

# The Genetic Control of Polyacetylenes Involved in Bitterness of Carrots (*Daucus Carota* L.): Identification of QTLs and Candidate Genes From the Plant Fatty Acid Metabolism

Frank Dunemann (✉ [frank.dunemann@julius-kuehn.de](mailto:frank.dunemann@julius-kuehn.de))

Julius Kühn-Institut

Wanying He

Julius Kühn-Institut

Christoph Böttcher

Julius Kühn-Institut

Sven Reichardt

Julius Kühn-Institut

Thomas Nothnagel

Julius Kühn-Institut

Paul Heuvelmans

Nunhems Netherlands BV

Freddy Hermans

Nunhems Netherlands BV


---

## Research Article

**Keywords:** *Daucus carota*, polyacetylene, falcarinol, falcarindiol, HPLC/DAD, QTL mapping, candidate gene, fatty acid desaturase, FAD2 gene family, CER1 decarbonylase gene

**Posted Date:** November 30th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1102862/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

## Background

Falcarinol-type polyacetylenes (PAs) such as falcarinol (FaOH) and falcarindiol (FaDOH) are produced by several Apiaceae vegetables such as carrot, parsnip, celeriac and parsley. They are known for numerous biological functions and contribute to the undesirable bitter off-taste of carrots and their products. Despite their interesting biological functions, the genetic basis of their structural diversity and function is widely unknown. A better understanding of the genetics of the PA levels present in carrot roots might support breeding of carrot cultivars with tailored PA levels for food production or nutraceuticals.

## Results

A large carrot  $F_2$  progeny derived from a cross of a cultivated inbred line with an inbred line derived from a *Daucus carota* ssp. *commutatus* accession rich in PAs was used for linkage mapping and quantitative trait locus (QTL) analysis. Ten QTLs for FaOH and FaDOH levels in roots were identified in the carrot genome. Major QTLs for FaOH and FaDOH with high LOD values of up to 40 were identified on chromosomes 4 and 9. To discover putative candidate genes from the plant fatty acid metabolism, we examined an extended version of the inventory of the carrot *FATTY ACID DESATURASE2* (*FAD2*) gene family. Additionally, we used the carrot genome sequence for a first inventory of *ECERIFERUM1* (*CER1*) genes possibly involved in PA biosynthesis. We identified genomic regions on different carrot chromosomes around the found QTLs, that contain several *FAD2* and *CER1* genes within their 2-LOD confidence intervals. With regard to the major QTLs on chromosome 9 three putative *CER1* decarboxylase gene models are proposed as candidate genes.

## Conclusion

The present study increases the current knowledge on the genetics of PA accumulation in carrot roots. Our finding, that carrot candidate genes from the fatty acid metabolism are significantly associated with major QTLs for both major PAs, will facilitate future functional gene studies and a further dissection of the genetic factors controlling PA accumulation. Characterization of such candidate genes will have a positive impact on carrot breeding programs aimed at both lowering or increasing PA concentrations in carrot roots.

## Background

Carrot is an economically important and world-wide grown root vegetable known as a multi-nutritional food source. The carrot root is one of the richest sources of provitamin A and contains several other bioactive natural products, which are recognized for their nutraceutical effects [1, 2]. Among the several chemical major classes the polyacetylenes (PAs) are a large group of phytochemicals, which are produced in higher plants of the families Apiaceae, Araliaceae and Asteraceae [3, 4]. Twelve different aliphatic  $C_{17}$ -PAs have been characterized in carrots [4], and Busta et al. [5] recently added two further compounds to this list. The major PAs of carrot and other Apiaceae species are falcarindiol (FaDOH) and its presumed precursor falcarinol (FaOH), previously also known as carotatoxin [6]. In carrot root periderm FaDOH is the quantitatively predominating PA [5, 7] but the concentrations of FaOH and FaDOH depend on the carrot cultivar [8]. In addition, roots of some wild carrot relatives such as the subspecies *D. c. maximus*, *D. c. maritimus*, or *D. c. halophilus* were shown to have up to 20 times higher PA concentrations than roots of cultivated carrots [9, 10]. Next to other phytochemicals such as 6-methoxymellein, laserin and epilaserin the presence of higher amounts of  $C_{17}$ -PAs has been shown to contribute to the undesirable bitter off-taste of carrots and their products [11, 12, 13]. The localization of PAs in exterior plant tissue layers such as the root periderm is also consistent with their assumed role in plant defense. Moreover, FaOH and FaDOH might action against some fungal carrot pathogens [14]. Accumulation of polyacetylenic phytoalexins, including FaOH, was observed in tomato fruits and leaves induced with *Cladosporium fulvum*, *Verticillium albo-atrum*, and *Fusarium oxysporum* [6]. In carrots, FaDOH could inhibit the *in vitro* development of the fungal leaf blight pathogen *Alternaria dauci*, and the levels increased in response to *A. dauci* inoculation [15]. With regard to putative effects of PAs on human health a wide range of biological activities have been reported for falcarinol-type PAs demonstrating that they exhibit potent anti-microbial, anti-inflammatory and anti-cancer effects [3, 4, 16, 17, 18, 19].

Despite their interesting biological functions, the biosynthesis of PAs and its genetic basis are poorly understood. In higher plants PAs are produced from unsaturated fatty acids such as oleic acid and linoleic acid. It has been hypothesized that a diverse pathway from linolenic acid to unusual fatty acids, such as crepenynic and dehydrocrepenynic acid, is the major route for the biosynthesis of falcarinol-type PAs [5, 6]. The enzyme primarily responsible for the synthesis of linoleic acid from oleic acid is a  $\Delta 12$ -fatty acid desaturase [20]. Divergent forms of Fatty Acid Desaturase 2 (*FAD2*) enzymes have diversified functionalities in fatty acid modification, e.g. catalyzing hydroxylation [21], epoxidation [22] and the formation of acetylenic bonds and conjugated double bonds [23]. The *Crep1* gene from *Crepis alpina* was the first cloned gene encoding a functional acetylenase. This enzyme accumulates large amounts of the acetylenic fatty acid crepenynic acid in the seeds of *C. alpina* [22].

*FAD2* is among the best-studied plant fatty acid desaturase gene families, in terms of both molecular and biochemical investigations. Since the cloning of a first plant *FAD2* gene in *A. thaliana* [24], *FAD2* genes have been isolated and characterized from many different plant species, including members of the *FAD2* gene family that possess diverse functional activities in fatty acid modification such as acetylenases. Only a single *FAD2* gene exists in *Arabidopsis*, but in most other plants multiple homologous *FAD2s* were found. Species with the largest known number of *FAD2s* are soybean (seven *FAD2* gene family members), cotton and tomato (9 genes), and safflower (11 genes) [25]. However, in the Apiaceae species *Petroselinum crispum* (parsley) 17 *FAD2* sequences were found [26], and in the carrot genome 24 *FAD2s* were previously reported [5]. This is the highest number of *FAD2* family members reported so far, indicating that functional divergence of *FAD2s* is not limited to oil seed crops [25].

The biochemical formation of FaDOH from its presumed precursor substance FaOH is largely unknown until now. Recently a pathogen-induced biosynthetic gene cluster was discovered in tomato and shown to be putatively involved in FaDOH production [27]. Among the four highly co-expressed clustered genes three genes were annotated as *FAD2* desaturases or acetylenases, and the fourth was shown to belong to five tomato homologs of *CER1 decarbonylase* [27]. In *Arabidopsis thaliana* *CER1 (ECERIFERUM1)* and *CER3 (ECERIFERUM3)* gene products catalyze fatty acid decarboxylation in the alkane biosynthetic pathway and play a major role in wax production [28]. Several orthologous genes of *CER1* and *CER3* have been identified in other plant species such as *Brassica napus*, *Camelina sativa*, and *Brachypodium distachyon* [29, 30, 31]. In carrot *CER1/CER3* family members have not been described yet.

Despite the extensive research concerning the biochemical identification, characterization and putative biological function, less knowledge exists about the genetic control of PA accumulation in carrot roots [4]. However, for breeding of carrot cultivars with a reduced bitter taste (i.e. for baby food or juice and pulp production), or with a higher pathogen resistance and enhanced health benefits more information about the inheritance of PA accumulation is needed. Genetic dissection by QTL analysis and the utilization of gene-specific (functional) molecular markers linked to QTLs for natural product biosynthesis could support such breeding strategies [32]. As a first attempt for carrot PAs, Le Clerc et al. [33] identified QTLs for the FaDOH levels in roots in some regions of a carrot linkage map and associated them with QTLs for bitterness and resistance to the fungus *Alternaria dauci*. However, no putative candidate genes were investigated in this study.

In this study, we performed a QTL analysis in a 400 plants large carrot F<sub>2</sub> family derived from a cross of inbred lines of a cultivated and a wild *Daucus* accession rich in PAs. Major QTLs were identified for both root levels of FaOH and FaDOH, enabling the detection of numerous genomic regions carrying QTLs and candidate genes from the plant fatty acid metabolism. To discover putative candidate genes, we examined and extended the *FAD2* gene family in carrot and, for the first time, we used the carrot genome sequence for an inventory of *CER1/3* genes putatively involved in the genetic control of PA biosynthesis.

## Results

### Polyacetylene phenotyping and QTL mapping

Two main C<sub>17</sub>-PAs FaOH and FaDOH were quantified in carrot root samples of 400 individual F<sub>2</sub> plants of the progeny 'CA' (P\_1870 x P\_345B). There was a large phenotypic variation among the F<sub>2</sub> genotypes, with a population mean of 197 µg/g dry weight (DW) for the FaOH level, and an about five-fold higher FaDOH content (1038 µg/g DW) (Fig. 1). We observed few individuals with less than 30 µg/g DW FaOH and three plants with a FaOH content above 1000 µg/g. The FaDOH contents varied between 200 µg/g and 3000 µg/g in the analyzed plants (Additional file 1: Table S1). Since a major prerequisite for bi-parental QTL mapping is a cross between contrasting parental genotypes, we performed a pre-test of the parents. In average, the plants of the *D. c. commutatus*-derived parent line (P\_1870) had a twice as high FaOH content compared to the parent plants of the cultivated carrot line (P\_345B), and the FaDOH content was about 3.5 times higher, respectively (Additional file 2: Fig. S1).

A total of 713 SNP markers were used to construct a 505 cM linkage map for the nine carrot chromosomes based on 374 F<sub>2</sub> plants of the CA progeny (Additional file 3: Fig. S2). The map constructed has an average map resolution of 1 cM with a maximum gap of 11.7 cM on chr\_8. With regard to the physical position of the markers, the chromosomes are well covered with the largest gap size of 1.48 Mbp. Severe segregation distortions were noted for chromosomes chr\_1, chr\_5 and chr\_9, all positive towards the P\_345B parent. To check if the segregation distortions were caused by bolting, the bolted individuals were genotyped for the distorted regions, and only chr\_1 showed a different allele distribution, showing a single incomplete recessive locus on chr\_1, between 16.6 - 25.6 Mbp (P\_1870 origin; data not shown).

QTL analysis mapped in total 10 QTLs for the concentrations of FaOH (6 QTLs) and FaDOH (4 QTLs), and two additional QTLs were identified for the concentration ratio of FaDOH and FaOH (Table 1). The six QTLs for the FaOH concentration were located on the chromosomes chr\_1, chr\_3, chr\_4, chr\_5, chr\_8 and chr\_9 (Fig. 2). For FaOH the QTL FaOH\_9.1 showed the highest LOD value (16.7) and explains 15.4% of the phenotypic variation. A second major QTL for FaOH with strong statistical support was found on chr\_4 (LOD 14.5). For FaDOH the four QTLs were identified on chr\_1, chr\_4, chr\_5 and chr\_9 (Fig. 2). As shown by the QTL peak positions (Table 1), on chr\_9 a strong QTL (LOD 40.4) explaining 24.6% of the variation was located at the same genomic position as the QTL for FaOH (*FaOH\_9.1*). With regard to the QTLs on chr\_4 the peak of QTL FaDOH\_4.1 (LOD 21.8) showed a distance of about 4.5 Mbp to QTL *FaOH\_4.1*, whereas the two QTL confidence intervals were in a similar magnitude and located in overlapping genomic regions. Analysis of QTL peak marker - trait associations showed that genotypes homozygous for the "wild" allele 'B' as well as the heterozygote genotypes showed significant higher PA contents (Additional file 4: Fig. S3). A major QTL (LOD 17.3) for the FaDOH/FaOH ratio (*Ratio\_3.1*) was identified at the end of chr\_3, and a second QTL for this parameter was located at the same position on chr\_4 as the QTL *FaOH\_8.1* (Table 1, Fig. 2).

Table 1  
QTLs for FaOH and FaDOH concentrations measured in roots of carrot F<sub>2</sub> progeny CA and parameters of 2-LOD confidence intervals

Substance	QTL	Chrom.	LOD	% Var.	Genetic 2-LOD interval (cM)			Physical 2-LOD interval (Mbp)			Markers 2-LOD Interval		
					Start	Peak	Stop	Start	Peak	Stop	mi start	mi peak	mi stop
FaOH	<i>FaOH_1.1</i>	1	4,34	5,57	4,8	22,4	53,6	824774	5,406415	47,969302	K0514	K4057	K1504
FaOH	<i>FaOH_3.1</i>	3	8,99	9,93	41,2	42,8	57,5	43,925869	45,134655	49,341067	K1064	K1792	K0474
FaOH	<i>FaOH_4.1</i>	4	14,48	8,41	32,9	40,4	47,8	13,388045	20.392970	33,347895	K3873	K3654	K1382
FaOH	<i>FaOH_5.1</i>	5	3,64	6,88	17,8	33,3	53,6	5,240197	22,019928	41,938893	K2562	K1670	mCA435
FaOH	<i>FaOH_8.1</i>	8	3,89	6,18	8,9	21,5	51,2	2,009986	5,214614	31,631123	K2093	K1762	K1770
FaOH	<i>FaOH_9.1</i>	9	16,71	15,35	13,0	16,1	21,8	4,342239	6,376534	11,832494	K0850	K3482	K3568
FaDOH	<i>FaDOH_1.1</i>	1	4,45	7,28	4,8	31,9	53,6	824774	21,266018	47,969302	K0514	K4090	K1504
FaDOH	<i>FaDOH_4.1</i>	4	21,81	8,51	32,9	35,5	48,4	13,388045	15,954204	33,764607	K3873	K1837	K0154
FaDOH	<i>FaDOH_5.1</i>	5	7,06	8,30	35,7	39,4	46,0	25,598051	33,366279	39,470106	K0871	K3010	K1404
FaDOH	<i>FaDOH_9.1</i>	9	40,35	24,55	13,0	16,1	18,9	4,342239	6,376534	8,494898	K0850	K3482	K3359
FaDOH/FaOH	<i>Ratio_3.1</i>	3	17,27	18,29	40,2	42,8	50,7	43,339133	45,134655	47,622391	K0980	K1792	K1640
FaDOH/FaOH	<i>Ratio_8.1</i>	8	5,00	5,11	8,9	21,5	51,2	2,009986	5,214614	31,631123	K2093	K1762	K1770

\*Alleles of QTL effects: A - Parent P\_345B; B - Parent P\_1870

#### Identification of FAD2 and CER1/3 genes in the carrot genome

We performed a reannotation of the whole carrot genome sequence vers.2 [34] to explore if there are carrot FAD2 genes which have not yet been identified in previous studies. In addition to the FAD2 inventory of [5], who described 24 members of the carrot FAD2 family, we detected seven undescribed putative FAD2 gene models, which are not annotated in the carrot genome, yet (Additional file 5: Data S1). They extend the total number of *Daucus* FAD2s to 31. The total inventory of carrot FAD2s including the gene names used in this study is shown in Additional file 6: Table S2. The newly predicted FAD2 genes *DcFAD2-29*, *DcFAD2-30*, and *DcFAD2-31* are located as single genes on chromosomes chr\_1, chr\_3, and chr\_4, respectively. The other four genes (*DcFAD2-25* to *DcFAD2-28*) are clustered together with the already known *DcFAD2-4* (DCAR\_027655) on chr\_8 (Additional file 6: Table S2). This five-gene cluster spans a genomic region of 12.4 kb. The seven newly identified FAD2 genes showed characteristic FAD2 sequence motifs, notably the three highly conserved histidine-rich motifs (Histidine box) which are part of the active sites of the enzyme [20]. Based on alignment of their deduced protein sequences the consensus sequences for the three His boxes in the seven genes are H-E-C-X-H, H-X-R-H-H, and H-V-X-H-H and correspond to the His boxes of the other carrot FAD2s (Additional file 7: Fig. S4). Also, the conserved aromatic amino acid-rich motif Y-X-X-K/R/D/E-L/F/M/I at the C-terminus is present with one exception (*DcFAD2-29*), where the amino acid P is present at the fifth position of the motif. This motif was reported to be involved in the localization of FAD2 enzymes in the ER [35] and was recently used to support categorization of tomato FAD2s [25]. Therefore, based on sequence information, we suppose that the newly found genes are members of the carrot FAD2 family. A phylogenetic analysis of the carrot FAD2s showed, that *DcFAD2-25*, *DcFAD2-26*, *DcFAD2-27* and *DcFAD2-28* are located in clade I (Additional file 8: Fig. S5) containing 10 known *Daucus* FAD2s which have been classified as divergent FAD2s [5]. All gene predictions from clade I show the amino acid G immediately preceding the first His box whereas the genes of clade II possess the amino acid A at this position and likely represent canonical FAD2 desaturases. It has been proposed that the amino acid G might indicate the functionally divergent FAD2s like acetylenases [36]. Clade I contains the genes *DcFAD2-6* (DCAR\_017011), *DcFAD2-7* (DCAR\_013552), and *DcFAD2-8* (DCAR\_013548) which were functionally characterized as  $\Delta$ 12-acetylenases [5]. *DcFAD2-6* is highly similar (amino acid identity 96%) to the parsley gene *PcELI12* which is also known as an FAD2-derived acetylenase [36]. Hence, we assume that *DcFAD2-25*, *DcFAD2-26*, *DcFAD2-27*, and *DcFAD2-28* might have an acetylenase function too. The three genes *DcFAD2-29*, *DcFAD2-30*, and *DcFAD2-31* build a separate clade III (Additional file 8: Fig. S5) but appear to be also FAD2s since the homology with a predicted chloroplast  $\omega$ -6 fatty acid desaturase gene (*DcFAD6\_DCAR\_019387*) is very low (results not shown).

To compile a first inventory of carrot CER1/3 genes, we used the predicted tomato CER1 decarboxylase gene *Solyc12g100270* [27] for BLAST searches in the carrot genome and in a whole plant protein database (NCBI). Totally 12 putative carrot CER1/3s were found in the carrot genome. Eleven gene models are already annotated in the carrot genome [34], but seven of them appeared to show uncomplete transcript sequences and were complemented by manual comparisons with other plant CER1/3s and bioinformatically predicted carrot CER1/3 gene models. The putative CER1/3 genes are located either as single or unclustered genes on the carrot chromosomes chr\_3, chr\_4, and chr\_5, respectively, or as two clustered genes each on chr\_6, chr\_7, and chr\_9 (Table 2). The gene model *DcCER1-3* on chr\_6 has not been annotated yet in the actual carrot genome sequence (Additional file 5: Data S1). The predicted *Daucus* CER1/3 proteins were compared with other plant CER1 and CER3 proteins as for instance those of *A. thaliana* [28] or *C. sativa* [29], and a phylogenetic tree was constructed which showed two major clades consisting each of six carrot CER1 or CER3 genes, respectively (Fig. 3). Sequence alignments of the 12 carrot gene models and other plant CER1/3s showed that the three typical His boxes [37] are also present in the carrot genes (Additional file 9: Fig. S6). Two of the genes (*DcCER1-1*, *DcCER3-1*) might be truncated and functionally inactive. The highest amino acid sequence identity was 85.5% for the genes *DcCER3-4* and *DcCER3-5* and 85.3% for *DcCER3-1* and *DcCER3-2*, whereas the maximum identity among the CER1 genes was 78.2%. The *DcCER1*s shared only 32 to 36% sequence identity with the *DcCER3* genes (Additional file 10: Fig. S7).

Table 2

List of *Daucus carota* *CER1* and *CER3* gene models sorted by their physical position on the assembled nine carrot chromosomes according to the whole genome sequence [34].

Chromosome <sup>1)</sup>	Gene Name	Genomic coordinates <sup>1)</sup>				Locus name <sup>1)</sup>	optimized	predicted	CDS	No. of	Protein
		Strand	Start	Stop			prediction	function	length	Introns	length
3	<i>DcCER3-6</i>	for	12173825	12178868	DCAR_009898	no	CER3	1899	10	633	
4	<i>DcCER1-4</i>	rev	8094127	8099851	DCAR_015766	yes	CER1	1884	8	628	
4	<i>DcCER3-5</i>	for	19421348	19424982	DCAR_014721	yes	CER3	1905	9	635	
5	<i>DcCER3-3</i>	for	2179381	2183098	DCAR_016321	yes	CER3	1914	9	638	
5	<i>DcCER3-4</i>	for	25089512	25093379	DCAR_017940	yes	CER3	1902	10	634	
6	<i>DcCER3-1</i>	rev	nd	23574832	DCAR_021381	yes	CER3	1713	nd	571	
6	<i>DcCER3-2</i>	rev	23591227	23597355	DCAR_021379	yes	CER3	1908	10	636	
7	<i>DcCER1-3</i>	rev	22041265	22045133	XP_017216154.1		CER1	1863	nd	621	
7	<i>DcCER1-5</i>	for	22046109	22048531	DCAR_025018	yes	CER1	1848	nd	616	
9	<i>DcCER1-2</i>	rev	5028865	5034295	DCAR_029345	no	CER1	1881	9	627	
9	<i>DcCER1-1</i>	rev	5038893	5042667	DCAR_029346	no	CER1	1719	7	573	
9	<i>DcCER1-6</i>	rev	8516093	8518841	DCAR_029533	no	CER1	1845	6	615	

<sup>1)</sup> Chromosomes, genomic coordinates, and locus names according the carrot whole genome sequence assembly vers.2 [34]; nd - not determined

## Association Of Qtls With Candidate Genes

Chromosomal start and end positions of the QTL intervals were utilized to compare the QTL regions detected with the chromosomal position of *FAD2* and *CER1/3* gene models described in the previous subsection. In total, 10 out of the 12 identified QTL regions comprise *FAD2* and/or *CER1/3* candidate genes (Table 3). The number of *FAD2* candidate genes per QTL region varies between one and nine, and the number of *CER1/3* candidate genes varies from one to three, respectively. Only for two QTLs on chr\_1 and chr\_3 no candidate genes belonging to the analyzed gene families could be found. With regard to the major QTL for FaDOH on chr\_9 (QTL peak at position 6.376534 Mbp) two putative *CER1* gene models (*DcCER1-1*, *DcCER1-2*) are located within the QTL interval which spans the region 4.342239 - 8.494898 Mbp. A third *CER1* gene (*DcCER1-6*) at position 8.516093 (Table 2) is located just outside the 2-LOD confidence interval. However, the support interval of the QTL for FaOH on chr\_9 (4.342239 - 11.832494 Mbp) contains all three *CER1* genes. The physical QTL peak position chr\_9 is about 1.4 Mbp distant from the two genes *DcCER1-1* (DCAR\_029346) and *DcCER1-2* (DCAR\_029345) and a fourth putative *CER1* gene (DCAR\_029349) which is not complete. Therefore, this gene fragment (called as *DcCER1-7*) has not been considered as candidate gene and was not included in this study. *DcCER1-1* amino acid sequence showed a deletion of 36 amino acids (Additional file 9: Fig. S6), which was confirmed by amplicon sequencing (results not shown). The support intervals for the QTLs for FaOH and FaDOH on chr\_4 covered a comparatively large genomic region and contain the six-gene *FAD2* cluster, starting at position 29.308974 (Additional file 6: Table S2). Furthermore, close to this position there is the candidate gene *DcCER3-5* at position 19.421348, too (Table 2). The largest *FAD2* gene cluster on chromosome chr\_8 consists of nine *FAD2* gene models. It is associated with minor QTLs for FaOH and QTLs for the ratio FaOH/FaDOH.

Table 3  
Putative *FAD2* and *CER1/3* candidate genes located in QTL intervals. For positions of QTL intervals and candidate genes, see Table 1 (QTL intervals), Table S2 (*FAD2*), and Table 2 (*CER1/3*)

QTL	Chrom.	<i>FAD2</i> candidate genes	<i>CER</i> candidate genes
<i>FaOH_1.1</i>	1	<i>DcFAD2-3, DcFAD2-18, DcFAD2-29</i>	
<i>FaOH_4.1</i>	4	<i>DcFAD2-7, DcFAD2-8, DcFAD2-16,</i> <i>DcFAD2-17, DcFAD2-19, DcFAD2-22</i>	<i>DcCER3-5</i>
<i>FaOH_5.1</i>	5	<i>DcFAD2-24</i>	<i>DcCER3-4</i>
<i>FaOH_8.1</i>	8	<i>DcFAD2-1, DcFAD2-2, DcFAD2-4,</i> <i>DcFAD2-9, DcFAD2-21, DcFAD2-25,</i> <i>DcFAD2-26, DcFAD2-27, DcFAD2-28</i>	
<i>FaOH_9.1</i>	9		<i>DcCER1-1, DcCER1-2, DcCER1-6</i>
<i>FaDOH_1.1</i>	1	<i>DcFAD2-3, DcFAD2-18, DcFAD2-29</i>	
<i>FaDOH_4.1</i>	4	<i>DcFAD2-7, DcFAD2-8, DcFAD2-16,</i> <i>DcFAD2-17, DcFAD2-19, DcFAD2-22</i>	<i>DcCER3-5</i>
<i>FaDOH_5.1</i>	5	<i>DcFAD2-11, DcFAD2-24</i>	
<i>FaDOH_9.1</i>	9		<i>DcCER1-1, DcCER1-2</i>
<i>Ratio_8.1</i>	8	<i>DcFAD2-1, DcFAD2-2, DcFAD2-4,</i> <i>DcFAD2-9, DcFAD2-21, DcFAD2-25,</i> <i>DcFAD2-26, DcFAD2-27, DcFAD2-28</i>	

### Expression analysis of *FAD2* and *CER1/3* candidate genes

To study the tissue-specific expression patterns of candidate genes, expression levels of 18 genes (14 *FAD2* and 4 *CER1/3* genes) were analyzed in four different tissues of three individual plants from each of the carrot cultivars Anthonina (AN), Breeding line (BRL) and Presto (PR) by qRT-PCR. Based on their location within the 2-LOD confidence intervals of strong QTLs (LOD > 10) the candidate genes *DcFAD2-7, DcFAD2-8, DcFAD2-16, DcFAD2-17, DcFAD2-19, DcFAD2-22* on chromosome chr\_4 and *DcCER1-1, DcCER1-2* and *DcCER1-6* on chr\_9 were analyzed first, and their relative expression levels are shown in Figure 4. A high expression level was observed for *DcFAD2-8* and *DcFAD2-19*, whereas transcripts were hardly detected for *DcFAD2-16, DcFAD2-22, DcCER1-1, DcCER1-2*, and *DcCER1-6*. For the majority of the genes, transcripts accumulate most in periderm and least in leaf tissue, with the exception of *FAD2-16* and *DcCER1-2*. They tend to be leaf-specific and petiole-specific, respectively. Among all tissues and cultivars, the highest gene expression level was found in the periderm of BRL for *DcFAD2-7, DcFAD2-8, DcFAD2-17, DcFAD2-19, DcFAD2-22, DcCER1-1* and *DcCER1-6*. BRL showed higher expression levels for *DcFAD2-7, DcFAD2-19, DcFAD2-22* in all tissues than in the corresponding tissues of AN and PR. For *DcFAD2-8* and *DcFAD2-19* in the periderm and the mixture of phloem and xylem, expression levels were highest in BRL and lowest in PR, which correspond with the levels of average PA contents in BRL (high) and PR (low), respectively (PA data not shown).

For *DcFAD2-4* associated with the minor QTL on chr\_8, the transcript levels showed the same pattern as for *DcFAD2-8* and *DcFAD2-19* located within the support interval of the strong QTL on chr\_4, and the expression in the periderm was highest among all three cultivars (Additional file 11: Fig. S8). For the genes which are not associated with QTLs, such as *DcFAD2-5, DcFAD2-6*, and *DcFAD2-12*, expression levels are high in periderm and low in leaf and petiole. However, *DcFAD2-12* was the only analysed *FAD2* gene, which had a higher expression level in the vascular system (mix of phloem and xylem) compared to the expression level in the periderm.

To get an impression about the genotypic variability of candidate gene expression patterns in different individuals of the parental lines, we analyzed the relative expression of the *FAD2* and *CER1* candidate genes located on chr\_4 and chr\_9 in each six individuals of parental lines P\_1870 and P\_345B. The total RNA was isolated from the 'PPX' samples (mixed periderm, phloem and xylem tissues) from the individual root harvested for each genotype. For P\_1870 individuals, the xylem was excluded due to lignification. The relative transcript levels were highly differentiated among the individuals even within the parental lines indicating a large genetic diversity for this trait in the parental lines (Additional file 12: Fig. S9). Similar to the tissue-specific expression study, transcripts of *DcFAD2-16, DcFAD2-22, DcCER1-1, DcCER1-2* were barely detected in all tissues. The individuals which have a high expression level of *DcFAD2-8*, expressed *DcFAD2-7, DcFAD2-17* and *DcCER1-6* in a high level as well.

## Discussion

### Polyacetylene quantification in a carrot $F_2$ family

In the present study a carrot  $F_2$  family derived from a cross of a cultivated carrot breeding line and the wild relative *D. c. commutatus* was used for quantification of major PAs FaOH and FaDOH. Since the roots of the original parental plants were not available for PA analysis, we measured the PA contents

of each six siblings raised for the pilot experiment with the aim to identify parental lines differing in PA contents and to justify a bi-parental QTL approach. The average PA contents in the wild parent appeared to be considerably higher than the average PA contents in the cultivated breeding line. In the rare PA studies based on carrot wild relatives it was shown, that *Daucus* species and subspecies can contain much higher PA contents than cultivated carrot forms [9]. Accessions of *D. c. commutatus* were among the (sub)species with the highest FaOH and FaDOH contents [10]. PA data collected from the pilot experiment as well as those from the main experiment showed that - with the exception of three plants - all investigated plants possessed a higher FaDOH level compared with FaOH. According to Czepa and Hofmann [7] the most abundant PA in cultivated orange carrots is FaDOH, and the highest total PA levels were found in the periderm tissue [5, 9, 38]. In our study we focused for QTL analysis on total FaOH and FaDOH contents measured in samples representing a mixture of the different root tissues (periderm, phloem, xylem). Considering the large number of  $F_2$  individuals involved in this study, this approach appeared to be appropriate to detect QTLs relevant for carrot breeding. A further dissection of the root tissues for QTL analysis might be applied in future studies to reveal tissue-specific changes in PA biosynthesis.

A large phenotypic variability was observed for the contents of both PAs in the  $F_2$  population indicating that the selection of the crossing parents was appropriate for the bi-parental QTL approach. Carrot is a highly heterozygous species, and genetic uniformity even in advanced breeding lines appears to be not the rule, as it is demonstrated by the phenotypic variation of the PA contents in the six individuals of the 345B parental line. The observed frequency distributions in the  $F_2$  progeny indicate a polygenic inheritance of the PA production in carrot. This is not surprising considering the complex biochemical pathway from which the falcarinol-type PAs derive. Nevertheless, the high phenotypic variation in the  $F_2$  family, which contains plants with no measurable contents of PAs and on the other side plants with extremely high contents above 1000  $\mu\text{g/g}$  FaOH or even 3000  $\mu\text{g/g}$  FaDOH suggest, that major structural or regulatory genes are involved in the control of PA biosynthesis. Generally, the inheritance of PA concentrations and distribution patterns are poorly understood. The only study on carrot PAs based on a segregating population showed that PA contents are a heritable trait [33]. In the carrot  $F_2$  family used in the 2-years experiment for QTL analyses of falcarinol-type PAs, a comparatively low impact of the environment on the accumulation of PAs was found, and the broad-sense heritability for, e.g. FaDOH was estimated with 0.88. Transgressive segregation was also observed in this study [33].

### Linkage mapping and QTL analysis

Carrot belongs to the vegetable species for which already several partial or saturated linkage maps have been developed and published (for review, see [39]).

For the current study publicly available SNP markers [40] were used for linkage mapping, allowing integration with other published genetic maps. Overall the genetic map is smaller compared to the first high-resolution SNP-based carrot map published by Cavagnaro et al. [41], but the marker order is similar, giving no indication of regions with recombination suppression. It has to be noted, that several regions of segregation distortions were identified, which could have an effect, though limited, on the general lower number of observed recombinations in this study. While segregation distortion does not affect QTL mapping especially in dense marker maps, it may lead to erroneous interpretations of the number of loci controlling a trait of interest [42].

In the present study we dissected the genetic basis of two key PAs, FaOH and FaDOH, by QTL analysis. Several major genomic regions seem to be involved in PA production in carrots. Overall, QTLs were localized on six out of the nine carrot chromosomes, which is a sign for the complexity of the genetic control of PA accumulation in carrot. Strong QTLs for FaOH and FaDOH (LOD > 10) were identified with overlapping 2-LOD confidence intervals on chromosomes chr\_4 and chr\_9 indicating a major involvement of these regions in PA biosynthesis. We propose that these QTL regions include major genes that control the primary or secondary PA metabolism in carrot. The QTL with the strongest statistical support was found for FaDOH in the upper region of chr\_9. Interestingly, in the study of Le Clerc et al. [43] the biggest QTL for FaDOH was also located on chr\_9, but according to their linkage map of chr\_9 the QTL was on the opposite side. Due to the usage of proprietary markers used by these researchers (V. le Clerc, pers. comm.) it was not possible to verify the orientation of chr\_9. However, it seems possible that both strong QTLs for FaDOH share a common biological origin. The QTL identified by us for both PAs in the lower region of chr\_4 could be the same QTL identified for FaDOH by Le Clerc et al. [43] in this genomic region. A QTL for falcariindiol-3-acetate, a minor PA probably metabolized from FaDOH, was detected by latter researchers on top of chr\_8, where we identified a QTL for the FaDOH/FaOH ratio. QTLs for PA ratio might be useful to identify loci controlling the interconversions among the PAs. It is presumed, that FaDOH is produced from the precursor substance FaOH [4, 6]. As the QTL effects of allele 'B' from the wild parent P\_1870 are responsible for a lower PA ratio in 'BB' genotypes, it is probably that the QTLs for the PA ratio indicate impaired conversion of FaOH into FaDOH (i.e. accumulation of FaOH). The generally increasing QTL effects of the 'B' allele for all other QTLs associated with total amounts of single PAs indicate, that rather the primary metabolism of common precursors for all falcarinol-type PAs, as for instance crepenynic and dehydrocrepenynic acid, is controlled by genetic factors underlying these QTLs. Likely there was a selection during the domestication process against unfavourable "wild" alleles associated with high FaDOH contents to avoid a too high level of bitterness in cultivated carrots.

Overall, the results from our QTL analysis confirmed the utility of a carrot family derived from a wide cross of a cultivated carrot with a *D. c. commutatus* accession rich in PAs. Previous studies based on the genetics of carotenoid accumulation used a mapping population derived from a cross between a cultivated orange carrot and a wild white carrot collected in North America [44, 45]. Important advances in the understanding of the genetic control of carrot anthocyanin pigmentation have been made by using a purple wild carrot parent from Turkey [41, 42].

To dissect the genetic basis of a quantitative trait, such as the content of a natural product, two main methods have been used in the past, bi-parental QTL mapping and genome-wide association studies (GWAS). Conventional QTL mapping depends on a diverging genetic diversity of two parents and is very time-consuming since in several crop plants including carrot a  $F_2$  mapping population has to be developed. Moreover, QTL regions can be quite large and may include many potential candidate genes. Nevertheless, this method has been used in carrot research and breeding to elucidate the genetics of important secondary metabolites determining root quality such as  $\beta$ -carotene, anthocyanins, polyacetylenes, and volatile terpenes [32, 33, 41, 44]. GWAS can overcome the limitations of bi-parental QTL mapping and has great potential for detection of QTLs with high resolution in diverse sequenced genotypes. Combinatorial approaches are also very useful to compensate the limitations of each method. Both QTL analysis and GWAS have been used for the detection of *DcTPS54*, a sabinene synthase gene putatively involved in carrot flavour [32, 46] and discovered the *Org* gene controlling carotene accumulation [47, 48]. GWAS aimed at PA

compounds in large sets of carrot cultivars, land races and wild relatives would allow the identification of additional QTLs, but might be especially helpful to narrow down the QTL intervals for both FaOH and FaDOH on chromosome chr\_4 allowing the more precise association with candidate genes.

### Candidate genes associated with QTLs

In this work, the combination of a PA metabolite quantification, bi-parental QTL analysis and the discovery of putative candidate genes from two different gene families involved in fatty acid metabolism have been used to get insights into the genetic control of PA biosynthesis. Few biochemical pathways have been studied in carrot by a combinatorial approach based on linkage mapping, QTL analysis and candidate gene identification. In contrast to previous investigations on carrot carotenoids, anthocyanins and terpenes (for review, see [49, 50, 51]), no studies have been, to date, reported on comparable candidate gene approaches for PA compounds. The reason is, that in higher plants little is known on the genetic control and the enzymes involved in the biosynthesis of these compounds. Until recently, no carrot genes involved in PA production have been described. Work in the past on parsley (*Petroselinum crispum*, Apiaceae) identified a divergent form of FAD2 that was upregulated in response to pathogen attack and, when expressed in soybean embryos, resulted in the production of crepenynic acid and dehydrocrepenynic acid [36, 52]. These results of the parsley studies are consistent with a pathogen-responsive, divergent FAD2-mediated pathway leading to the accumulation of acetylenic fatty acids needed for PA formation [5]. The FAD2 inventory work of Busta et al. [5] resulted in the identification of 24 carrot FAD2 members, and it was shown after functional analysis of six genes in yeast and *Arabidopsis thaliana*, that these genes represented the major entry point into carrot PA biosynthesis. We performed a reannotation of the carrot genome sequence [34] and detected seven new FAD2 gene models, which increases the total number of carrot FAD2 family member to the large total number of 31. It is a striking feature, that carrots seem to contain the highest number of FAD2 family members reported so far in the plant kingdom. Knowledge about the FAD2 family including the divergent forms is based largely on research on *A. thaliana* and on oil seed crops. Despite the evidence for functional diversification of FAD2 in multiple plant families, the potential of such diversification in major crops including possible roles in biotic stress resistance has been largely unexplored [25].

Recently a biosynthetic gene cluster was discovered in tomato, that was required for FaDOH production. Interestingly, among the four highly co-expressed clustered genes there was a *CER1 decarbonylase* [27]. To our knowledge, this was the first time, that a possible *CER1* gene function was reported, that is different from the so far recognized function in the alkane biosynthetic pathway. The *Arabidopsis* CER1 protein is known as a decarbonylase that converts fatty acid metabolites into alkanes. Alkanes are components of waxes in the plant cuticle, a waterproof barrier serving to protect land plants from both biotic and abiotic stimuli [53]. We searched for carrot orthologous genes and found six putative *CER1* genes showing a close phylogenetic relationship with the previously characterized tomato *CER1* and additionally six putative *CER3* gene models. Few reports are available about *CER* gene families in plants. Characterization of the *CER1* family genes in *Brachypodium distachyon* identified eight *CER1* homologs [31], and in rice seven putative *CER1* paralogs were identified [54].

We identified two genomic regions on two different carrot chromosomes around QTLs for FaOH and FaDOH with high LOD values of > 14 and > 21, respectively, that contain several FAD2 and CER1/3 genes. The six FAD2 genes clustering on chr\_4 and the three CER1 genes on chr\_9 are preliminary considered as candidate genes, based on their location within 2-LOD QTL confidence intervals. However, a causal relationship between CER1 genes and the production of FaOH and FaDOH remains to be biochemically explained. Concerning the six FAD2 genes associated with the major QTLs on chr\_4 their distance to the QTL main peak is larger. However, in the upper region of this chromosome (outside of the 2-LOD interval) the gene *DcFAD2-31* is located, which might have an influence on the calculated QTL peak positions. The six carrot FAD2 genes associated with the strong QTL on chr\_4 reside close to one another within a small region (29.3 – 29.4 Mbp) and might have been originated from local tandem duplications, as data from microsyntenic and phylogenomic analyses suggest [5]. This cluster also contains the two genes *DcFAD2-7* and *DcFAD2-8*, which have been functionally characterized as  $\Delta$ 12-fatty acid acetylenases, and additionally the gene *DcFAD2-19* which was described as a bifunctional enzyme capable of catalysing both  $\Delta$ 12 desaturation of oleate and  $\Delta$ 14 desaturation of crepenynate [5]. Interestingly, in the plants of BRL, which showed high PA levels in comparison to other carrot cultivars, a co-expression seem to have occurred in the periderm tissue samples for five out of the six FAD2 candidates including *DcFAD2-7*, *DcFAD2-8*, and *DcFAD2-19*. Taken together, it is likely that the QTL region on chr\_4 plays a major role in carrot PA production. Nevertheless, the situation, that 10 out of the 12 QTLs, with the exception of those on chr\_3, are associated with FAD2 candidate genes, support the suggested importance of the FAD2 family on the PA biosynthesis in Apiaceae. Further functional and biochemical studies are needed to identify the most relevant FAD2s. Even beyond identifying FAD2 and CER candidate genes for QTLs identified in this study, more genes involved in PA biosynthesis might be identified elsewhere in the carrot genome. Since the PAs in this study are related in terms that they represent compounds produced at different steps in the same biosynthetic pathway, it is also possible that a single regulatory gene like a transcription factor affects the compounds produced in this pathway.

### Implications for carrot breeding

Bitterness is considered as an undesirable taste of carrot roots, which can cause consumer rejection and is one of the main reasons for low preference scores in sensory evaluations of carrots [13]. It is a very complex quality trait because numerous chemically different compounds may contribute to bitter taste. Potential bitter compounds in carrot are volatile mono- and sesquiterpenes, PAs, phenylpropanoids, and isocoumarins. Schmid et al. [8] recently listed 14 known bitter off-taste compounds including the major PAs FaOH and FaDOH. In the study of Le Clerc et al. [33] the total PA content was closely related to bitterness, and the highest quantities accumulated in the most bitter genotypes, whereas the lowest amounts were measured in the least bitter genotypes. There is some evidence from quantitative chemical analyses combined with sensory analysis, that FaDOH is highly correlated with bitterness, whereas FaOH is not [7, 13]. On the other hand, several *in vitro* studies have confirmed, that FaOH is one of most cytotoxic PAs in Apiaceae vegetables and, according the current state of knowledge, has a higher bioactivity than FaDOH [19]. A better understanding of the genetics of the PA levels present in carrot roots might support breeding carrot cultivars with low bitterness but high health potential for the consumers. To reach this goal it will be necessary to reveal the genes that control the decisive steps in the complex PA biosynthesis pathway, i.e. the formation of FaDOH. Carrot chemotypes with acceptable amounts of bioactive PAs may contribute significantly to the known positive effects of carrots on human health.



## Conclusions

The aim of the current study was to dissect the complex genetic control of PA accumulation in carrot roots. We used a carrot  $F_2$  family derived from a carrot wild relative rich in PAs for SNP-based QTL mapping and detected several major QTLs for FaOH and FaDOH with high LOD values on chromosomes 4 and 9. To study the association of these QTLs with *FAD2* candidate genes known to be involved in PA biosynthesis, we compared the genomic positions of *FAD2* gene models with QTL intervals and found significant relationships, such as the putative involvement of a known and partly functionally characterized six-*FAD2* gene cluster located on carrot chromosome 4 which contain acetylenase genes. To discover other putative candidate genes from the plant fatty acid metabolism, we performed a first inventory of *CER1* genes, which were recently reported to be involved in the production of FaDOH in tomato. The genetic association of three putative *CER1 decarboxylase* candidates with the strongest QTLs on chromosome chr\_9 might be a first indication for this hypothesis. Our finding, that more or less functionally characterized carrot genes from the plant fatty acid metabolism are significantly associated with major QTLs for key PAs, will facilitate future functional gene studies and a further dissection of the genetic factors controlling PA accumulation. The results of this research are considered to have a positive impact on carrot breeding programs aiming at lowering or increasing PA concentrations and PA patterns, dependent on the requirements of the consumers.

## Methods

### Plant material and sample preparation

$F_2$  population P11054 (called in this study 'CA') was derived from a single  $F_1$  plant out of a cross between a cultivated carrot inbred line '345B' (parent 'P\_345B'; proprietary BASF) and an inbred from *Daucus carota* ssp. *commutatus* (parent 'P\_1870'). Both parental inbred lines were produced at Nunhems Netherlands BV (Nunhem, The Netherlands). Seeds from the original *D. c. commutatus* accession no. JKI-1870 are available from the carrot seed repository of Julius Kühn-Institute (JKI) upon request (corresponding author). The  $F_2$  population CA was chosen based on a pilot test with PA content measurements (spring 2020), including the parental lines (6 individuals each) and a small subset of 9 individuals from  $F_2$  population CA.

For the main experiment  $F_2$  population CA was sown in the field in May 2020, grown during the summer and harvested in October 2020 at the BASF research station in Nunhem (Limburg, Netherlands). About 15 % of the plants in the field showed to bolt. A total of 550  $F_2$  individuals were evaluated in the field and 400 non-bolted  $F_2$  individuals were selected for further PA content measurements. Next 80 bolted  $F_2$  individuals were sampled, but excluded from further analysis, as the roots were less developed, possible hampering reliable PA content measurements. For PA analysis of the  $F_2$  plants three plugs were excised from an individual carrot root at top, middle and bottom position using a cork borer ( $\varnothing$  8 mm, depth 3 mm). These three plugs were then pooled for PA analysis. Plant material was shock frozen in liquid nitrogen immediately after harvest and stored at  $-80^\circ\text{C}$  until freeze-drying.

A pilot test with PA content measurements was carried out to identify parental lines differing in PA content and to prove if there is sufficient phenotypic variability amongst the individuals of the  $F_2$  progeny. For the pilot test a slightly modified sampling method called 'PPX' was used, because some roots of the parental line P\_1870 were extremely thin (Additional file 2: Fig. S1) and therefore it was not possible to sample a 3mm-thick and 8 mm-in-diameter plug for these roots. For PPX sampling the root was first cut longitudinally into two halves and the PPX sample (a plug cylinder consisted of periderm, phloem and xylem from a half of the root) was taken by a 8-mm cork-borer. The samples were shock frozen in liquid nitrogen immediately after harvest and stored at  $-80^\circ\text{C}$  until PA analysis and expression analysis of candidate genes.

### Analysis of polyacetylenes by HPLC/DAD

Deep-frozen root plugs (main experiment) or PPX samples (pre-test) were freeze-dried for four days (Christ Gamma 1-16 LSC, condenser temperature  $-50^\circ\text{C}$ , pressure 0.04 mbar). Freeze-dried plugs of individual roots were precisely weighed into 1.5-mL polypropylene centrifuge tubes (approx. 20 – 30 mg dry weight). After addition of three steel balls ( $\varnothing$  3 mm) plugs were homogenized using a mixer mill (Retsch MM 400, 30 Hz,  $2 \times 60$  s). After homogenization 50  $\mu\text{L}$  internal standard solution (0.2  $\text{g L}^{-1}$  *N*-vanillylnonamide in MeOH) and 300  $\mu\text{L}$  acetone were added. The mixture was homogenized once again using a mixer mill (30 Hz, 60 s, room temperature), sonicated (ElmaSonic P, 37 kHz, 100 W, 5 min,  $20^\circ\text{C}$ ) and shaken ( $2400 \text{ min}^{-1}$ , 10 min, room temperature). After centrifugation ( $13000 \text{ g}$ , 5 min,  $20^\circ\text{C}$ ) a 200- $\mu\text{L}$  aliquot of the supernatant was transferred into an HPLC vial with micro-insert and stored at  $6^\circ\text{C}$  until analysis.

Polyacetylene analyses were performed on an 1100 Series HPLC system (Agilent Technologies) comprising a degasser (G1322A), a binary pump (G1312A), an autosampler (G1329A), an autosampler thermostat (G1330A), a column compartment (G1316A) and a diode array detector (G1315A). Extracts (injection volume 2.5  $\mu\text{L}$ ) were separated on a Zorbax Eclipse XDB-C18 column (3 mm  $\times$  150 mm, 3.5  $\mu\text{m}$  particle size, Agilent Technologies) using water and acetonitrile as eluent A and B, respectively. The following binary gradient program at a flow rate of  $1 \text{ mL min}^{-1}$  was used: 0-10 min, linear from 50 to 80 % B; 10-10.5 min, linear from 60 to 100 % B; 10.5-13 min, isocratic, 100 % B; 13-15 min, isocratic 50 % B. The column and autosampler temperature was maintained at  $40^\circ\text{C}$  and  $6^\circ\text{C}$ , respectively. The diode array detector response time was set at 0.2 s, the optical slit width at 4 nm. Polyacetylenes were detected at 196 nm with a spectral bandwidth of 4 nm, the internal standard *N*-vanillylnonamide at 204 nm with a spectral bandwidth of 4 nm. ChemStation software (version B.03.02) was applied for controlling the instrument, data acquisition and quantitative analysis. *N*-Vanillylnonamide ( $t_R$  2.53 min), faltarindiol ( $t_R$  5.09 min) and faltarinol ( $t_R$  9.48 min) were quantified based on peak area using external standard calibration method. Therefore, the following calibration curves were established: (i) *N*-vanillylnonamide: calibration range 1 – 600 ng, 14 points, linear regression model ( $y = m x$ ), equal weighting,  $R^2 = 0.99987$ ; (ii) Faltarindiol: calibration range 5 – 500 ng, 9 points, linear regression model ( $y = m x$ ), equal weighting,  $R^2 = 0.99996$ ; (iii) Faltarindiol: calibration range 500 – 1000 ng, 6 points, logarithmic regression model ( $y = m \ln(x) + b$ ), equal weighting,  $R^2 = 0.99919$ ; (iv) Faltarinol: calibration range 5 – 1000 ng, 14 points, linear regression model ( $y = m x$ ), equal weighting,  $R^2 = 0.99970$ . Polyacetylene levels were corrected using the recovery rate of the internal standard.

### Construction of a SNP-based linkage map and QTL analysis

DNA of 400 CA F<sub>2</sub> plants was extracted from dried young leaflets using LGC/SBeadex<sup>tm</sup> plant DNA purification kit (standard protocol). Genotyping was performed with a proprietary SNP array. A dataset, comprised of 374 F<sub>2</sub> individuals and 705 polymorphic literature markers [40] and 8 proprietary SNP markers, was used for genetic mapping. The SNP scores were translated to parental scores, with the A genotype for homozygous P<sub>-345B</sub> parent allele, B genotype for homozygous P<sub>-1870</sub> parent allele, H for heterozygous genotype and U for undetermined genotype. For construction of the genetic linkage map Joinmap 5.0 software [55] was used with standard settings. Markers were grouped at a recombination frequency threshold of 0.25 into nine linkage groups, which were renamed to their representing chromosomes (chr\_1 until chr\_9). Maximum Likelihood (ML) mapping was used to calculate the genetic distances. After the calculation overlapping genetic positions were pruned, leaving a genetic map only represented with unique genetic positions. The QTL analysis was conducted in R 3.6.2, using R/qtl package [56]. Interval Mapping (IM) and Composite Interval Mapping (CIM) were used for a preliminary QTL detection, followed by a “scantwo” two-dimensional scan to detect possible QTL interactions. For each analysis, 5,000 permutations were performed (n.perm = 5000). The significance level was set to 95% (alpha = 0.05). Final QTL models were made with FitQTL and LODINT with a 2-LOD drop was used to determine the confidence intervals.

### Discovery of *FAD2* and *CER1/3* candidate genes

The homology-based gene prediction program GeMoMa (version 1.6.beta, [57]) was used to reannotate the published carrot genome [34] and to determine potential *FAD2* candidate genes in carrot, which have not yet been previously described. Totally 12 genomes (*Arabidopsis thaliana* (TAIR10), *Brachypodium distachyon* (v3.1), *Glycine max* (Wm82.a2.v1), *Lactuca sativa* (v5), *Mimulus guttatus* (v2.0), *Oryza sativa* (v7.0), *Prunus persica* (v2.1), *Populus trichocarpa* (v3.1), *Sorghum bicolor* (v3.1.1), *Setaria italica* (v2.2), *Solanum lycopersicum* (v2.5), *Theobroma cacao* (v1.1)) were used as reference with the software parameters described in [32]. The newly annotated carrot genome was used for BLASTP searches based on the putative 24 members of the carrot *FAD2* family [5]. For a first approach to identify putative carrot *CER1/3* genes the annotated tomato gene *Solyc12g100270* predicted as putative *CER1 decarboxylase* [27] was used for BLAST of the carrot genome vers.2 [34] by Phytozome vers.13 [58]. All identified putative *CER1/3* sequences annotated in the carrot genome were used for a BLASTP search at NCBI using the non-redundant protein sequence database. Predicted carrot *CER1/3* protein sequences and known plant *CER1/3* sequences were used for comparisons with the original DCAR sequences. In case that the DCAR locus appeared to be incompletely annotated, the DCAR gene model was edited by adding the missing regions (optimized prediction).

### Expression analysis of *FAD2* and *CER1/3* candidate genes

The tissue-specific expression analysis of candidate genes was carried out with the purple cultivar ‘Anthonina’ (AN; Seminis, USA), the orange cultivar ‘Presto’ (PR; Vilmorin, France) and a selected orange carrot breeding line (BRL; Nunhems BV, Netherlands). For the expression study seeds were sown and cultivated in 19 cm/30 cm W/H plastic pots in a sand-humus mixture (v/v 3/1) under optimized greenhouse conditions at 25/20°C D/N and 16 h photoperiod until harvest. Samples from leaves, petioles, root periderm and mixed root phloem/xylem were taken from three individual plants about 100 days after sowing, immediately shock frozen and stored at -80°C until RNA isolation. Total RNA was isolated from leaf, petiole, periderm, and the mixture of phloem and xylem of carrot cultivars AN, BRL, and PR. RNA was extracted from 80-100 mg plant material utilizing innuPREP PlantRNA Kit (Analytik Jena, Jena, Germany) with an additional DNase treatment by innuPREP DNase I Digest Kit (Analytik Jena), according to the manufacturer’s instructions. First-strand cDNA was synthesized by Revert-Aid First Strand cDNA Synthesis Kit (Thermo Scientific, Massachusetts, USA) following the manufacturer’s instructions. The Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) reactions were performed in C1000 Touch<sup>TM</sup> Thermal Cycler (Bio-Rad Laboratories, California, USA). The amplification was carried out in duplicate in 10µl reaction volumes consisting of 1× iTaq SYBR Green Supermix (Bio-Rad), 0.4 µM each primer and 1µl cDNA. Primers were designed using NCBI Primer-BLAST tool (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). The information for the primer used for this study is listed in Additional file 13: Table S3. The qPCR thermal cycling conditions were as follows: initial denaturation at 95°C for 5 min, followed by denaturation for 5 s, annealing and extension at 60°C for 30 s for 40 cycles. The data from three individual plants were normalized to the housekeeping genes *actin1* (*ACT1M*), *elongation factor* (*EF-1a*) and *poly-ubiquitin 10* (*UBQ*) [59]. The relative gene expression levels were calculated using the delta Ct method with calculated primer efficiencies [60].

## Declarations

### Acknowledgements

The authors wish to thank Vicky Bartels (JKI Quedlinburg), Astrid Hansen (JKI Berlin), Mariana Dias Domingos and Bram Colbers (Nunhems BV) for their excellent technical assistance.

### Authors' contributions

FD, CB, TN, PH and FH conceived and designed the experimental layout. PH created the mapping population. CB conducted the PA analysis. FH performed the linkage and QTL mapping. SR prepared the reannotation of the carrot genome and identified the additional *FAD2* genes. FD performed the search and evaluation of carrot *CER1/3* genes. WH conducted the qRT-PCRs. FD, CB and FH analysed the data. FD, WH, CB and FH drafted the manuscript. FD, CB, SR, TN, PH and FH contributed to the discussion and interpretation of results. All authors read, reviewed and approved the final manuscript.

### Funding

This project was partly funded by the vegetable seed company BASF-Nunhems BV, The Netherlands.

### Availability of data and materials

Seeds of the original *D. c. commutatus* accession no. JKI-1870 are available from the carrot seed repository of Julius Kühn-Institute (JKI) upon request (contact: corresponding author). All data generated or analysed during this study are included in this published article and its supplementary information files. The predicted sequences of eight gene models *DcFAD2-25* to *DcFAD2-31* and *DcCER1-3*, which are not yet annotated in the current carrot genome vers.2 [34], are included in the supplement of this article (Additional File 5: Data S1).

### Ethics approval and consent to participate

Not applicable

### Consent of publication

Not applicable

### Competing interests

The authors declare that they have no competing interests.

## References

1. Arcscott SA, Tanumihardjo SA. Carrots of many colors provide basic nutrition and bioavailable phytochemicals acting as a functional food. *Compr. Rev. Food Sci.* 2010;9:223–39.
2. Ahmad T, Cawood M, Iqbal Q, Ariño A, Batool A, et al. Phytochemicals in *Daucus carota* and their health benefits - Review Article. *Foods.* 2019;8:424. <https://doi.org/10.3390/foods8090424>.
3. Christensen LP. Aliphatic C(17)-polyacetylenes of the falcarinol type as potential health promoting compounds in food plants of the Apiaceae family. *Recent Patents on Food, Nutrition & Agriculture.* 2011;3:64–77.
4. Dawid C, Dunemann F, Schwab W, Nothnagel T, Hofmann T. Bioactive C<sub>17</sub>-polyacetylenes in carrots (*Daucus carota* L.): Current knowledge and future perspectives. *J. Agric. Food Chem.* 2015;63:9211–22. <https://doi.org/10.1021/acs.jafc.5b04357>.
5. Busta L, Yim WC, LaBrant EW, Wang P, Grimes L, Malyszka K, et al. Identification of genes encoding enzymes catalyzing the early steps of carrot polyacetylene biosynthesis. *Plant Physiol.* 2018;178:1507–21.
6. Minto RE, Blacklock BJ. Biosynthesis and function of polyacetylenes and allied natural products. *Prog. Lipid Res.* 2008;47:233–306.
7. Czepa A, Hofmann T. Quantitative studies and sensory analyses on the influence of cultivar, spatial tissue distribution, and industrial processing on the bitter off-taste of carrots (*Daucus carota* L.) and carrot products. *J. Agric. Food Chem.* 2004; 52:4508–14. <https://doi.org/10.1021/jf0496393>.
8. Schmid C, Sharma S, Stark TD, Günzkofer D, Hofmann TF, et al. Influence of the abiotic stress conditions, waterlogging and drought, on the bitter sensometabolome as well as agronomical traits of six genotypes of *Daucus carota*. *Foods.* 2021;10:1607. <https://doi.org/10.3390/foods10071607>.
9. Baranska M, Schulz H, Baranski R, Nothnagel T, Christensen L. In situ simultaneous analysis of polyacetylenes, carotenoids and polysaccharides in carrot roots. *J. Agric. Food Chem.* 2005;53:6565–71. <https://doi.org/10.1021/jf0510440>.
10. Schulz-Witte J. Diversität wertgebender Inhaltsstoffe bei *Daucus carota* L. Thesis, 2011. Technical University, Braunschweig
11. Czepa A, Hofmann T. Structural and sensory characterization of compounds contributing to the bitter off-taste of carrots (*Daucus carota* L.) and carrot puree. *J. Agric. Food Chem.* 2003;51:3865–73. <https://doi.org/10.1021/jf034085>.
12. Schmiech L, Uemura D, Hofmann T. Reinvestigation of the bitter compounds in carrots (*Daucus carota* L.) by using a molecular sensory science approach. *J. Agric. Food Chem.* 2008;56:10252–260. <https://doi.org/10.1021/jf8023358>.
13. Kreuzmann S, Christensen LP, Edelenbos M. Investigation of bitterness in carrots (*Daucus carota* L.) based on quantitative chemical and sensory analyses. *LWT - Food Science and Technology.* 2008; 41:193–205.
14. Hansen L, Boll PM. Polyacetylenes in Araliaceae: Their chemistry, biosynthesis and biological significance. *Phytochemistry.* 1986;25:285–93.
15. Lecomte M, Berruyer R, Hamama L, Boedo C, Hudhomme P, Bersihand S, Arul J, et al. Inhibitory effect of the carrot metabolites 6-methoxymellein and falcarindiol on development of the fungal leaf blight pathogen *Alternaria dauci*. *Phys. Mol. Plant Path.* 2012;80:58–67.
16. Zidorn C, Jöhner K, Ganzera M, Schubert B, Sigmund EM, Mader J, et al. Polyacetylenes from the Apiaceae vegetables carrot, celery, fennel, parsley, and parsnip and their cytotoxic activities. *J. Agric. Food Chem.* 2005; 53:2518–23. <https://doi.org/10.1021/jf048041s>.
17. Christensen LP, Brandt K. Bioactive polyacetylenes in food plants of the Apiaceae family: occurrence, bioactivity and analysis. *J. Pharm. Biomedical Analysis.* 2006;41:683–93. <https://doi.org/10.1016/j.jpba.2006.01.057>.
18. Kobaek-Larsen M, Baatrup G, Notabi MK, El-Houri RB, Pipó-Ollé E, Christensen-Arnspang E, Christensen LP. Dietary polyacetylenic oxylipins falcarinol and falcarindiol prevent inflammation and colorectal neoplastic transformation: A mechanistic and dose-response study in a rat model. *Nutrients.* 2019;11:2223. <https://doi.org/10.3390/nu11092223>.
19. Christensen LP. Bioactive C17 and C18 acetylenic oxylipins from terrestrial plants as potential lead compounds for anticancer drug development. *Molecules.* 2020;25:2568. <https://doi.org/10.3390/molecules25112568>.
20. Shanklin J, Cahoon EB. Desaturation and related modifications of fatty acids. *Annu Rev Plant Physiol Plant Mol Biol.* 1998;49:611–41.
21. Van Der Loo FJ, Broun P, Turner S, Somerville C. An oleate 12-hydroxylase from *Ricinus communis* L. is a fatty acyl desaturase homolog. *Proc Natl Acad Sci USA.* 1995;92:6743–47.

22. Lee M, Lenman M, Banaś A, Bafor M, Singh SP, Schweizer M, et al. Identification of non-heme diiron proteins that catalyze triple bond and epoxy group formation. *Science*. 1998;280:915–18.
23. Cahoon EB, Kinney AJ. Dimorphecolic acid is synthesized by the coordinate activities of two divergent delta12-oleic acid desaturases. *J Biol Chem*. 2004;279:12495–502.
24. Okuley J, Lightner J, Feldmann K, Yadav N, Lark E, Browse J. Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell*. 1994;6:147–58.
25. Lee MW, Padilla CS, Gupta C, Galla A, Pereira A, Li J, Goggin FL. The *fatty acid desaturase2* family in tomato contributes to primary metabolism and stress responses. *Plant Physiol*. 2020; 182:1083–99.
26. Somssich IE, Bollmann J, Hahlbrock K, Kombrink E, Schulz W. Differential early activation of defense-related genes in elicitor-treated parsley cells. *Plant Mol Biol*. 1989;12:227–34.
27. Jeon JE, Kim JG, Fischer CR, Mehta N, Dufour-Schroif C, Wemmer K, et al. A pathogen-responsive gene cluster for highly modified fatty acids in tomato. *Cell*. 2020;180:176–87. <https://doi.org/10.1016/j.cell.2019.11.037>.
28. Aarts MG, Keijzer CJ, Stiekema WJ, Pereira A. Molecular characterization of the *CER1* gene of Arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. *Plant Cell*. 1995;7:2115–27. <https://doi.org/10.2307/3870155>.
29. Pu Y, Gao J, Guo Y, Liu T, Zhu L, Xu P, et al. A novel dominant glossy mutation causes suppression of wax biosynthesis pathway and deficiency of cuticular wax in *Brassica napus*. *BMC Plant Biol*. 2013;13:215.
30. Lee SB, Kim H, Kim RJ, Suh MC. Overexpression of Arabidopsis MYB96 confers drought resistance in *Camelina sativa* via cuticular wax accumulation. *Plant Cell Rep*. 2014;33:1535–46.
31. Wu H, Shi S, Lu X, Li T, Wang J, Liu T, et al. Expression Analysis and Functional Characterization of *CER1* Family Genes Involved in Very-Long-Chain Alkanes Biosynthesis in *Brachypodium distachyon*. *Front. Plant Sci*. 2019; 10:1389. <https://doi.org/10.3389/fpls.2019.01389>.
32. Reichardt S, Budahn H, Lamprecht D, Riewe D, Ulrich D, Dunemann F, Koperthek L. The carrot monoterpene synthase gene cluster on chromosome 4 harbours genes encoding flavour-associated sabinene synthases. *Hort. Res*. 2020;7:190. <https://doi.org/10.1038/s41438-020-00412-y>.
33. Le Clerc V, Aubert C, Cottet V, Yovanopoulos C, Piquet M, Suel A, et al. Breeding for carrot resistance to *Alternaria dauci* without compromising taste. *Mol Breeding*. 2019;39:59. <https://doi.org/10.1007/s11032-019-0966-7>.
34. Iorizzo M, Ellison S, Senalik D, Zeng P, Satapoomin P, et al. A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat. Genet*. 2016; 48:657–66.
35. McCartney AW, Dyer JM, Dhanoa PK, Kim PK, Andrews DW, McNew JA, Mullen RT. Membrane-bound fatty acid desaturases are inserted co-translationally into the ER and contain different ER retrieval motifs at their carboxy termini. *Plant J*. 2004;37:156–73.
36. Cahoon EB, Schnurr JA, Huffman EA, Minto RE. Fungal responsive fatty acid acetylases occur widely in evolutionarily distant plant families. *Plant J*. 2003;34: 671–83.
37. Pascal S, Bernard A, Deslous P, Gronnier J, Fournier-Goss A, Domergue F, et al. Arabidopsis *CER1-LIKE1* functions in a cuticular very-long-chain alkane-forming complex. *Plant Physiol*. 2019;179:415–32. <https://doi.org/10.1104/pp.18.01075>.
38. Baranska M, Schulz H. Spatial tissue distribution of polyacetylenes in carrot root. *Analyst*. 2005;130:855–59.
39. Iorizzo M, Ellison S, Pottorff M., Cavagnaro PF. Carrot Molecular Genetics and Mapping. In: Simon P, Iorizzo M, Grzebelus D, Baranski R, editors. *The Carrot Genome*. Cham: Springer Nature Switzerland. 2019. p. 101–17. [https://doi.org/10.1007/978-3-030-03389-7\\_7](https://doi.org/10.1007/978-3-030-03389-7_7).
40. Iorizzo M, Senalik D, Ellison S, Grzebelus D, Cavagnaro P, Allender C, et al. Genetic structure and domestication of carrot (*Daucus carota* subsp. *sativus* L.) (Apiaceae). *Am. J. Bot*. 2013;100:930–38.
41. Cavagnaro PF, Iorizzo M, Yildiz M, Senalik D, Parsons J, Ellison S, Simon PW. A gene-derived SNP-based high resolution linkage map of carrot including the location of QTL conditioning root and leaf anthocyanin Pigmentation. *BMC Genomics*. 2014;15:1118.
42. Bannoud F, Ellison S, Paolinelli M, Horejsi T, Senalik D, Fanzone M, et al. Dissecting the genetic control of root and leaf tissue-specific anthocyanin pigmentation in carrot (*Daucus carota* L.). *Theor. Appl. Genet*. 2019;132:2485–2507. <https://doi.org/10.1007/s00122-019-03366-5>.
43. Le Clerc V, Aubert C, Cottet V, Yovanopoulos C, Piquet M, Suel A, et al. Resistant carrots to *Alternaria dauci* or tasty carrots: should we choose? *Acta Hort*. 2019;1264:199–204. <https://doi.org/10.17660/ActaHortic.2019.1264.24>.
44. Santos CAF, Simon PW. QTL analyses reveal clustered loci for accumulation of major provitamin A carotenes and lycopene in carrot roots. *Mol Genet Genom*. 2002;268:122–129.
45. Just BJ, Santos CAF, Yandell BS, Simon PW. Major QTL for carrot color are positionally associated with carotenoid biosynthetic genes and interact epistatically in a domesticated x wild carrot cross. *Theor Appl Genet*. 2009;119:1155–69. <https://doi.org/10.1007/s00122-009-1117-z>.
46. Keilwagen J, Lehnert H, Berner T, Budahn H, Nothnagel T, Ulrich D, Dunemann F. The terpene synthase gene family of carrot (*Daucus carota* L.): Identification of QTLs and candidate genes associated with terpenoid volatile compounds. *Front. Plant Sci*. 2017;8:1930. <https://doi.org/10.3389/fpls.2017.01930>.
47. Ellison SL, Luby CH, Corak KE, Coe KM, Senalik D, Iorizzo M, et al. Carotenoid presence is associated with the *Or* gene in domesticated carrot. *Genetics*. 2018;210:1497–1508.
48. Coe KM, Ellison S, Senalik D, Dawson J, Simon P. The influence of the *Or* and *Carotene Hydroxylase* genes on carotenoid accumulation in orange carrots [*Daucus carota* (L.)]. *Theor. Appl. Genet*. 2021;134:3351–62. <https://doi.org/10.1007/s00122-021-03901-3>.
49. Simon PW, Geoffriau E, Ellison S, Iorizzo M. Carrot Carotenoid Genetics and Genomics. In: Simon P, Iorizzo M, Grzebelus D, Baranski R, editors. *The Carrot Genome*. Cham: Springer Nature Switzerland. 2019. p. 247–60. [https://doi.org/10.1007/978-3-030-03389-7\\_14](https://doi.org/10.1007/978-3-030-03389-7_14).

50. Cavagnaro PF, Iorizzo M. Carrot Anthocyanin Diversity, Genetics, and Genomics. In: Simon P, Iorizzo M, Grzebelus D, Baranski R, editors. The Carrot Genome. Cham: Springer Nature Switzerland. 2019. p. 261-77. [https://doi.org/10.1007/978-3-030-03389-7\\_15](https://doi.org/10.1007/978-3-030-03389-7_15).
51. Ibdah M, Muchlinski A, Yahyaa M, Nawade B, Tholl D. Carrot Volatile Terpene Metabolism: Terpene Diversity and Biosynthetic Genes. In: Simon P, Iorizzo M, Grzebelus D, Baranski R, editors. The Carrot Genome. Cham: Springer Nature Switzerland. 2019. p. 279–93. [https://doi.org/10.1007/978-3-030-03389-7\\_16](https://doi.org/10.1007/978-3-030-03389-7_16).
52. Kirsch C, Hahlbrock K, Somssich IE. Rapid and transient induction of a parsley microsomal  $\Delta 12$  fatty acid desaturase mRNA by fungal elicitor. *Plant Physiol.* 1997;115:283–89.
53. Chaudhary K, Geeta R, Panjabi P. Origin and diversification of *ECERIFERUM1* (*CER1*) and *ECERIFERUM3* (*CER3*) genes in land plants and phylogenetic evidence that the ancestral *CER1/3* gene resulted from the fusion of pre-existing domains. *Molecular Phylogenetics and Evolution.* 2012;159:107101.
54. Islam MA, Du H, Ning J, Ye H, Xiong L. Characterization of *Glossy1*-homologous genes in rice involved in leaf wax accumulation and drought resistance. *Plant Mol. Biol.* 2009;70:443–56. <https://doi.org/10.1007/s11103-009-9483-0>.
55. Van Ooijen JW. Joinmap 5. Software for the calculation of genetic linkage maps in experimental populations of diploid species. 2018. Kyazma BV, Wageningen, The Netherlands.
56. Broman KW, Sen S. A Guide to QTL Mapping with R/qtl. New York: Springer; 2009.
57. Keilwagen J, Wenk M, Erickson JL, Schattat MH, Grau J, Hartung F. Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res.* 2016;44:e89. <https://doi.org/10.1093/nar/gkw092>.
58. 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:D1178-D1186.
59. Tian C, Jiang Q, Wang F, Wang GL, Xu ZS, et al. Selection of Suitable Reference Genes for qPCR Normalization under Abiotic Stresses and Hormone Stimuli in Carrot Leaves. *PlosOne.* 2015;10:e0117569. <https://doi.org/10.1371/journal.pone.0117569>.
60. Pfaffl MW. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* 2001;29:e45. <https://doi.org/10.1093/nar/29.9.e45>.

## Figures

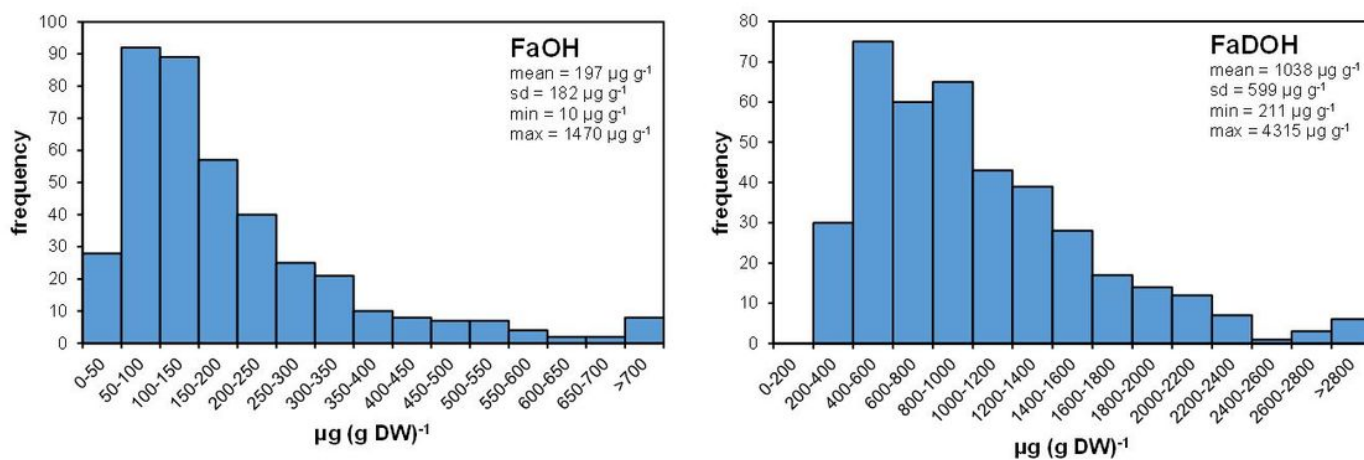
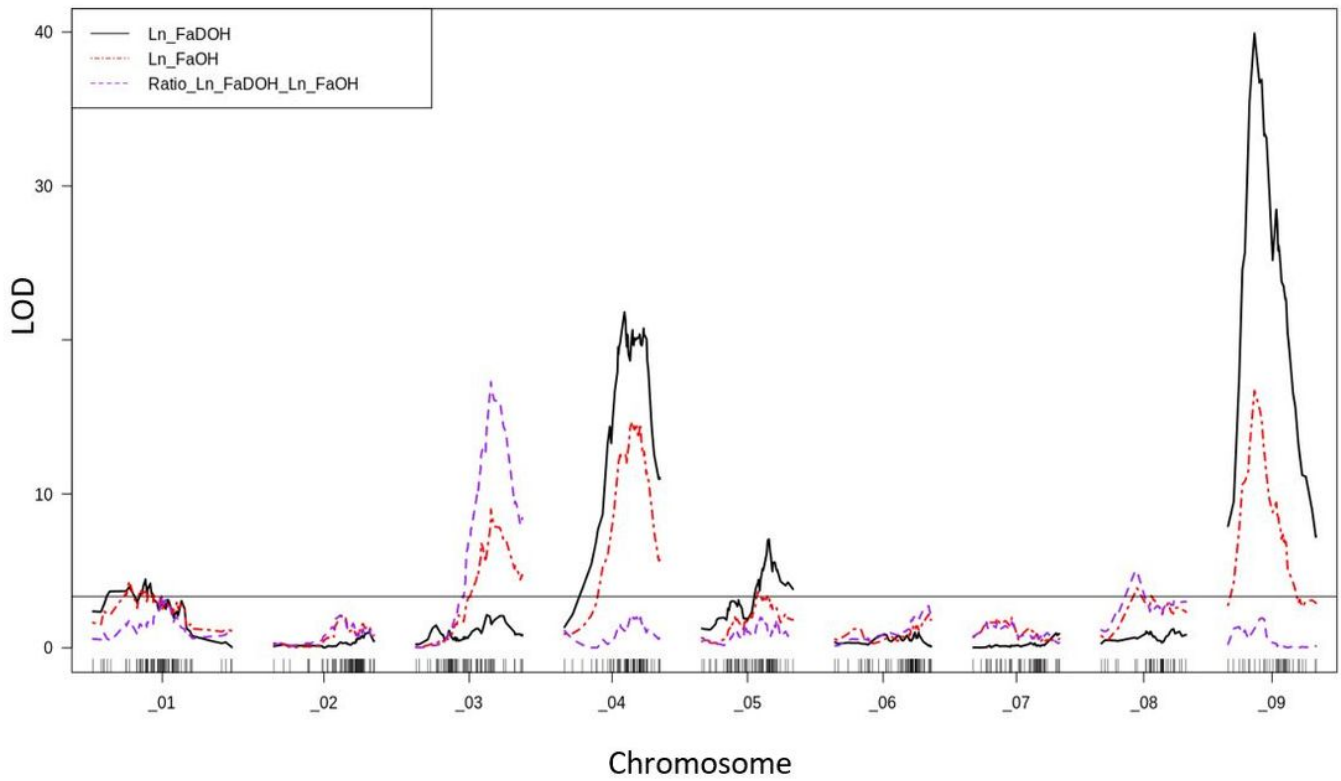
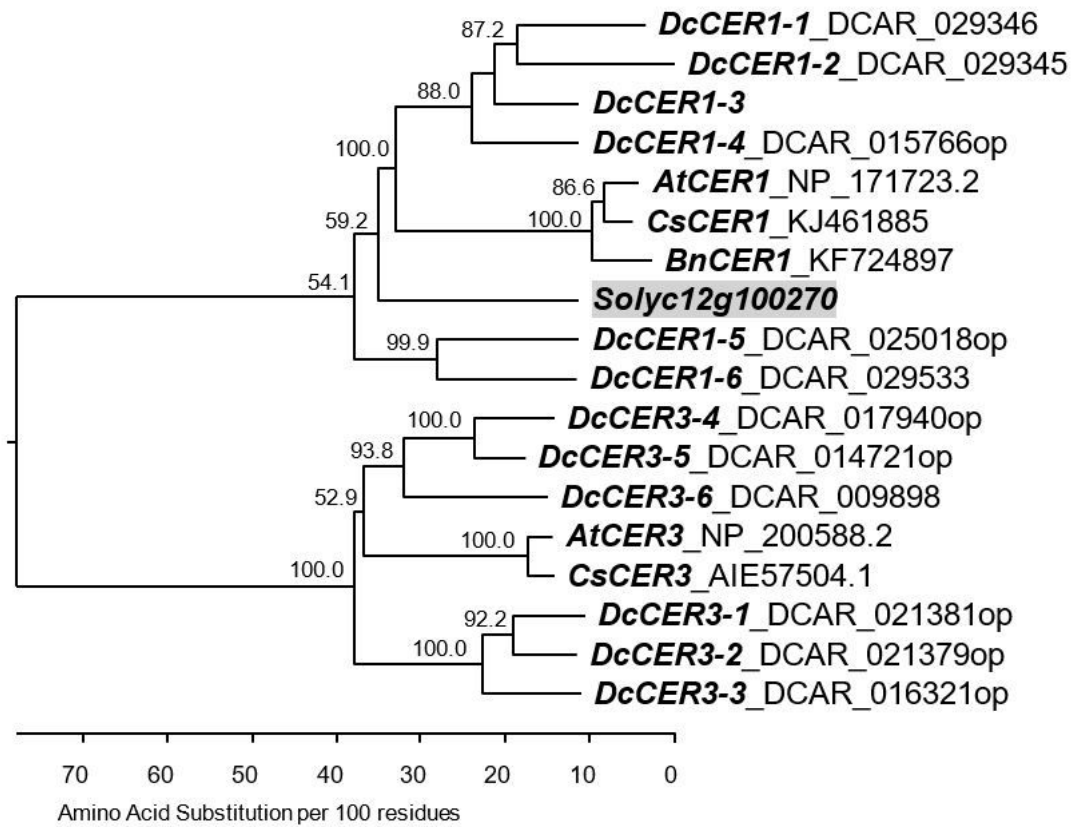


Figure 1

Distribution of falcarinol (FaOH) and falcarindiol (FaDOH) levels in carrot roots of 400 individuals of F2 progeny CA

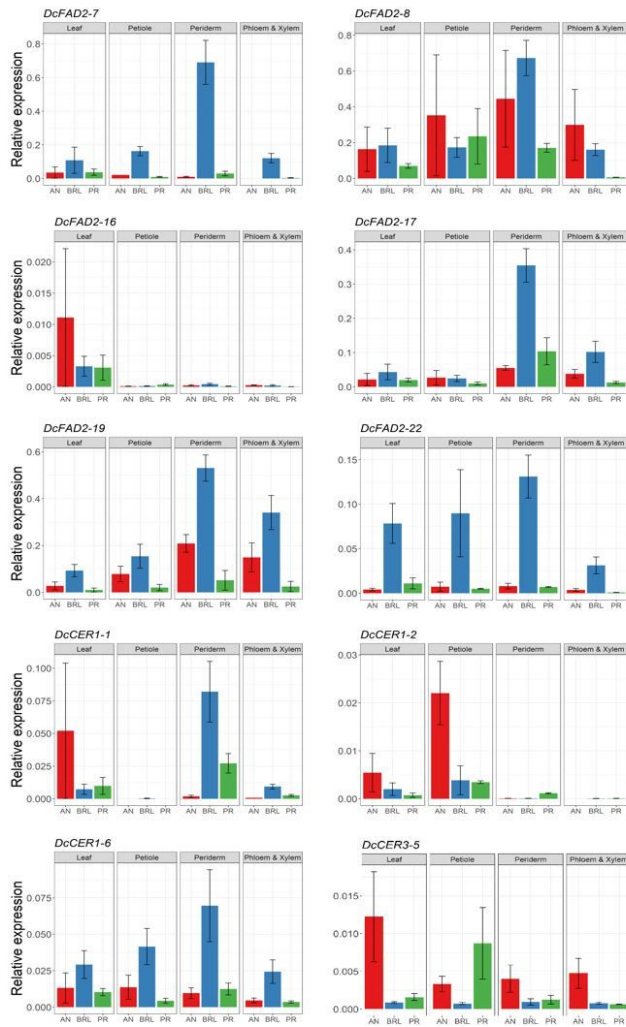


**Figure 2**  
Interval Mapping's logarithm of odds (LOD) scores of markers for PA contents and PA ratio (FaDOH/FaOH), shown alongside the nine carrot chromosomal linkage groups. The LOD threshold for significance ( $p < 0.05$ ) calculated by 5000 permutations is shown as black line. Linkage map is presented in Additional file 3: Figure S2.



**Figure 3**

Phylogenetic tree of predicted *Daucus* CER1/3 proteins (DCAR numbers according [34]; op - optimized prediction) and known plant CER1/3s (AtCER1\_NP\_171723.2 and AtCER3\_NP\_200588.2 from *A. thaliana*; CsCER1\_KJ461885 and CsCER3\_AIE57504.1 from *Camelina sativa*; BnCER1\_KF724897 from *Brassica napus*). The putative tomato CER1 gene Solyc12g100270 [27] is shaded in grey. Multiple sequence alignment was performed by ClustalW using the Lasergene (DNASTAR) software package. A phylogenetic tree was constructed using the Kimura distance formula to calculate distance values and bootstrap analysis (1,000 replicates). Numbers indicate bootstrap replication, and branch length is scaled below the tree indicating the number of amino acid substitutions per 100 amino acids.



**Figure 4**

Tissue-specific expression profiles of FAD2 and CER candidate genes associated with major QTLs on chromosomes chr\_4 and chr\_9. The levels of RNA transcripts were analysed in leaf, petiole, periderm, and a mixture of phloem and xylem of Anthonina (AN), Breeding line (BRL) and Presto (PR). Data represent means of three individual plants of each cultivar with error bars indicating standard error.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Add.file1TableS1.xlsx](#)
- [Add.file2FigureS1.docx](#)
- [Add.file3FigureS2.docx](#)
- [Add.file4FigureS3.docx](#)
- [Add.file5DataS1.docx](#)
- [Add.file6TableS2.docx](#)
- [Add.file7FigureS4.docx](#)
- [Add.file8FigureS5.docx](#)
- [Add.file9FigureS6.docx](#)
- [Add.file10FigureS7.docx](#)
- [Add.file11FigureS8.docx](#)
- [Add.file12FigureS9.docx](#)
- [Add.file13TableS3.xlsx](#)