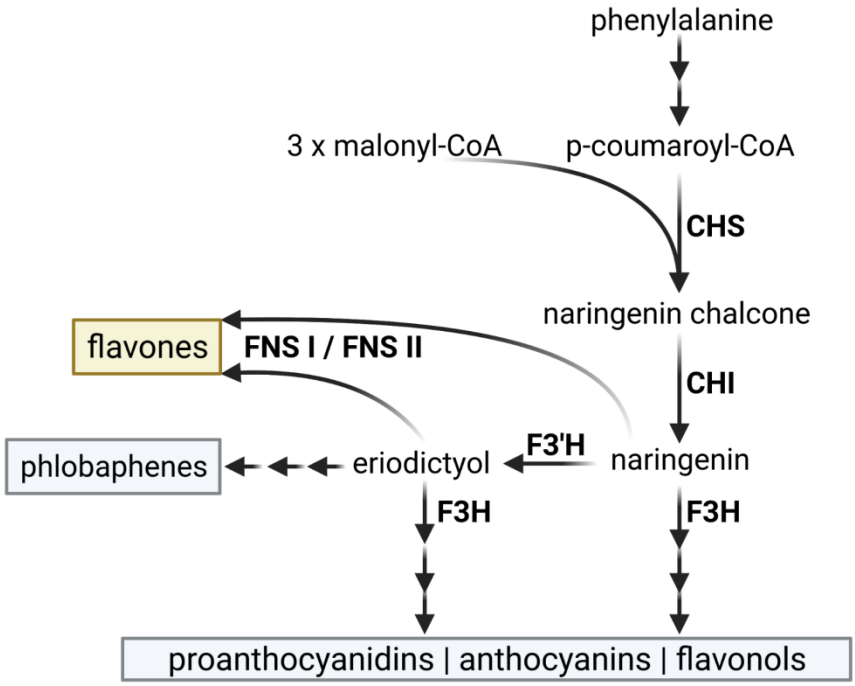

Fig. 1: Simplified illustration of the flavonoid biosynthesis with focus on the flavone biosynthesis. CHS (naringenin-chalcone synthase), CHI (chalcone isomerase), F3H (flavanone 3-hydroxylase), F3'H (flavonoid 3'-hydroxylase), and FNS (flavone synthase).

Although previous work hypothesized that Apiaceae FNS I originated from F3H through duplication and neofunctionlization, this hypothesis has not yet been validated. The recent release of high quality genome sequences representing most angiosperm lineages including members of the Apiacea family [21–24] opens the opportunity to address this hypothesis. Here, we investigated the evolution of FNS I in the Apiaceae through phylogenetic analysis and comparative genomics. The results indicate that *FNS I* originated from a tandem duplication of *F3H* that was followed by a neofunctionalization event.

Results

Apiaceae FNS I sequences show high sequence similarity to F3H that suggest a close phylogenetic relationship of both lineages. A phylogenetic tree was constructed to visualize this relationship. FNS I sequences of the non-Apiaceae species *Arabidopsis thaliana*, *Oryza sativa*, *Zea mays*, and *Parmotrema appendiculatum* were located outside the F3H clade in this tree. The FNS I sequences of seven Apiaceae species form a distinct clade (Fig. 2). This FNS I clade is embedded within a large clade of F3H sequences of a wide phylogenetic range of plants. The position of the FNS I sequences within the F3H clade suggests that Apiaceae FNS Is originated from F3H. The pattern also supports a single FNS I origin within the Apiaceae. The FNS I sequences of non-Apiaceae species seem to belong to an independent origin. The separation of the *FNS I*-like lineage from the *FNS I* lineage seems to predate a duplication in the *FNS I*-like lineage that produced the two copies discovered in *D. carota* and *A. graveolens*.

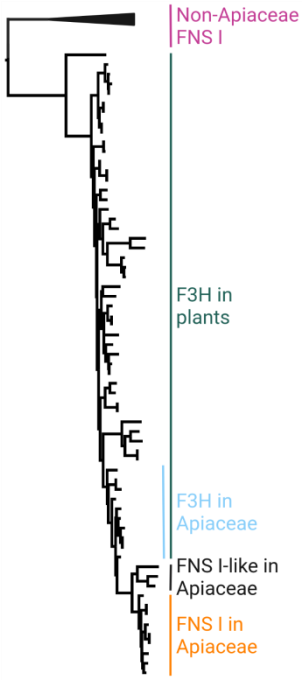

Fig. 2: Phylogenetic tree of F3H and FNS I sequences. Apiaceae FNS I sequences form a nested cluster within the F3H context. The FNS I sequences of non-Apiaceae species are placed outside the F3H clade. Polypeptide sequences used for the construction of this tree are included in Additional file 1. A full phylogenetic tree with sequence names is included as Additional file 2.

A previous study identified diagnostic amino acid residues that determine the FNS or F3H activity, respectively [10]. They demonstrated that a substitution of selected amino acid residues can convert one enzyme into the other. We inspected these characteristic features of the FNS I and F3H sequences of *Daucus carota* and *Apium graveolens* (Table 1). The results suggest that there is one *bona fide* F3H in *D. carota* (DCAR_009483) and *A. graveolens* (CM020901_g36676), respectively (Additional file 1). We also identified one FNS I in each of these species: DCAR_009489 and CM020901_g36861, respectively. In addition, there are FNS I-like copies which lack some of the functionally important amino acid residues of a *bona fide* FNS I (Table 1). Since the residues do not always match F3H, we cannot tell whether (1) these sequences have lost their FNS function in secondary events or (2) represent intermediates in the evolution from F3H towards FNS I. However, the incorporation of additional residues places these sequences in an intermediate clade (Fig. 2).

Table 1: Inspection of diagnostic amino acid residues in FNS I and F3H candidates of *Daucus carota* and *Apium graveolens*. FNS I residues are highlighted in orange, F3H residues are highlighted in skyblue. Positions are based on the FNS I of *Petroselium crispum* (AAP57393.1).

Sequence	Name	106	115	116	131	195	200	215	216
DCAR_009489	*DcFNS I*	T	T	I	F	E	I	V	R
DCAR_009487	*DcFNS I- like*	P	I	V	F	E	I	C	R
DCAR_009488	*DcFNS I-like*	T	T	V	F	E	I	V	R
DCAR_009483	*DcF3H*	M	I	V	I	D	V	L	K
CM020904_g37715	*AgFNS I-like*	T	T	I	F	K	I	C	R
CM020901_g36861	*AgFNS I*	T	T	I	F	E	I	V	R
CM020901_g36676	*AgF3H*	M	I	V	I	D	V	L	K

To narrow down the origin of the Apiaceae *FNS I*, we compared highly contiguous genome sequences of Apiaceae and outgroup species. The Apiaceae members *Daucus carota* and *Apium graveolens* show microsynteny in a region that harbours both, *F3H* and *FNS I* genes (Fig. 3). Both species differ from the *Centella asiatica* (basal Apiaceae species) and *Panax ginseng* (outgroup species) which do not show a *FNS I* gene in this region or elsewhere in the genome sequence. However, the presence of *F3H* and flanking genes indicates that the correct region is analyzed.

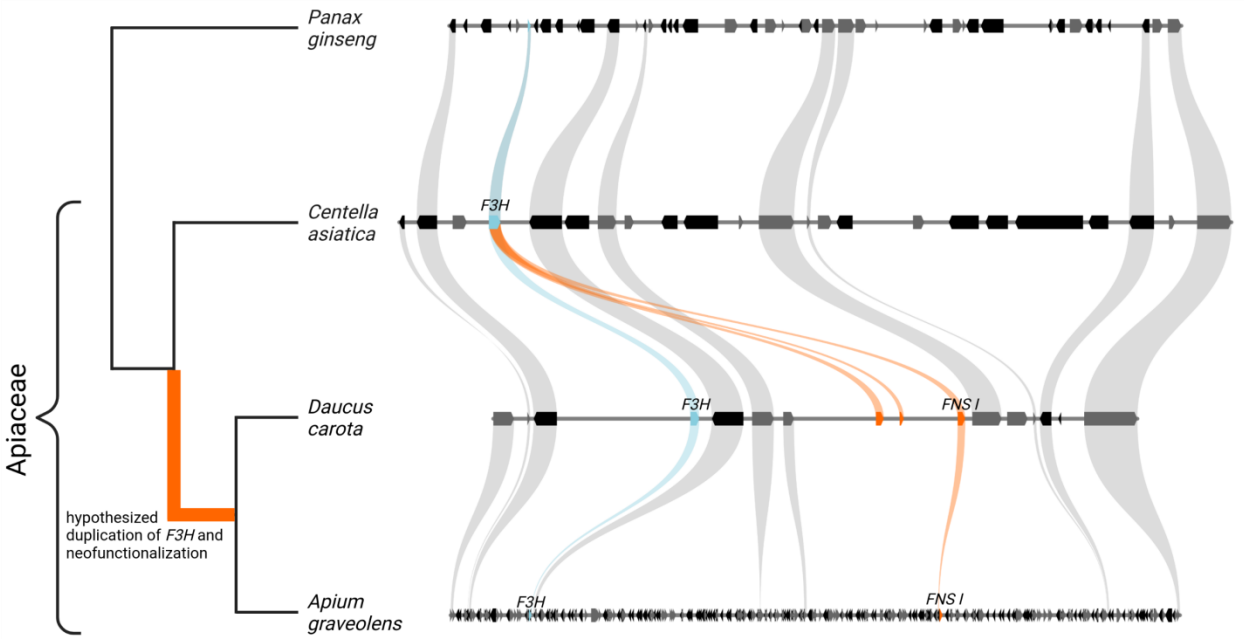

Fig. 3: Syntenic region between the Apiaceae species *Daucus carota* and *Apium graveolens* shows *F3H* (skyblue) and *FNS I* (orange) in close proximity, while *FNS I* was not observed in the basal Apiaceae species *Centella asiatica* or in the outgroup *Panax ginseng*.

Expression of the *F3H*, *FNS I*, and *FNS I*-like genes in carrots was analyzed across 146 RNA-seq samples (Fig. 4). The results show that *F3H* and *FNS I* show substantially higher expression than one of the *FNS I*-like genes (DCAR_009487) while the other *FNS I*-like gene (DCAR_009488) is almost not expressed. Strongest expression of DCAR_009487 was observed in the phloem and xylem of the root and in the petiole. DCAR_009488 showed the highest expression in whole flowers and stressed leaves of orange cultivars (Additional file 3).

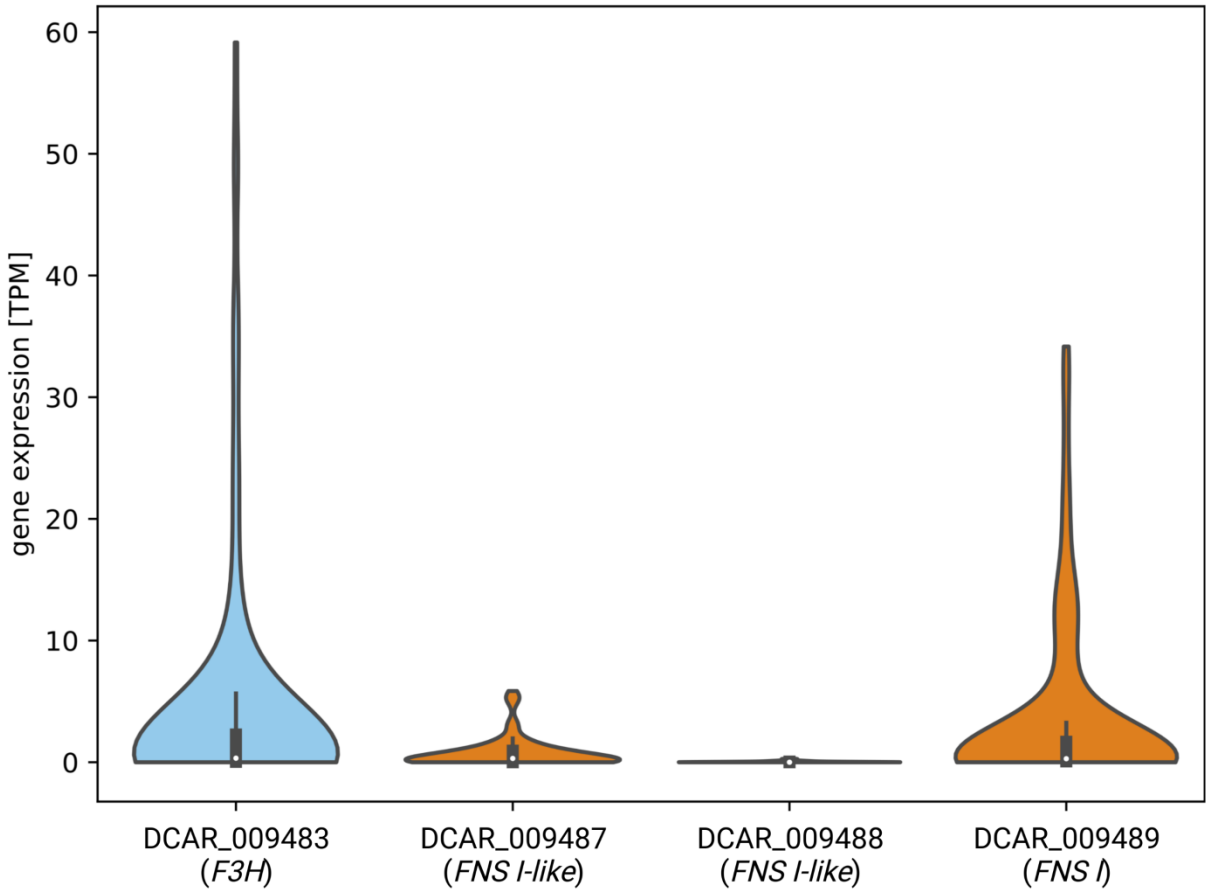

Fig. 4: Expression of *F3H*, *FNS I*, and *FNS I*-like genes in carrots. These plots show the distribution of transcript per million (TPM) values across various RNA-seq data sets derived from different tissues/organs and conditions (Additional file 3).

Discussion

We provide genomic and phylogenetic evidence for the evolution of the Apiaceae *FNS I* from *F3H* through tandem duplication followed by neofunctionalization. These results confirm a hypothesis about the evolution of *FNS I* from *F3H* [19, 20] and narrow down the initial duplication event. We show that the F3H duplication took place in a shared ancestor of *D. carota* and *A. graveolens*. Since there is no evidence for this gene duplication in *C. asiatica* which branches early in the Apiaceae, we hypothesize that the F3H duplication took place after the separation of *C. asiatica* (Fig. 3). Additional genome sequences will help to support this hypothesis and to narrow down the precise duplication event on the phylogenetic tree.

The inspection of conserved amino acid residues in *D. carota* and *A. graveolens* candidate sequences confirmed the presence of one F3H and FNS I in each species. Additionally, both species have at least one sequence that lacks some of the functionally important amino acid residues of a *bona fide* FNS I

without having all residues of a F3H (Table 1). This might indicate a different enzymatic activity or promiscuity of these enzymes. According to substitution experiment [10], the presence of at least T106, F131, and E195 in DCAR_009488 indicates that this enzyme has at least some basal FNS activity. It is possible that FNS I-like enzymes have multiple activities [8]. Only based on the diagnostic amino acid residues, we cannot tell whether (1) these sequences have lost their FNS function in secondary events or (2) represent intermediates in the evolution from F3H towards FNS I. However, the incorporation of additional residues places these sequences in an intermediate clade (Fig. 2). Based on their phylogenetic relationship, we hypothesize that two *FNS I* copies were present in the common ancestor of *D. carota* and *A. graveolens*. One of these copies was again duplicated in *D. carota* after separation of the lineages leading to *D. carota* and *A. graveolens* hence explaining the presence of three copies in *D. carota*. The preservation of these sequences since the separation of both species indicates a relevance of these FNS I-like sequences. The expression analysis suggests that these genes are active in specific tissues.

The physical clustering of *FNS I* and *F3H* in the genome could be due to the recent tandem duplication and a lack of time that would be necessary for dispersal of the cluster. However, it could be interesting to investigate whether this clustering does also provide an evolutionary benefit. Biosynthetic gene clusters (BGCs) were previously described in numerous plant species [25, 26]. These BGCs are often associated with an evolutionary young trait that provides a particular advantage e.g. in the defense against a pathogen [25]. Given the relevance of flavones in the defense against pathogens [7, 27], it seems possible that the flavone biosynthesis could be a similar trait that evolved in the Apiaceae. Nevertheless, a large number of additional functions of flavones [7] do not allow a definite answer to this question yet.

FNS I genes were also discovered in a small number of non-Apiaceae species [11, 13, 14]. However, these genes belong to an independent *FNS I* lineage [8]. As more high quality genome sequences of seed plants are released, a systematic search for additional non-Apiaceae *FNS I* sequences could become feasible in the near future. The number of independent FNS I origins remains unknown. Exploration and comparison of additional *FNS I* lineages across plants has the potential to advance our understanding of enzyme evolution.

Conclusions

A tandem gene duplication of *F3H* followed by neofunctionalization resulted in the evolution of *FNS I* in Apiaceae. This origin appears independent of FNS I in *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays*.

Methods

Datasets

The genome sequences and the corresponding annotation of *Daucus carota* 388_v2.0 [21] and *Panax ginseng* GCA_020205605.1 [22] were retrieved from Phytozome [28]. The genome sequences of *Apium graveolens* GCA_009905375.1 [23] and *Centella asiatica* GCA_014636745.1 [24] were downloaded from

the NCBI. Sequences of F3H and FNS I were retrieved from KIPEs [29] and are included in Additional file 1. The phylogenetic relationships of Apiaceae species were inferred from a previously constructed species tree [30].

Gene prediction

Since no complete annotation of the coding sequences was publicly available for *Apium graveolens* and *Centella asiatica*, we applied AUGUSTUS v3.3 [31] for an *ab initio* gene prediction with previously described settings [32]. The *Daucus carota* annotation of *F3H* and *FNS I* was manually checked in the Integrated Genome Viewer [33] and revised (Additional file 1). This was based on TBLASTN v2.8.1 [34] results of the *Petroselinum crispum* FNS I sequence against the *D. carota* genome sequence. Additionally, RNA-seq reads were retrieved from the Sequence Read Archive (Additional file 3) and aligned to the *D. carota* genome sequence using STAR v2.7.3a [35] with previously described parameters [36].

Phylogenetic tree construction

F3H and FNS I sequence collections [29] were used to search for additional candidates in *C. asiatica*, *D. carota*, and *A. graveolens* using a BLAST-based Python script [37]. Initial candidates were validated through a phylogenetic tree constructed with FastTree 2 [38]. A final tree was constructed with RAxML-NG [39] using LG+G8+F and 100 rounds of bootstrapping.

Synteny analysis

JCVI [40] was applied to compare the genome sequences of *P. ginseng*, *C. asiatica*, *D. carota*, and *A. graveolens*. The region around *F3H* and *FNS I* was manually selected. Connections of genes between the species were manually validated and revised based on phylogenetic trees (Additional file 2). TBLASTN was run with the *P. crispum* FNS I against the genome sequence of *C. asiatica* and *A. graveolens* to identify gene copies that might be missing in the annotation. The results of this search were compared against the annotation to find hits outside of annotated genes [37]. The best hits were assessed in a phylogenetic tree with previously characterized F3H and FNS I sequences.

Gene expression analysis

Paired-end RNA-seq data sets were retrieved from the Sequence Read Archive via fastq-dump [41] (Additional file 3). kallisto v0.44 [42] was applied with default parameters for the quantification of gene expression (counts and TPMs). A Python script was developed for the generation of violin plots to show the variation of gene expression (TPMs) across various samples [37]. Outliers were suppressed in this visualization. They are defined as data points which are more than three interquartile ranges away from the median.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All datasets underlying this study are publicly available or included within the additional files. Scripts developed for this work are freely available on github: https://github.com/bpucker/ApiaceaeFNS1.

Competing interests

The authors declare that they have no competing interests.

Funding

MI was supported by the United States Department of Agriculture National Institute of Food and Agriculture, Hatch project 1008691. We acknowledge support by the Open Access Publication Funds of Technische Universität Braunschweig.

Authors' contribution

BP and MI performed the analyses and wrote the manuscript.

Acknowledgements

Many thanks to the German network for bioinformatics infrastructure (de.NBI, grant 031A533A) and the Bioinformatics Resource Facility (BRF) at the Center for Biotechnology (CeBiTec) at Bielefeld University for providing an environment to perform the computational analyses. We used bioRender.com for the construction of some figures. We thank Hanna Marie Schilbert for discussion and comments on the manuscript.

Additional files

Additional file 1. Collection of F3H, FNS I, and FNS I-like sequences that were used for the analyses of this study.

Additional file 2. Phylogenetic tree of F3H, FNS I, and FNS I-like sequences.

Additional file 3. Gene expression values (TPMs) of *Daucus carota F3H*, *FNS I*, and *FNS I-like* genes.

References

1. Winkel-Shirley B. Biosynthesis of flavonoids and effects of stress. Current Opinion in Plant Biology. 2002;5:218–23.

2. Nakabayashi R, Mori T, Saito K. Alternation of flavonoid accumulation under drought stress in Arabidopsis thaliana. Plant Signal Behav. 2014;9:e29518.

3. Winkel-Shirley B. Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. Plant Physiology. 2001;126:485–93.

4. Grotewold E. The genetics and biochemistry of floral pigments. Annu Rev Plant Biol. 2006;57:761–80.

5. Todd JJ, Vodkin LO. Pigmented Soybean (Glycine max) Seed Coats Accumulate Proanthocyanidins during Development. Plant Physiology. 1993;102:663–70.

6. Emiliani J, Grotewold E, Falcone Ferreyra ML, Casati P. Flavonols protect Arabidopsis plants against UV-B deleterious effects. Mol Plant. 2013;6:1376–9.

7. Jiang N, Doseff AI, Grotewold E. Flavones: From Biosynthesis to Health Benefits. Plants. 2016;5:27.

8. Li D-D, Ni R, Wang P-P, Zhang X-S, Wang P-Y, Zhu T-T, et al. Molecular Basis for Chemical Evolution of Flavones to Flavonols and Anthocyanins in Land Plants. Plant Physiol. 2020;184:1731–43.

9. Martens S, Forkmann G. Cloning and expression of flavone synthase II from Gerbera hybrids. The Plant Journal. 1999;20:611–8.

10. Gebhardt YH, Witte S, Steuber H, Matern U, Martens S. Evolution of Flavone Synthase I from Parsley Flavanone 3β-Hydroxylase by Site-Directed Mutagenesis. Plant Physiol. 2007;144:1442–54.

11. Lee YJ, Kim JH, Kim BG, Lim Y, Ahn J-H. Characterization of flavone synthase I from rice. BMB Rep. 2008;41:68–71.

12. Bredebach M, Matern U, Martens S. Three 2-oxoglutarate-dependent dioxygenase activities of Equisetum arvense L. forming flavone and flavonol from (2S)-naringenin. Phytochemistry. 2011;72:557–63.

13. Han X-J, Wu Y-F, Gao S, Yu H-N, Xu R-X, Lou H-X, et al. Functional characterization of a Plagiochasma appendiculatum flavone synthase I showing flavanone 2-hydroxylase activity. FEBS Lett. 2014;588:2307–14.

14. Falcone Ferreyra ML, Emiliani J, Rodriguez EJ, Campos-Bermudez VA, Grotewold E, Casati P. The Identification of Maize and Arabidopsis Type I FLAVONE SYNTHASEs Links Flavones with Hormones and Biotic Interactions. Plant Physiol. 2015;169:1090–107.

15. Wang Q-Z, Downie SR, Chen Z-X. Genome-wide searches and molecular analyses highlight the unique evolutionary path of flavone synthase I (FNSI) in Apiaceae. Genome. 2018;61:103–9.

16. Prescott AG, Stamford NPJ, Wheeler G, Firmin JL. In vitro properties of a recombinant flavonol synthase from Arabidopsis thaliana. Phytochemistry. 2002;60:589–93.

17. 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.

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

19. Martens S, Forkmann G, Britsch L, Wellmann F, Matern U, Lukačin R. Divergent evolution of flavonoid 2-oxoglutarate-dependent dioxygenases in parsley 1. FEBS Letters. 2003;544:93–8.

20. Gebhardt Y, Witte S, Forkmann G, Lukacin R, Matern U, Martens S. Molecular evolution of flavonoid dioxygenases in the family Apiaceae. Phytochemistry. 2005;66:1273–84.

21. Iorizzo M, Ellison S, Senalik D, Zeng P, Satapoomin P, Huang J, et al. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. Nat Genet. 2016;48:657–66.

22. Jiang Z, Tu L, Yang W, Zhang Y, Hu T, Ma B, et al. The chromosome-level reference genome assembly for Panax notoginseng and insights into ginsenoside biosynthesis. Plant Communications. 2021;2:100113.

23. Li M-Y, Feng K, Hou X-L, Jiang Q, Xu Z-S, Wang G-L, et al. The genome sequence of celery (Apium graveolens L.), an important leaf vegetable crop rich in apigenin in the Apiaceae family. Hortic Res. 2020;7:1–10.

24. Pootakham W, Naktang C, Kongkachana W, Sonthirod C, Yoocha T, Sangsrakru D, et al. De novo chromosome-level assembly of the Centella asiatica genome. Genomics. 2021;113:2221–8.

25. Polturak G, Osbourn A. The emerging role of biosynthetic gene clusters in plant defense and plant interactions. PLOS Pathogens. 2021;17:e1009698.

26. Polturak G, Liu Z, Osboum A. New and emerging concepts in the evolution and function of plant biosynthetic gene clusters. Current Opinion in Green and Sustainable Chemistry. 2022;33:100568.

27. Du Y, Chu H, Wang M, Chu IK, Lo C. Identification of flavone phytoalexins and a pathogen-inducible flavone synthase II gene (SbFNSII) in sorghum. J Exp Bot. 2010;61:983–94.

28. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2012;40 Database issue:D1178–86.

29. Pucker B, Reiher F, Schilbert HM. Automatic Identification of Players in the Flavonoid Biosynthesis with Application on the Biomedicinal Plant Croton tiglium. Plants. 2020;9:1103.

30. Downie SR, Katz-Downie DS, Watson MF. A phylogeny of the flowering plant family Apiaceae based on chloroplast DNA rpl16 and rpoC1 intron sequences: towards a suprageneric classification of subfamily Apioideae. Am J Bot. 2000;87:273–92.

31. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res. 2006;34 suppl_2:W435–9.

32. Pucker B, Holtgräwe D, Weisshaar B. Consideration of non-canonical splice sites improves gene prediction on the Arabidopsis thaliana Niederzenz-1 genome sequence. BMC Research Notes. 2017;10:667.

33. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. Nat Biotechnol. 2011;29:24–6.

325 34. Gertz EM, Yu Y-K, Agarwala R, Schäffer AA, Altschul SF. Composition-based statistics and translated
326 nucleotide searches: Improving the TBLASTN module of BLAST. BMC Biology. 2006;4:41.

327 35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq
328 aligner. Bioinformatics. 2013;29:15–21.

329 36. Haak M, Vinke S, Keller W, Droste J, Rückert C, Kalinowski J, et al. High Quality de Novo
330 Transcriptome Assembly of Croton tiglium. Front Mol Biosci. 2018;5.

331 37. Pucker B. Apiaceae FNS I. https://github.com/bpucker/ApiaceaeFNS1. Python. 2022.

332 38. Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for Large
333 Alignments. PLOS ONE. 2010;5:e9490.

334 39. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly
335 tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35:4453–5.

336 40. Tang H, Krishnakumar V, Li J. jcvi: JCVI utility libraries (v0.5.7). 2015.
337 https://doi.org/10.5281/zenodo.31631.

338 41. Leinonen R, Sugawara H, Shumway M, on behalf of the International Nucleotide Sequence Database
339 Collaboration. The Sequence Read Archive. Nucleic Acids Research. 2011;39 suppl_1:D19–21.

340 42. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat
341 Biotechnol. 2016;34:525–7.

342