

***Cannabis sativa* oxylipin biosynthesis: Genome-wide characterization of lipoxygenase, allene oxide synthase, allene oxide cyclase, hydroperoxide lyase, and 12-oxo-phytodienoic acid reductase gene families**

Eli J. Borrego^{1*}, Mariah Robertson¹, James Taylor², Elida Espinoza³

¹Rochester Institute of Technology, Thomas H. Gosnell School of Life Sciences, Rochester, NY, USA

²Texas A&M University, Department of Plant Pathology and Microbiology, College Station, TX, USA

³EnviroFlight, Apex, NC, USA

*** Corresponding author:**

E-mail: eli.borrego@gmail.com

Keywords: *Cannabis*, oxylipins, cannabinoids, green leaf volatiles, jasmonic acid, lipoxygenase, allene oxide synthase, allene oxide cyclase, hydroperoxide lyase, 12-oxo-phytodienoic acid reductase, green leaf volatiles

Abstract

Cannabis sativa is a global multi-billion-dollar cash crop with numerous industrial uses, including in medicine and recreation where its value is largely owed to the production of pharmacological and psychoactive metabolites known as cannabinoids. Often underappreciated in this role, the lipoxygenase (LOX)-derived green leaf volatiles (GLVs), also known as the scent of cut grass, are the hypothetical origin of hexanoic acid, the initial substrate for cannabinoid biosynthesis. The LOX pathway is best known as the primary source of plant oxylipins, molecules analogous to the eicosanoids from mammalian systems. These molecules are chemically and functionally diverse group of fatty acid-derived signals that govern nearly all biological processes including plant defense and development. The interaction between oxylipin and cannabinoid biosynthetic pathways remains to be explored.

Despite their unique importance in this crop, there has not been a comprehensive investigation focusing on the genes responsible for oxylipin biosynthesis in any *Cannabis* species. This study documents the first genome-wide catalogue of the *Cannabis sativa* oxylipin biosynthetic genes and identified 21 *LOX*, five allene oxide synthases (*AOS*), three allene oxide cyclases (*AOC*), one hydroperoxide lyase (*HPL*), and five 12-oxo-phytodienoic acid reductases (*OPR*). Gene collinearity analysis found chromosomal regions containing several isoforms maintained across *Cannabis*, *Arabidopsis*, and tomato. Promoter, expression, weighted co-expression genetic network, and functional enrichment analysis provide evidence of tissue- and cultivar-specific transcription and roles for distinct isoforms in oxylipin and cannabinoid biosynthesis.

This knowledge facilitates future targeted approaches towards *Cannabis* crop improvement and for the manipulation of cannabinoid metabolism.

1 Introduction

For millennia, the diploid dioecious shrub *Cannabis sativa*, has been cultivated for its fiber, grain, and pharmacological properties (1). Originating in Asia and now grown globally (2), *C. sativa* is a multibillion-dollar cash crop (3, 4) with many uses (5) including feed (6), textile (7), biofuel (8), medicine (9), and recreation (10, 11). The latter two of which are due to the production of biologically active metabolites, known as cannabinoids, within the cannabis trichomes (12). Cannabinoids are a large chemical family with 130 currently known distinct chemical species (13) grouped into 10 structural types (14). Undeniably, the best understood are the cannabidiol (CBD) (15) and psychoactive Δ^9 -tetrahydrocannabinol (THC) (16). Given their growing therapeutic uses and social acceptance, efforts are underway to produce chemotypes with explicit cannabinoid content (17, 18), platforms for heterologous production (19), and novel cannabinoid analogues (20).

The understanding of cannabinoid biosynthesis has made strides in recent years (12, 19, 21). Briefly, cannabinoids are synthesized from the convergence of a fatty acid and terpene pathway. Hexanoic acid is activated into hexanoyl-CoA by acyl-activating enzyme 1 and subsequently elongated with three molecules of malonyl-CoA by TETRAKETIDE SYNTHASE (also known as OLIVETOL SYNTHASE. TKS functions in concert with OLIVETOLIC ACID CYCLASE to generate olivetolic acid. Geranyl pyrophosphate produced from the non-mevalonate-dependent isoprenoid pathway (MEP) is used by CANNABIGEROLIC ACID SYNTHASE to prenylated olivetolic acid into cannabigerolic acid (CBGA), the first *bona fide* cannabinoid. Cannabigerolic acid can then be shunted into one of at least three oxidocyclization subbranches to produce tetrahydrocannabinolic acid, cannabidiolic acid, and cannabichromenic acid (CBCA) by THCAS SYNTHASE, CBDA-SYNTHASE, and CBCA SYNTHASE, respectively. Later, THCA, CBDA, CBCA may undergo nonenzymatic decarboxylation into the better-known THC, CBD, or

cannabichromene. Lastly, cannabinoids may also undergo spontaneous rearrangements that are responsible the structural variation found within this chemical family (12, 13, 22).

Despite the recent advances in the understanding of the cannabinoid biosynthetic pathway, the origin of hexanoic acid (also known as caproic acid) remains unclear. It is often the rate-limiting step in heterogenous systems (20, 23) requiring hexanoic acid as a feedstock (19). In *planta*, trichome-specific transcriptomic analysis observed co-expression of genes involved in oxylipin biosynthesis, namely *LIPOXYGENASE (LOX)* and *HYDROPEROXIDE LYASE (HPL)*, with genes previously established in cannabinoid biosynthesis (24, 25). This has prompted the hypothesis that the oxylipin pathway provides the substrate required for cannabinoid biosynthesis. Further, this putative oxylipin-cannabinoid interaction invites exploration to understand the role of the oxylipins in cannabinoid biosynthesis and *Cannabis* biology.

Oxylipins are analogous to the mammalian eicosanoids, and are a large group of oxidized fatty acid-derivatives possessing potent signaling activities that govern a multitude of physiological processes including, growth, development, and defense (26). In plants, the majority of oxylipin biosynthesis occurs via the LOX pathway. It begins with the regio- and stereo-specific incorporation of molecular oxygen at either the 9th - or 13th -carbon of linoleic (C18:2) or linolenic (C18:3) acid by 9- or 13-LOX isoforms (27). The resulting 9- or 13-oxylipins can be fluxed into at least one of seven sub-branches to produce chemically diverse groups of distinct chemical species, including alcohols, aldehydes, divinyl ethers, esters, epoxides, hydroxides, hydroperoxides, ketols, ketones, and triols. Though the biological role for the vast majority of plant oxylipins remains to be deciphered, investigations of selected members of 13-oxylipins, namely jasmonates and green leaf volatiles (GLVs) have spearheaded a framework towards understanding their physiological and ecological roles.

86 Jasmonates are best known for their roles in providing defense against insects and necrotrophic
87 pathogens (28). Jasmonates are cyclopentones produced through the ALLENE OXIDE SYNTHASE
88 (AOS) subbranch (29, 30). Following the oxygenation of C18:3 to 13(*S*)-
89 hydroperoxyoctadecatrienoic acid (13-HPOT) by 13-LOX activity, AOS (29, 31, 32) converts 13-
90 HPOT to 12,13(*S*)-epoxy-octadecatrienoic acid (12,13-EOT) and ALLENE OXIDE CYCLASE
91 (AOC) (33) converts 12,13-EOT to (+)-*cis*-12-oxo-phytodienoic acid (12-OPDA), the first jasmonate
92 in the pathway. LOX, AOS, and AOC are closely associated with each other and participate in
93 substrate channeling during jasmonate biosynthesis (34). Aside from serving as a substrate for
94 downstream reactions, 12-OPDA possesses its own distinct activity (35-37). 12-OPDA is reduced via
95 12-OXO-PHYTODIENOIC ACID REDUCTASE (OPR) (38, 39) to 9*S*,13*S*-OPDA to 3-oxo-2(2'[*Z*]-
96 pentenyl)-cyclopentane-1-octanoic acid (OPC-8:0) and then undergoes three rounds of β -oxidations
97 to produce (+)-7-iso-jasmonic acid (JA) (40). Hexadecatrienoic acid (C16:3) may also serve as
98 substrate for jasmonate biosynthesis, eliminating the need for a round of β -oxidation (41-43). While a
99 multitude of JA derivatives have been identified (29), (+)-7-iso-jasmonoyl-L-isoleucine (JA-Ile) is the
100 best explored, owed to its service as a ligand in receptor-mediated perception and subsequent
101 signaling (44).

102 GLVs are best known as the scent of cut grass and are involved in defense signaling, plant-to-
103 plant and plant-to-insect communication (45). They are C₆ volatile aldehydes, alcohols, and their
104 acetyl esters produced through the HPL subbranch (46, 47). Here, the 13-hydroperoxy fatty acids
105 produced from 13-LOX activity from either C18:2 or C18:3 are cleaved, respectively, by 13-HPL
106 (48-50) into the hexanal or (3*Z*)-hexenal and 12-oxo-(9*Z*)-dodecenoic acid. The latter of which is the
107 progenitor of the traumatin sub-group, of which some members display signaling activity
108 independent of GLVs (51). Additionally, C16:3 may also serve as a fatty acid substrate for GLV
109 biosynthesis, yielding (7*Z*)-10-oxo-decenoic acid in place of traumatin (52). The GLV aldehydes are

reduced to alcohols through reductase (53) and acetylated through acetyltransferase (54). Finally, (3Z)-hexenal can also be isomerized through (3Z):(2E)-ENAL ISOMERASE (55, 56).

Thus, this study sought to elucidate the *Cannabis* oxylipin biosynthetic pathway as a source for cannabinoid substrate and molecular signals important in defense and development. Here, a census was conducted on the *C. sativa* genome to catalog the major oxylipin biosynthetic genes from the LOX, AOS, AOC, and OPR gene families. Respectively, their chromosomal locations, conserved protein domains, genetic structures, and phylogenetic relationships were determined. Collinearity analysis identified several genes in syntenic regions with *Arabidopsis* and tomato. The promoter analysis revealed evidence for tissue- and stimulus-specific transcriptional regulation, which stands in agreement with expression observed from a *Cannabis* transcriptome atlas and publicly-available transcriptomes. Finally, gene co-expressional network and functional enrichment analysis identified genetic modules that implicate oxylipin biosynthetic genes in distinct *Cannabis* physiological processes.

2 Materials and Methods

Gene model identification, protein domain analysis, and subcellular localization prediction. The *Cannabis sativa* representative genome [cs10](#) (57) was surveyed for LOX, AOS, HPL, AOC, and OPR gene models using the NCBI BLASTP algorithm against the *Arabidopsis* sequences as the input queries. The corresponding public transcriptome data was explored through the NCBI Genome Data Viewer (58) to obtain gene models and evidence of transcription and putative translation. Amino acid sequences were examined against the NCBI Conserved Domain Database for the presence of canonical domains using [Batch CD-Search](#) (59, 60). Conserved peptide sequence motifs were determined using [MEME 5.4.1](#) (61, 62) for up to 25 motifs and any number of repetitions.

Subcellular localization prediction analysis was performed using [DeepLoc 1.0](#) (63), [LOCALIZER](#) (64), [Plant-mSubP](#) (65), and [TargetP 2.0](#) (66) through their web servers with default settings.

Multiple sequence alignment, phylogenetic analysis, and genetic structure visualization. Peptide sequences were aligned through the online version of [MAFFT](#) (67, 68) using the FFT-NS-I iterative refinement strategy, leaving gappy regions, and with other settings as default. Phylogenetic trees were generated through the MAFFT website using neighbor-joining, Jones-Taylor-Thornton substitution model, with estimated heterogeneity. Trees were tested with bootstrapping of 1000 resampling and drawn with [FigTree 1.4.1](#). Gene structures were visualized with [TBtools v1.098689](#) (69) using the NCBI *Cannabis sativa* Annotation Release 100.

Promoter analysis of cis-regulatory elements. Putative cis-acting regulatory elements were determined using the 1.5 kb upstream nucleotide sequences of gene models, analyzed with [PlantCARE](#) (70), and promoter diagrams were constructed with TBtools.

Comparative genomic analysis. Gene collinearity between *C. sativa*-*A. thaliana* (reference genome [TAIR10.1](#)) and *C. sativa*-*S. lycopersicum* (representative genome [SL3.0](#)) was performed using MCScanX (71) feature of TBtools with an e-value 1e-10 and 5 BLAST hit cutoffs.

Transcriptomic profiling, regulatory network analysis. RNA-Seq SRP achieved data was retrieved from GEO Datasets, [SRP234963](#) (72) and [SRP168446](#) (73), and trimmed with [fastp v0.23.1](#) (74) using default settings. Transcripts Per Million (TPM) were calculated using [salmon](#) (75) mapping-based mode against a decoy-aware transcriptome file. Corrections were made for fragment-

level GC content and random hexamer primer biases. Transcriptome indices were computed from the cs10 reference transcriptome against its respective genome.

Gene co-expression networks were generated through an iterative weighted correlation network analysis (WGCNA) using [iterativeWGCNA](#) with default parameters (76, 77). Relationships between highly interconnected gene modules were explored through an eigengene score correlation matrix visualized through [Cytoscape](#) (78).

Functional enrichment analysis was performed using [g:Profiler](#) (79) using orthologues identified through [PLAZA 5.0](#) (80) and cross-referenced against the [Kyoto Encyclopedia of Genes and Genomes](#) (81-83) and [Gene Ontology Database](#) (84, 85).

3 Results

Cannabis oxylipin biosynthetic genes. To establish a foundation for *Cannabis* oxylipin biology, the *C. sativa* oxylipin biosynthetic genes were determined. The *C. sativa* representative genome was queried against selected *Arabidopsis* LOX, AOS, AOC, and OPR amino (19)acid sequences with BLASTP to detect candidate *C. sativa* orthologues. The analysis identified 21 LOX, six CYP74, three AOC, and five OPR distinct gene models (Table 1).

Table 1. Genes encoding LOX, CYP74, AOC, and OPR gene families in *C. sativa* and their predicted peptide sublocalization

Annotation	Locus	Chromosome	Location	Sense	mRNA	Protein	AA	Subcellular localization analysis			
								DeepLoc	LOCALIZER	mSubP	TargetP
CsLOX1	LOC115719608	2	48717953..48726423	-	XM_030648706.1	XP_030504566	860	Cytoplasm	-	Cyto	-
CsLOX2	LOC115718693	2	48399032..48404403	-	XM_030647516.1	XP_030503376	955	Plastid	-	Plastid	-
CsLOX3	LOC115724062	9	45035547..45040076	-	XM_030653517.1	XP_030509377	872	Cytoplasm	-	Cyto	-
CsLOX4	LOC115709296	3	1987993..1993483	+	XM_030637370.1	XP_030493230	859	Cytoplasm	-	Cyto	-
CsLOX5	LOC115720291	2	48437789..48443653	-	XM_030649445.1	XP_030505305	848	Cytoplasm	-	Cyto	-
CsLOX6	LOC115721132	2	48616833..48624053	-	XM_030650384.1	XP_030506244	859	Cytoplasm	-	Cyto	-
CsLOX7	LOC115719336	2	48633009..48641095	-	XM_030648324.1	XP_030504184	871	Cytoplasm	-	Cyto	-
CsLOX8	LOC115722276	9	58835655..58849094	+	XM_030651442.1	XP_030507302	868	Cytoplasm	-	Cyto	-
CsLOX9	LOC115721268	2	3929913..3935080	+	XM_030650533.1	XP_030506393	875	Cytoplasm	-	Cyto	-
CsLOX10	LOC115722275	9	58804514..58809886	+	XM_030651441.1	XP_030507301	868	Cytoplasm	-	Cyto	-
CsLOX11	LOC115723988	9	58775266..58790926	+	XM_030653445.1	XP_030509305	873	Cytoplasm	-	Cyto	-
CsLOX12	LOC115718785	2	181580..185609	-	XM_030647611.1	XP_030503471	838	Mitochondrion	Mitochondrion	Endoplasm	mTP
CsLOX13	LOC115712696	4	8039511..8043316	+	XM_030641024.1	XP_030496884	935	Plastid	cTP	Plastid	cTP
CsLOX14	LOC115707105	1	89971772..89976562	-	XM_030634953.1	XP_030490813	931	Plastid	cTP	Plastid	-
CsLOX15	LOC115719612	2	93914675..93922215	+	XM_030648714.1	XP_030504574	929	Plastid	cTP	Cyto	-
CsLOX16	LOC115719614	2	93928923..93935412	+	XM_030648717.1	XP_030504577	926	Plastid	-	Cyto	-
CsLOX17	LOC115720530	2	93947032..93955004	+	XM_030649675.1	XP_030505535	1293	Plastid	-	Plastid	-
CsLOX18	LOC115719613	2	93802540..93809737	+	XM_030648716.1	XP_030504576	929	Plastid	cTP	Plastid	cTP
CsLOX19	LOC115719615	2	93892217..93901839	+	XM_030648718.1	XP_030504578	922	Plastid	cTP	Plastid	-
CsLOX20	LOC115719616	2	93836391..93848870	+	XM_030648719.1	XP_030504579	917	Plastid	-	Plastid	-
CsLOX21	LOC115719617	2	93822415..93829745	+	XM_030648721.1	XP_030504581	909	Plastid	cTP	Plastid	-
CsAOS1	LOC115723125	9	2271373..2273482	+	XM_030652553.1	XP_030508413	525	Plastid	cTP	Plastid	cTP
CsAOS2	LOC115703182	X	6627388..6629208	+	XM_030630703.1	XP_030486563	482	Peroxisome	-	Celmemb	-
CsAOS3	LOC115708244	1	22613465..22615178	+	XM_030636464.1	XP_030492324	483	Plastid	-	Celmemb	-
CsAOS4	LOC115722144	9	55962020..55963631	+	XM_030651275.1	XP_030507135	503	Plastid	-	Celmemb	-
CsAOS5	LOC115703186	X	6682344..6684188	-	XM_030630707.1	XP_030486567	503	Plastid	-	Celmemb	-
CsHPL	LOC115698766	8	51951025..51958560	+	XM_030625939.1	XP_030481799	485	Plastid	-	Plastid	-
CsAOC1	LOC115701324	8	54784220..54785234	+	XM_030629083.1	XP_030484943	183	Cytoplasm	-	Celmemb	-
CsAOC2	LOC115700991	8	54774852..54776132	-	XM_030628671.1	XP_030484531	249	Plastid	cTP	Plastid	cTP
CsAOC3	LOC115699327	8	35236892..35238108	+	XM_030626678.1	XP_030482538	257	Plastid	cTP	Plastid	cTP
CsOPR1	LOC115706391	1	27999501..28002834	-	XM_030634031.1	XP_030489891	452	Peroxisome	cTP	Plastid	-
CsOPR2	LOC115706393	1	27991920..27994962	-	XM_030634033.1	XP_030489893	392	Peroxisome	-	-	-
CsOPR3	LOC115706394	1	27977420..27994954	-	XM_030634034.1	XP_030489894	392	Peroxisome	-	Cyto	-
CsOPR4	LOC115703964	1	27889405..27891439	-	XM_030631189.1	XP_030487049	392	Peroxisome	-	Cyto	-
CsOPR5	LOC115722904	9	35363370..35365841	-	XM_030652270.1	XP_030508130	371	Peroxisome	cTP	Cyto	cTP

The genes were found asymmetrically distributed both across and within the 10 *C. sativa* chromosomes (Fig 1) with the largest concentration of genes was found on Chromosome 1, 2, and 9. Over 60% of the LOX isoforms were found on Chromosome 2, and with the exception of two, were in large gene clusters. *CsLOX1*, 2, 5, 6, and 7 were located in a 327 Kb region in the center of Chromosome 2, while *CsLOX15*, 16, 17, 18, 19, 20, and 21 were near the outer arm in a 152 Kb region (S1 Fig.). A smaller 74 Kb region on Chromosome 9 contained the cluster of *CsLOX8*, 10, and 11 towards the end of a chromosomal arm. *CsLOX9* and 12 were relatively close together on Chromosome 2, while Chromosomes 1, 3, and 4 contained single LOX isoforms, *CsLOX14*, 4, and 13, respectively.

Fig 1. Distribution of the LOX, AOS, HPL, AOC, and OPR gene families across the *C. sativa* genome. Ruler depicts chromosome length.

Similar to the *LOX* genes, members of the CYP74, AOC, and OPR gene families were found in gene clusters at nearly the same proportions. On Chromosome 8, CsAOS2 and 5 were within 57 Kb from each other (S2 Fig.). The other *CYP74* genes, *CsAOS1*, *CsHPL*, *CsAOS1*, and *CsAOS4* were on Chromosomes 1, 8, and 9, respectively. *AOC* genes were found exclusively on Chromosome X with *CsAOC1* and 2 only 10 Kb apart (S2 Fig.). Strikingly, with the exception of *CsOPR5* located on Chromosome 9, the entire OPR gene family was confined to a 113 Kb segment on Chromosome 1.

The *C. sativa* LOX gene family contains 21 members. To verify that candidate LOX gene models encode functional isoforms, their peptide sequences were examined for the presence of the lipid-associated PLAT and catalytic LOX domains archetypical of LOX proteins (27). Amino acids were examined through the NCBI Conserved Domain Database function and all 21 gene models encoded peptides that displayed both domains, providing support that *C. sativa* possesses 21, *bona fide*, LOX isoforms (Fig. 2A). Unexpectedly, the analysis revealed that *CsLOX17* also contained a RVT2 and a RNase H-like domain which may indicate the insertion of a retrotransposon in its upstream region (86, 87).

Fig 2. Phylogenetic, genetic, motif, and domain analysis of the CsLOX gene family. **(A)** Cladogram of peptide sequences and conserved domains. **(B)** Distribution of conserved peptide sequence motifs. Colors are described in legend; x-axis represents length of peptides in amino acids. **(C)** Diagram of

genetic structure. Blue bars, orange bars, and gray lines represent untranslated regions, exons, and introns, respectively; x-axis represents length of gene in nucleotides. **(D)** Polar cladogram depicting evolutionary relationship with gene families of selected species. Node labels show confidence values from 1000 bootstrap replications.

Analysis of conserved motifs showed that the majority of the LOX isoforms contain similar composition and distribution of amino acid sequence patterns (Fig. 2B). Nearly all isoforms displayed the same 13 motif arrangement, typical of LOX proteins, on the N-terminus, a highly conserved region which requires stringent composition for enzymatic activity. The exception was CsLOX12 which was absent of motif 13. Aside from CsLOX1 and 14, the arrangement of the first 5 motifs on the C-terminus was broadly consistent across all the proteins. However, several isoforms possessed regions where no motifs were detected. In LOX proteins, these regions typically contain highly divergent, transient peptide sequences that direct subcellular localization (88). In support of this notion, their amino acid sequences were tested against four subcellular prediction algorithms, DeepLOC, LOCALIZER, mSUBP, and TargetP (Table 1). A consensus of plastid localization was reached for only CsLOX13 and CsLOX18, along with an agreement of the majority of the software localizing CsLOX2, CsLOX14, CsLOX15, CsLOX17, CsLOX18, CsLOX19, CsLOX20, and CsLOX21 to this organelle. CsLOX1 and CsLOX3-11 were predicted to be found in the cytoplasm. Interestingly, CsLOX12 was predicted to be associated with the mitochondria.

Motifs 17 and 20 located within the LOX domain showed the largest variability across the LOX isoforms. The former was absent from CsLOX3, 13, and 14, the latter was missing from CsLOX19 and 20, and both were missing from CsLOX12, 15, 16, 17, and 18, prompting the speculation of divergent catalytic activity among even closely related members (89).

The *C. sativa* LOX gene models ranged in size from around 4 to 16 Kb and over 75% possessed eight introns (Fig. 2C). A recurrent observation in the genetic structures of some *LOX* family members was the presence of a large first or second intron, e.g., CsLOX1, 6, 7, 8, 11, and 20 comprising of about 50 to 60% of the entire gene. No data is yet available for alternative splicing or transcript variants of these genes to provide insight into the role of these features.

Plant LOXs are typically grouped as 9- or 13-LOXs according to their major enzymatic product and isoforms with similar activity display higher sequence similarity and evolutionary relationships (27). To designate the classification of *C. sativa* LOXs, their peptide sequences were compared to dicot and monocot plant species with well-characterized LOX protein families, namely, *A. thaliana*, *O. sativa*, *S. lycopersicum*, and *Z. mays* (Fig. 2D). CsLOX1, 2, and 3 grouped with the *Arabidopsis* 9-LOXs, AtLOX1 and 5, respectively. Remarkably, eight *Cannabis* LOXs peptides, CsLOX4-11, formed a *Cannabis* specific clade nested within monocot and dicot 9-LOXs. CsLOX12 was difficult to place within the tree, a result which was likely related to the absence of the three variable peptide motifs in its catalytic domain (Fig. 2B). CsLOX13 and 14 clustered with clades containing dicot 13-LOXs. Interestingly, seven *Cannabis* LOX peptides, CsLOX15-21 grouped with the 13-LOX clade, and several of which may represent *Cannabis*-specific 13-LOXs. Thus, *C. sativa* possess eleven 9-LOXs (CsLOX1 – 11), nine 13-LOXs (CsLOX13 – 21), and one that remains to be empirically determined (CsLOX12).

The *C. sativa* CYP74 gene family contains six members. AOS and HPL belong to the atypical CYP74 clade of the P450 family involved in hydroperoxyl fatty acid rearrangement or dismutation (90). Amino acid sequence analysis found a single CYP74 domain in all six *Cannabis* candidate CYP74s and four distinct patterns of conserved sequence motifs (Fig. 3A, B). The pattern variations were most pronounced towards the N-termini where each possessed a distinctive starting motif. The

N-terminus of CYP74s typically possess transient peptide signals involved in directing their subcellular localizations towards plastids, and while subsequent analysis prediction supported this notion in CsHPL and CsAOS1, a consensus among the four software prediction tools could not be reached for CsAOS2-5 (Table 1). For gene structure analysis, CsHPL was notable for the presence of two introns (Fig. 3C), an unusual feature of CYP74 family members, one of which comprised the majority of its gene model sequence.

Fig 3. Phylogenetic, genetic, motif, and domain analysis of the *C. sativa* CYP74 gene family. **(A)** Cladogram of peptide sequences and conserved domains. **(B)** Distribution of conserved peptide sequence motifs. Colors are described in legend; x-axis represents length of peptides in amino acids. **(C)** Diagram of genetic structure. Blue bars, orange bars, and gray lines represent untranslated regions, exons, and introns, respectively; x-axis represents length of gene in nucleotides. **(D)** Polar cladogram depicting evolutionary relationship with gene families of selected species. Node labels show confidence values from 1000 bootstrap replications.

While many CYP74 isoforms display multifunctionality (91), catalyzing varying proportions of AOS, EAS, DES, and HPL products, the CYP74A and CYP74B clade members show predominantly AOS and HPL activities, respectively. To understand the potential function of *C. sativa* CYP74 family members, a phylogenetic analysis assessed their relationship within the relatively well-characterized CYP74 subclades from diverse plant species (Fig. 3D). CsAOS1 and CsHPL were grouped within the CYP74A and CYP74B clades, respectively. However, the placement of CsAOS2-4 proved challenging. Nonetheless, CsAOS2 and 3 and CsAOS4 and 5 grouped as pairs roughly within CYP74C.

Taken together, this analysis suggests that *C. sativa* contains at least two dedicated CYP74 enzymes that, given their phylogenies and subcellular localization, are specialized for 13-AOS and HPL activity, thus capable of providing substrate for JA or GLV biosynthesis. Cannabis also contains four CYP74 isoforms with unassigned activities.

The *C. sativa* AOC gene family contains three members. AOC provides steric hindrance for the stereospecific cyclization of allene oxide to 12-OPDA, the parent jasmonate species. All three CsAOC isoforms contain an AOC domain and similar amino acid sequence patterns on the last six motifs of their C-termini (Fig. 4A, B). CsAOC2 and CsAOC3 displayed substantial variability in predicted motifs at their N-termini, likely corresponding to putative plastid transient peptide sequences (Table 1). CsAOC1 was predicted to localize outside of the plastids. The intron-exon distribution of this gene family differed mildly with CsAOC1 and 2 having three exons while CsAOC3 contained one (Fig. 4C). To understand the evolutionary relationship of the *C. sativa* AOC gene family, their peptide sequences were analyzed phylogenetically with AOC members from other plant species (92). CsAOC1 and 2 clustered close to each other within a dicot specific clade and CsAOC3 grouped into a separate dicot-specific clade that contained all *Arabidopsis* AOCs (Fig. 4D).

Fig 4. Phylogenetic, genetic, motif, and domain analysis of the CsAOC gene family. **(A)** Cladogram of peptide sequences and conserved domains. **(B)** Distribution of conserved peptide sequence motifs. Colors are described in legend; x-axis represents length of peptides in amino acids. **(C)** Diagram of genetic structure. Blue bars, orange bars, and gray lines represent untranslated regions, exons, and introns, respectively; x-axis represents length of gene in nucleotides. **(D)** Polar cladogram depicting evolutionary relationship with gene families of selected species. Node labels show confidence values from 1000 bootstrap replications.

299

300 **The *C. sativa* OPR gene family contains five members.** All five identified *C. sativa* OPR
301 candidates were found to contain the conserved Old Yellow Enzyme-like domain necessary for their
302 activity (Fig. 5A) (93). CsOPR1-4 had nearly identical patterns of amino acid sequence motifs and
303 were largely similar to the motif arrangement of CsOPR5, however, the latter isoform lacked four
304 conserved motifs (Fig. 5B). In regard to their genetic structure, a substantial first intron was
305 identified in CsOPR3, resulting in a gene model roughly three times larger than the other isoforms
306 (Fig. 5C). It is important to note that the current gene model of CsOPR3 appears chimeric with
307 CsOPR2 in the current *C. sativa* representative genome (S2 Fig.).

308

309 **Fig 5.** Phylogenetic, genetic, motif, and domain analysis of the CsOPR gene family. **(A)** Cladogram
310 of peptide sequences and conserved domains. **(B)** Distribution of conserved peptide sequence motifs.
311 Colors are described in legend; x-axis represents length of peptides in amino acids. **(C)** Diagram of
312 genetic structure. Blue bars, orange bars, and gray lines represent untranslated regions, exons, and
313 introns, respectively; x-axis represents length of gene in nucleotides. **(D)** Polar cladogram depicting
314 evolutionary relationship with gene families of selected species. Node labels show confidence values
315 from 1000 bootstrap replications.

316

317 To investigate the evolutionary relationship between the *C. sativa* OPR gene family members,
318 the amino acid sequences were compared to OPR family members from dicots and monocots.
319 CsOPR1-4 formed a phylogenetic clade with the JA-producing AtOPR3. CsOPR5, however, was
320 grouped with Type-II OPRs (Fig. 5D). Taken together, *C. sativa* contains four Type-I OPR isoforms
321 that are likely involved in classical JA biosynthesis.

322

323 **Some LOX, CYP74, and AOC members are syntenic across *Cannabis*, *Arabidopsis*, and tomato.**

324 To infer the ancestry of the *C. sativa* oxylipin pathway, a genetic collinearity analysis was performed
 325 to compare physical co-localization of oxylipin biosynthetic genes in the genomes of *C. sativa*
 326 against two dicot genomes with well characterized oxylipin biosynthetic pathways, *A. thaliana*, and
 327 *S. lycopersicum* (tomato) (Fig. 6). Remarkably, three LOX genes displayed collinearity across all
 328 three species. Of the 9-LOXs, *CsLOX2* was in a syntenic region with *AtLOX1* and *SILOX5* (also
 329 known as TOMLOXE). Two 13-LOXs display collinearity, *CsLOX14* with both *AtLOX3* and *AtLOX4*
 330 and with *SILOX4/LOXD*, while *CsLOX13* was collinear with *AtLOX6* and *SILOX11*. Two *C. sativa*
 331 CYP74 genes were in syntenic regions with tomato genes. *CsAOS2* matched with *SIDES* (also known
 332 as *LeDES*) and *CsHPL* matched with both *SIHPL* and *SLAOS3*. Two *CsAOC* genes were collinear
 333 with either plant species although with dissimilarities. *CsAOC2* was in a syntenic gene block with
 334 *SLAOC* and *CsAOC3* matched with both *AtAOC1* and *AtAOC4*.

335

336 **Fig 6.** Oxylipin biosynthetic gene collinearity between *A. thaliana*, *C. sativa*, and *S. lycopersicum*.

337 The grey lines represent collinear genetic blocks between the genomes of *A. thaliana* and *C. sativa*
 338 and between *C. sativa* and *S. lycopersicum*. Bars represent chromosomes and the black, blue, and
 339 orange lines represent the collinear pairs of LOX, CYP74, and AOC gene families, respectively.

340

341 **Promoters of *C. sativa* oxylipin biosynthetic genes carry tissue-specific, conditional, and**

342 **phytohormone-inducible cis-acting regulatory elements (CAREs).** To elucidate the involvement

343 of *Cannabis* oxylipin genes in diverse developmental and stress responses, an *in silico* promoter

344 analysis was conducted to identify cis-regulatory elements involved in plant tissue-specific

345 expression (endosperm, meristem, mesophyll, seed), adaptation to environmental conditions

(anaerobic, circadian rhythm, drought, light, and low temperature), and phytohormone signaling (ABA, GA, IAA, JA, and SA).

All promoters varied considerably in the content, arrangement, and position of their CAREs (Fig. 7 and S1 Table). Overall, the oxylipin biosynthetic gene promoters possessed CARE motifs with lengths of 8 to 32 nucleotides. For the LOX gene family, the promoter of *CsLOX12* gene was found to contain the most regulatory elements with 32 and followed by *CsLOX16* and *21*, each with 28, while *CsLOX2*, *7*, *11*, and *17* displayed the least, with around 10 each. The *Cannabis*-specific 9-LOX group (*CsLOX4-11*) displayed the most motifs followed by the amplified GLV-producing 13-LOXs (*CsLOX15-21*). With the exception of *CsLOX5*, all contained several regulatory elements involved in conditional responses, though few contained elements involved in tissue expression. CAREs involved in light responses dominated most promoters with all genes having at least three such motifs. Interestingly, *CsLOX12* light-responsive motifs constituted over 60% of all its identified motifs. Promoters associated with the presence of phytohormone inducible cis-regulatory elements varies considerably, spanning from 0 to 8 regulatory elements. With the exception of *CsLOX2*, *4*, *7*, *8*, and *17*, all had at least one CARE related to ABA responses. Incidentally, *CsLOX2*, *4*, *7*, *8*, and *17* were also among the few members that showed motifs involved in GA responses, suggesting a role mediated by ABA-GA antagonism (94). SA-responsive motifs were limited and predominately associated with the 9-LOXs.

Fig 7. *Cis*-acting regulatory element distribution across *C. sativa* oxylipin biosynthetic gene promoters. (A) Physical distribution of motifs throughout a 1.5 kb region upstream of corresponding gene model. Gene symbols are presented on the left with grey line representing promoter regions with motif positions. Legend describes motifs associate with tissue-specific expression, environmental responses, and phytohormone inducibility. (B) Summary of *cis*-acting regulatory

motif content. Heatmap depicts number of motifs identified in gene promoter region. The y-axis depicts *C. sativa* oxylipin biosynthetic genes and x-axis depicts associated condition for motif. Cells values are number of motifs identified and are colored according to legend.

Of the CYP74 gene family, the largest number of CAREs were identified on promoters of CsAOS4 and 5, followed by CsHPL (Fig. 7). Relatively few were found on the putative JA-producing CsAOS1, namely those associated with anaerobic conditions, light, and ABA signaling responses. Unlike the pattern observed for LOXs, the disparity between light-related motifs and other CAREs was not as dramatic in the CYP74 family, with the exception of the CsAOS4 and 5 pair. Nearly all CYP74s had ABA-responsive motifs, with the exception of CsAOS3 and CsHPL which were also the only ones to possess GA-related motifs. CsAOS3, 4, and CsHPL each had two JA-related motifs, while CsAOS1 had none; this suggests a limited contribution of positive feedforward control of JA biosynthesis.

Analysis of the AOC gene family found only a single tissue-type related to CAREs across the promoters of all members: CsAOC1 for meristem expression. While non-light conditional-response motifs were the most numerous in CsAOC2 compared to the other two isomers, the opposite was seen for light-responsive motifs, where only three were found in CsAOC2. This was a sharp contrast with CsAOC1, which showed more than twice the number of these CAREs compared to its paralogues (Fig. 7). No AOC was found to have SA-related, and only CsAOC2 had JA-related motifs, while all had ABA-related motifs.

Promoters of the OPR gene family contained similar patterns of CAREs compared with the other gene families, albeit with reduced CARE content. No tissue-related CARE was detected for any OPR member. However, all other members possessed motifs for anaerobic responses, with the exception of CsOPR1 (Fig. 7). ABA motifs were found in CsOPR1, CsOPR2, and CsOPR5. CsOPR5

was also the only member to show GA motifs, and JA motifs were found in CsOPR1, 3, and 4. The notable difference in CARE patterns found within the JA-producing OPRs (CsOPR1-4), suggests that the 12-OPDA to JA balance is maintained in *Cannabis* through transcriptional control of distinct OPR members responses to the plant's environment and during phytohormone signaling (95).

***C. sativa* oxylipin biosynthetic genes show tissue-, developmental-, and cultivar-specific expression.** Promoter analysis suggested that expression of *C. sativa* oxylipin biosynthetic genes would display distinct patterns. To understand the transcriptional profile of oxylipin biosynthetic genes and to elucidate their role in *Cannabis* development and physiology, the gene expression of 23 tissues from a *C. sativa* transcriptome atlas (72) were examined (S2 Table). Overall, the major contributor associated with levels of expression appeared to be the identity of the specific gene family member, i.e., patterns of expression levels were generally consistent across the tissues examined (Fig. 8A). Moreover, several members in all gene families had few if any detected transcripts across all tissues, suggesting a limited role for those isoforms under basal conditions or in the tissues profiled.

Fig 8. Heatmaps showing expression of oxylipin biosynthetic genes from the *C. sativa* gene expression atlas (A) or female trichomes of diverse marijuana lines (B). Cells values are rounded TPM values and are colored according to legend

Of the 9-LOXs, *CsLOX1*, 2, 4, 7 and 8 showed pronounced levels of expression across most tissues. *CsLOX4* is particularly notable as the highest expressed oxylipin biosynthetic gene analyzed, by one order of magnitude, in root and trichome tissues (Fig. 8A), highlighting the probable

importance of this isoform grouping within the *Cannabis*-specific 9-LOX clade (Fig. 2). Curiously, *CsLOX9* and *10* had minimal levels of expression in all but root tissues, where their levels were 20-fold higher compared to other tissues. With respect to the 13-LOXs, with the exception of *CsLOX13* and *18*, all had elevated expression in reproductive-related tissue. However, under these basal conditions, only *CsLOX14* showed modest expression in roots. Remarkably, despite possessing the greatest number of identified CAREs (Fig. 7B), nearly no expression was detected for *CsLOX12* (Fig. 8A) which implies a non-basal role for this gene.

In regards to the CYP74-related genes, the likely non-JA producing *CsAOS2* displayed the most abundant expression levels, predominantly in early stages of flower and trichome development (Fig. 8A, S3). The JA-producing, *CsAOS1*, also showed its greatest levels of expression in the female flowers, while minimally expressed in other tissues. While *CsAOS5* showed modest levels of expression across tissues, transcripts of its closest paralogue *CsAOS4* were only detected at low levels in few tissues. For the GLV- and cannabinoid-producing *CsHPL*, expression remained consistent across tissues, disregarding low levels in root tissues (Fig. 8A). Interestingly, *CsHPL* expression remained relatively consistent across all stages of flower and trichome development (S3 Fig.).

In the majority of tissues, *CsAOC2* showed the most uniform levels of expression from this gene family (Fig. 8), suggesting this member is the prominent isoform involved in JA biosynthesis under during basal conditions. *CsAOC1* and *CsAOC3* showed moderate to modest levels of expression across most tissue types with several exceptions, prompting the idea that these members participate in inducible processes.

Only *CsOPR2* and *5*, displayed mentionable levels of expression. *CsOPR2* transcript levels were an order of magnitude higher than the other Type II paralogues, suggesting this is the major JA-producing isoform under basal conditions (Fig. 8). It is also notable that the majority of its expression

was in tissues related to flower structures. The sole Type I OPR, *CsOPR5* displayed even greater levels of expression in flower tissues, especially developing trichomes, raising the possibility for the involvement of *CsOPR5* during reduction or detoxification of metabolites during trichome development.

To understand the variation in gene expression across *Cannabis* diversity, transcriptomes of trichomes from nine cultivars of mixed ancestry (73) were profiled for their expression of oxylipin biosynthetic genes (S3 Table). With the exception of expression of *CsLOX15*, of which transcript levels showed nearly a bimodal distribution across the cultivar tested, i.e., Black Berry Kush, Cherry Chem, Mamma Thai, and Terple possessed half of the expression compared to the five other lines (Fig. 8B). Few differences were observed in the other oxylipin biosynthetic gene expression. In particular, decreased expression was seen in *CsLOX14*, *CsLOX16*, and *CsOPR2* from Cherry Cham, Terple, and Black Berry Kush, respectively, relative to the other *Cannabis* lines. On the other hand, Mama Thai and Black Berry Kush showed elevated levels for transcripts of *CsLOX4* and *CsAOCL*, respectively, and trichomes of White Cookies had increased expression of both *CsAOS3* and *CsAOCL*. Interestingly, in sharp contrast to expression levels seen in the transcriptome atlas, transcripts of *CsLOX15* dominated in the trichomes of these cultivars followed by those of *CsLOX4*. In agreement with the oxylipin-derived hexanoic acid hypothesis (96), the expression of *CsHPL* levels were also two- to 10-fold higher compared to the other CYP74s, likely owed to the selection of these cultivars for high cannabinoid production.

***Cannabis* oxylipin biosynthetic genes are found in gene networks associated with stress responses, growth, and development.** To assign molecular and cellular functions to the *Cannabis* oxylipin biosynthetic genes, an iterative weighted gene co-expression network analysis was performed on a *Cannabis* transcriptome atlas (72) to identify clusters of genes with highly correlated

expression. To inform their roles in physiological processes, the resulting gene modules were analyzed for functional term enrichment. In total, 23 modules containing at least one *Cannabis* oxylipin biosynthetic gene were identified with seven modules containing multiple genes (Fig. 9A). The four largest modules contained between 1130-1852 genes and comprised nearly 30% of the current *C. sativa* gene models; one module contained *CsAOS5*, while the other three modules contained multiple oxylipin biosynthetic genes: *CsLOX3* and *10*, *CsOPR3* and *4*, and *CsLOX5*, *9* and *CsOPR1*. The 15 smallest modules had under 200 genes each and together accounted for only 5% of the *C. sativa* transcriptome.

Fig 9. Weighted co-expression genetic network derived from the *C. sativa* transcriptome expression atlas. **(A)** Clusters represent modules of highly connected genes containing at least one oxylipin biosynthetic gene. **(B)** Top ten terms for functional enrichment analysis of gene modules containing oxylipin biosynthetic genes: Gene Ontologies for Molecular Functions (MF), Biological Processes (BP), and Cellular Components (CC) and KEGG.

Functional enrichment analysis of the entire oxylipin gene-associated co-expression network for Gene Ontology and KEGG identified terms related to molecular functions, biological process, and cellular component orthologs (Fig. 9B). Among the top 10 terms for molecular functions, 60% were related to nucleic acid binding and transcription. Of the biological processes, 80% were related to development or physiological responses. For cellular components, 40% and 30% described organelle- and membrane-related terms, respectively. KEGG terms described secondary metabolism, particularly alpha-linolenic acid metabolism and branch-chain amino acid degradation. As expected, individual modules were enriched for differential functional terms (S4 Table). Taken together, these

results suggest that even closely related *Cannabis* oxylipin biosynthetic genes are specialized for discrete functions.

4 Discussion

Enthusiasm is growing to understand the physiological and ecological functions of plant oxylipins (90). JA has long served (97) as the model oxylipin, and over the last fifty years, many aspects of its biosynthesis, metabolism, signaling, and activities have started to be defined (98-101). However, this knowledge has been generated predominantly in the model plant species, *Arabidopsis*, and investigations into oxylipin biology in crop species have been launched only in recent years. Here, we cataloged the major oxylipin biosynthetic genes families of the agro-economically important crop species, *C. sativa*, described their expression patterns, and delineated connections between their putative physiological functions.

LOX gene family. This study adds Cannabaceae to the collection of plant families with described LOX gene families, a group which so far includes Araceae: duckweed (102); Actinidiaceae: kiwi (103); Brassicaceae: *Arabidopsis thaliana* (104-106), radish (107), and turnip (108); Caricaceae: papaya (106); Cucurbitaceae: cucumber (109, 110), melon (111), and watermelon (112); Fabaceae: *Medicago truncatula*, peanut (113), and soybean (114); Malvaceae: cotton (115) Muscaceae: banana (116); Poaceae: maize (117), rice (105), *Setaria italica* (118), sorghum (119); Polygonaceae: buckwheat (120); Rhamnaceae: Jujubie (121); Rosaceae: apple (122), peach (123), pear (124); Salicaceae: poplar (125); Solanaceae: pepper (126) and tomato (127, 128); Theaceae: *Camellia sinensis* (129); and Vitaceae: grape (130).

Of the 21 *C. sativa* LOX isoforms identified (Fig. 1), 11 are likely responsible for the production of 9-oxylipins, nine are producers of 13-oxylipins, and one member (CsLOX12) remains to be empirically determined. Of the hallmark 9-LOXs AtLOX1 and AtLOX5, two and one *C. sativa* LOX members grouped with each respectively, while eight formed their own clade. Within 9-LOX phylogenies, monocots typically display their own grouping (102, 116, 118), while the divergence of 9-LOXs in dicots is less common; yet still observed in Cucurbitaceae (111), Salicaceae (125), and Vitaceae (130). Interestingly, seven members were identified in nearly one large tandem gene array (S1 Fig.), which fell within the clade containing both monocot and dicot isoforms required for GLV biosynthesis (45, 131). Similar substantial gene amplification of this LOX clade has been found to various degrees, in Cucurbitaceae (109-112), Salicaceae (125), and Vitaceae (130). As *AtLOX2* was not found to be syntenic across the Rosids (106), the gene duplications likely occurred independently in ancestors following the divergence of Rosales, Cucurbitales, Malpighiales, and Vitales.

Proteomic analysis (132) has verified that several lipoxygenases are found with spatial specificity in *Cannabis* flowers and components of their glandular trichomes. Of particular interest was the presence of isoforms from the tandem duplicated 13-LOXs, namely, CsLOX15, 16 and 17 within the trichome head. It is tempting to speculate that these gene duplications may serve to maintain the pool of hexanoic acid substrate available for cannabinoid biosynthesis. In *Cannabis*, tandem gene arrays have previously been implicated in regulating lipid biosynthesis, particularly the ratios of volatile terpenes that subsequently determines the scent of specific cultivars (133). This amplification of the 13-LOX paralogues in a tandem gene array may be due to the origin of the representative *C. sativa* genome (57) from CBDRx, a high CBD producing line. The ubiquity of tandem gene arrays among other *Cannabis* species and cultivars remains to be observed, especially in

regards to the agronomic characteristics that drive the selection of economically important cultivars. Interestingly, the 9-LOXs, CsLOX4 and 7 were also found within the flower and trichomes where they are likely to contribute to the 9-oxylipin volatiles of *Cannabis* (134).

CYP74 gene family. Despite its importance in oxylipin metabolism, few studies have examined CYP74 gene families from a genome-wide perspective (135-137). In *C. sativa*, six CYP74 isoforms were identified. Specifically CsAOS1 and CsHPL, were distinctly found within the JA- and GLV-producing CYP74 clades, respectively (Fig. 3D). Both contained a transient peptide signal sequence predicting localization in plastids (Table 1), although the unique motifs found at their N-terminus suggest an association with different sub-plastid locations. This is similar to tomato where LeAOS was found with the inner chloroplast envelope while LeHPL was targeted to the outer plastid envelope (138). Four isoforms were found with the CYP74C subfamily, a more irregular group of enzymes with members possessing varying levels of AOS, EAS, and/or HPL activity (139). It is interesting to note that CsAOS2 was syntenic to the DES of tomato (Fig. 6), however, a comprehensive survey and validation of *Cannabis* oxylipins remains to be performed to understand if this species indeed produces divinyl ethers. Ultimately, these four isoforms will require biochemical characterization, as even a single amino acid change can result in a new function or protein activity (140, 141). Four CYP74 peptides (CsAOS1,2,5 and CsHPL) have been detected in association with *Cannabis* flowers, and of these, CsHPL was the only isoform found within trichome heads and stalks (132).

AOC gene family. AOC activity yields the first JA and is the last enzymatic step of JA biosynthesis performed in the chloroplast (101). Three members of the AOC gene family were identified in *C. sativa*, a number consistent with four of *Arabidopsis* (33), six of soybean (142), and two of maize

(26). Isoforms typically display organ- and tissue-specific expression and function as heterodimers (143). Only CsAOC2 and CsAOC3 were predicted to localize to plastids (Table 1), supporting the function of these isoforms in JA-biosynthesis. Remarkably, CsAOC3 showed synteny with both *Arabidopsis* (*AtAOC1* and *AtAOC4*) and tomato (*SLAOC*) orthologues. In *Arabidopsis*, heteromers containing AtAOC1 and AtAOC4 display the greatest enzymatic activity (144), suggesting an evolutionary advantage of this conserved genomic region for JA production.

OPR gene family. A larger variability in OPR gene content exists across plant species analyzed so far: three in *Arabidopsis* (38), five in pea (145), 5 in watermelon (146), 8 in maize (147), 10 in cotton (148), 13 in rice (149), and 48 in wheat (150). This study identified one Type I OPR and four Type II OPRs in the *C. sativa* genome. Both subgroups reduce α , β -unsaturated double bonds and are involved in reactive electrophilic species (RES) detoxification (151, 152). While Type II OPRs are well-characterized for their peroxisomal role in reducing 12-OPDA to OPC8:0 during JA biosynthesis, Type I OPRs are less understood. They have been shown to reduce 4, 6-trinitrotoluene (TNT) (93) and were recently demonstrated to be involved in producing JA through reduction of an endogenous cyclopentenone JA-analog (42). Tandem duplication has been implicated as playing a major role in the size of OPR gene families across plant species (153) and interestingly, while all Type II OPRs were found together in a gene cluster, no OPR displayed synteny across the species tested, suggesting OPR gene duplication was a relatively recent event in the *C. sativa* ancestor.

Conclusion. While the LOX pathway is the hypothetical origin of the hexanoic acid moiety used for cannabinoid biosynthesis (12, 19, 24, 25, 96), the precise mechanisms for its production are largely unknown. Here, we provide a comprehensive description of the *C. sativa* LOX pathway, including

the major oxylipin biosynthetic gene families, to establish a working model (Fig. 10) for further investigations.

Fig. 10. Working model of *C. sativa* oxylipin biosynthetic enzyme isoforms and pathways involved in cannabinoid, GLV, and JA biosynthesis. Abbreviations: [enzymes]: allene oxide cyclase (AOC), allene oxide synthase (AOS), 2-alkynal reductase (2-AR), hydroperoxide lyase (HPL), lipoxygenase (LOX), methylerythritol 4-phosphate pathway (MEP), 12-oxo-phytodienoic reductase (OPR); [metabolites]: (9Z, 11E, 13S, 15Z)-12,13-epoxy-9,11,15-octadecatrienoic acid (12,13-EOT), (9Z, 11E, 13S)-13-hydroperoxy-9,11-octadecadienoic acid (13-HPOD), (9Z, 11E, 13S, 15Z)-13-hydroperoxy-9,11,15-octadecatrienoic acid (13-HPOT), cannabigerolic acid (CBGA), linoleic acid (C18:2), alpha-linolenic acid (C18:3), jasmonic acid (JA), 12-oxo-phytodienoic acid (12-OPDA), 8-[3-oxo-2-cis-(25)cyclopentyl]octanoic acid (OPC8:0).

Within this context, it will become important to understand the co-localization of the substrates with their corresponding enzymes. LOX activity depends on a 1,4-dipentadiene structure found in either C18:2 or C18:3 (27), both of which, in tomato, account for the most abundant fatty acids in its trichomes (154). However, while HPL-cleavage of 13-hydroperoxydioenoic acid would generate a saturated C₆-aldehyde, the GLV-producing 13-LOXs and HPL are likely localized to the C18:3-rich chloroplasts (155, 156). Additionally, the GLV-producing AtLOX2 oxygenates C18:3 more efficiently than C18:2 (104), which together may explain why unsaturated GLVs prevail as the major C₆ volatile in plants (157).

Enzymatic production of hexanoic acid from the C18:3 HPL-product, Z-3-hexenal is unexplored, but would reasonably require distinct enzymatic reactions. Chemical or enzymatic isomerization by (3Z):(2E)-enal isomerase would generate (2E)-hexenal (46, 55) which could be

reduced to hexanal by 2-alkenal reductase (158) or by OPR (159). The conversion of the fatty aldehyde to fatty acid likely requires an aldehyde dehydrogenase (160), the product of which would be readily diffused across cell membranes (161). In an alternative scenario, 13-LOX may oxygenate C18:2 directly, however this would presumably require the cytosolic localization of both 13-LOX and HPL. This strategy would take advantage of the relatively high C18:2 content in the inflorescences of some *Cannabis* cultivars (162).

Understanding and applying knowledge of the *Cannabis* oxylipin biosynthetic pathway is expected to provide novel environmentally friendly approaches towards improving the crop and its desirable consumer traits. Similar technologies in maize and soybeans in development aim at increasing disease resistance (163), drought tolerance (164), and seed flavor (165). Exploiting the diversity in *Cannabis* cultivars via oxylipin biosynthetic gene expression and variability in the predominant oxylipin biosynthetic pathway branch across cultivars and species (e.g., *C. sativa* vs *C. indica*) should also provide opportunities for manipulating cannabinoid content and other traits through marker-assisted selection or identification of superior alleles.

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

EB conceived and designed the study. EB, MR, JT, and EE collected the data. EB, JT, and EE performed analysis. EB wrote the manuscript with input from EE and all authors reviewed and agree with the final version.

7 Acknowledgments

Thank you to Drs. Michael Kolomiets and Zachary Schultzhause for helpful discussions and to Dr. Michael Osier for use of computational resources.

8 Funding

We would like to thank the Thomas H. Gosnell School of Life Sciences (GSoLS) and the College of Science (COS) at the Rochester Institute of Technology (RIT) for ongoing support.

9 References

1. Bonini SA, Premoli M, Tambaro S, Kumar A, Maccarinelli G, Memo M, et al. Cannabis sativa: A comprehensive ethnopharmacological review of a medicinal plant with a long history. *J Ethnopharmacol.* 2018;227:300-15.
2. Kovalchuk I, Pellino M, Rigault P, van Velzen R, Ebersbach J, Ashnest JR, et al. The Genomics of Cannabis and Its Close Relatives. *Annu Rev Plant Biol.* 2020;71:713-39.
3. de la Fuente A, Zamberlan F, Ferrán AS, Carrillo F, Tagliazucchi E, Pallavicini C. Relationship among subjective responses, flavor, and chemical composition across more than 800 commercial cannabis varieties. *Journal of cannabis research.* 2020;2(1):1-18.
4. Torkamaneh D, Jones AMP. Cannabis, the multibillion dollar plant that no genebank wanted. *Genome.* 2021;99(999):1-5.
5. Rehman M, Fahad S, Du G, Cheng X, Yang Y, Tang K, et al. Evaluation of hemp (*Cannabis sativa* L.) as an industrial crop: a review. *Environ Sci Pollut Res Int.* 2021;28(38):52832-43.
6. Callaway JC. Hempseed as a nutritional resource: An overview. *Euphytica.* 2004;140(1-2):65-72.

- 652 7. Duque Schumacher AG, Pequito S, Pazour J. Industrial hemp fiber: A sustainable and
653 economical alternative to cotton. *Journal of Cleaner Production*. 2020;268:122180.
- 654 8. Finnan J, Styles D. Hemp: A more sustainable annual energy crop for climate and energy
655 policy. *Energy Policy*. 2013;58:152-62.
- 656 9. Citti C, Braghiroli D, Vandelli MA, Cannazza G. Pharmaceutical and biomedical analysis of
657 cannabinoids: A critical review. *J Pharm Biomed Anal*. 2018;147:565-79.
- 658 10. Keul A, Eisenhauer B. Making the high country: cannabis tourism in Colorado USA. *Annals*
659 *of Leisure Research*. 2018;22(2):140-60.
- 660 11. Peng H, Shahidi F. Cannabis and Cannabis Edibles: A Review. *J Agric Food Chem*.
661 2021;69(6):1751-74.
- 662 12. Gulck T, Moller BL. Phytocannabinoids: Origins and Biosynthesis. *Trends Plant Sci*.
663 2020;25(10):985-1004.
- 664 13. Lange BM, Zager JJ. Comprehensive inventory of cannabinoids in *Cannabis sativa* L.: Can
665 we connect genotype and chemotype? *Phytochemistry Reviews*. 2021:1-41.
- 666 14. Flores-Sanchez IJ, Verpoorte R. Secondary metabolism in cannabis. *Phytochemistry reviews*.
667 2008;7(3):615-39.
- 668 15. Iseger TA, Bossong MG. A systematic review of the antipsychotic properties of cannabidiol
669 in humans. *Schizophr Res*. 2015;162(1-3):153-61.
- 670 16. Mechoulam R, Parker LA. The endocannabinoid system and the brain. *Annu Rev Psychol*.
671 2013;64:21-47.
- 672 17. Toth JA, Stack GM, Cala AR, Carlson CH, Wilk RL, Crawford JL, et al. Development and
673 validation of genetic markers for sex and cannabinoid chemotype in *Cannabis sativa* L. *Global*
674 *Change Biology Bioenergy*. 2020;12(3):213-22.
- 675 18. Johnson MS, Wallace JG. Genomic and Chemical Diversity of Commercially Available
676 High-CBD Industrial Hemp Accessions. *Front Genet*. 2021;12(1160):682475.
- 677 19. Blatt-Janmaat K, Qu Y. The Biochemistry of Phytocannabinoids and Metabolic Engineering
678 of Their Production in Heterologous Systems. *Int J Mol Sci*. 2021;22(5):2454.
- 679 20. Luo X, Reiter MA, d'Espaux L, Wong J, Denby CM, Lechner A, et al. Complete biosynthesis
680 of cannabinoids and their unnatural analogues in yeast. *Nature*. 2019;567(7746):123-6.
- 681 21. Tahir MN, Shahbazi F, Rondeau-Gagne S, Trant JF. The biosynthesis of the cannabinoids. *J*
682 *Cannabis Res*. 2021;3(1):7.
- 683 22. Hanus LO, Meyer SM, Munoz E, Taglialatela-Scafati O, Appendino G. Phytocannabinoids: a
684 unified critical inventory. *Nat Prod Rep*. 2016;33(12):1357-92.
- 685 23. Tan Z, Clomburg JM, Gonzalez R. Synthetic Pathway for the Production of Olivetolic Acid
686 in *Escherichia coli*. *ACS Synth Biol*. 2018;7(8):1886-96.
- 687 24. Livingston SJ, Quilichini TD, Booth JK, Wong DCJ, Rensing KH, Laflamme-Yonkman J, et
688 al. Cannabis glandular trichomes alter morphology and metabolite content during flower maturation.
689 *Plant J*. 2020;101(1):37-56.
- 690 25. Stout JM, Boubakir Z, Ambrose SJ, Purves RW, Page JE. The hexanoyl-CoA precursor for
691 cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. *The*
692 *Plant Journal*. 2012;71(3):353-65.

- 693 26. Borrego EJ, Kolomiets MV. Synthesis and Functions of Jasmonates in Maize. *Plants* (Basel).
694 2016;5(4):41.
- 695 27. Feussner I, Wasternack C. The lipoxygenase pathway. *Annu Rev Plant Biol.* 2002;53(1):275-
696 97.
- 697 28. Yan Y, Christensen S, Isakeit T, Engelberth J, Meeley R, Hayward A, et al. Disruption of
698 OPR7 and OPR8 reveals the versatile functions of jasmonic acid in maize development and defense.
699 *Plant Cell.* 2012;24(4):1420-36.
- 700 29. Yan Y, Borrego E, Kolomiets MV. Jasmonate biosynthesis, perception and function in plant
701 development and stress responses. *Lipid metabolism InTech.* 2013:393-442.
- 702 30. Wasternack C. Jasmonates: an update on biosynthesis, signal transduction and action in plant
703 stress response, growth and development. *Ann Bot.* 2007;100(4):681-97.
- 704 31. Laudert D, Pfannschmidt U, Lottspeich F, Hollander-Czytko H, Weiler EW. Cloning,
705 molecular and functional characterization of *Arabidopsis thaliana* allene oxide synthase (CYP 74),
706 the first enzyme of the octadecanoid pathway to jasmonates. *Plant Mol Biol.* 1996;31(2):323-35.
- 707 32. Park JH, Halitschke R, Kim HB, Baldwin IT, Feldmann KA, Feyereisen R. A knock-out
708 mutation in allene oxide synthase results in male sterility and defective wound signal transduction in
709 *Arabidopsis* due to a block in jasmonic acid biosynthesis. *The Plant Journal.* 2002;31(1):1-12.
- 710 33. Stenzel I, Hause B, Miersch O, Kurz T, Maucher H, Weichert H, et al. Jasmonate
711 biosynthesis and the allene oxide cyclase family of *Arabidopsis thaliana*. *Plant Mol Biol.*
712 2003;51(6):895-911.
- 713 34. Pollmann S, Springer A, Rustgi S, von Wettstein D, Kang C, Reinbothe C, et al. Substrate
714 channeling in oxylipin biosynthesis through a protein complex in the plastid envelope of *Arabidopsis*
715 *thaliana*. *J Exp Bot.* 2019;70(5):1483-95.
- 716 35. Park SW, Li W, Viehhauser A, He B, Kim S, Nilsson AK, et al. Cyclophilin 20-3 relays a 12-
717 oxo-phytodienoic acid signal during stress responsive regulation of cellular redox homeostasis. *Proc*
718 *Natl Acad Sci U S A.* 2013;110(23):9559-64.
- 719 36. Stintzi A, Weber H, Reymond P, Browse J, Farmer EE. Plant defense in the absence of
720 jasmonic acid: the role of cyclopentenones. *Proc Natl Acad Sci U S A.* 2001;98(22):12837-42.
- 721 37. Wang KD, Borrego EJ, Kenerley CM, Kolomiets MV. Oxylipins Other Than Jasmonic Acid
722 Are Xylem-Resident Signals Regulating Systemic Resistance Induced by *Trichoderma virens* in
723 Maize. *Plant Cell.* 2020;32(1):166-85.
- 724 38. Mussig C, Biesgen C, Lisso J, Uwer U, Weiler EW, Altmann T. A novel stress-inducible 12-
725 oxophytodienoate reductase from *Arabidopsis thaliana* provides a potential link between
726 Brassinosteroid-action and Jasmonic-acid synthesis. *Journal of Plant Physiology.* 2000;157(2):143-
727 52.
- 728 39. Stintzi A, Browse J. The *Arabidopsis* male-sterile mutant, *opr3*, lacks the 12-oxophytodienoic
729 acid reductase required for jasmonate synthesis. *Proc Natl Acad Sci U S A.* 2000;97(19):10625-30.
- 730 40. Vick BA, Zimmerman DC. Biosynthesis of jasmonic Acid by several plant species. *Plant*
731 *Physiol.* 1984;75(2):458-61.
- 732 41. Wasternack C, Hause B. A Bypass in Jasmonate Biosynthesis - the OPR3-independent
733 Formation. *Trends Plant Sci.* 2018;23(4):276-9.

- 734 42. Chini A, Monte I, Zamarreno AM, Hamberg M, Lassueur S, Reymond P, et al. An OPR3-
735 independent pathway uses 4,5-didehydrojasmonate for jasmonate synthesis. *Nat Chem Biol.*
736 2018;14(2):171-8.
- 737 43. Weber H, Vick BA, Farmer EE. Dinor-oxo-phytodienoic acid: a new hexadecanoid signal in
738 the jasmonate family. *Proc Natl Acad Sci U S A.* 1997;94(19):10473-8.
- 739 44. Fonseca S, Chini A, Hamberg M, Adie B, Porzel A, Kramell R, et al. (+)-7-iso-Jasmonoyl-L-
740 isoleucine is the endogenous bioactive jasmonate. *Nature chemical biology.* 2009;5(5):344-50.
- 741 45. Christensen SA, Nemchenko A, Borrego E, Murray I, Sobhy IS, Bosak L, et al. The maize
742 lipoxygenase, Zm LOX 10, mediates green leaf volatile, jasmonate and herbivore-induced plant
743 volatile production for defense against insect attack. *The Plant Journal.* 2013;74(1):59-73.
- 744 46. Ameye M, Allmann S, Verwaeren J, Smagghe G, Haesaert G, Schuurink RC, et al. Green leaf
745 volatile production by plants: a meta-analysis. *New Phytol.* 2018;220(3):666-83.
- 746 47. Sugimoto K, Iijima Y, Takabayashi J, Matsui K. Processing of Airborne Green Leaf Volatiles
747 for Their Glycosylation in the Exposed Plants. *Front Plant Sci.* 2021;12(2514):721572.
- 748 48. Bate NJ, Sivasankar S, Moxon C, Riley JM, Thompson JE, Rothstein SJ. Molecular
749 characterization of an Arabidopsis gene encoding hydroperoxide lyase, a cytochrome P-450 that is
750 wound inducible. *Plant Physiol.* 1998;117(4):1393-400.
- 751 49. Matsui K, Wilkinson J, Hiatt B, Knauf V, Kajiwarra T. Molecular cloning and expression of
752 Arabidopsis fatty acid hydroperoxide lyase. *Plant Cell Physiol.* 1999;40(5):477-81.
- 753 50. Duan H, Huang MY, Palacio K, Schuler MA. Variations in CYP74B2 (hydroperoxide lyase)
754 gene expression differentially affect hexenal signaling in the Columbia and Landsberg erecta
755 ecotypes of Arabidopsis. *Plant Physiol.* 2005;139(3):1529-44.
- 756 51. Kallenbach M, Gilardoni PA, Allmann S, Baldwin IT, Bonaventure G. C12 derivatives of the
757 hydroperoxide lyase pathway are produced by product recycling through lipoxygenase-2 in *Nicotiana*
758 *attenuata* leaves. *New Phytologist.* 2011;191(4):1054-68.
- 759 52. Nakashima A, von Reuss SH, Tasaka H, Nomura M, Mochizuki S, Iijima Y, et al. Traumatins
760 and dinortraumatins-containing galactolipids in Arabidopsis: their formation in tissue-disrupted leaves
761 as counterparts of green leaf volatiles. *J Biol Chem.* 2013;288(36):26078-88.
- 762 53. Tanaka T, Ikeda A, Shiojiri K, Ozawa R, Shiki K, Nagai-Kunihiro N, et al. Identification of a
763 Hexenal Reductase That Modulates the Composition of Green Leaf Volatiles. *Plant Physiol.*
764 2018;178(2):552-64.
- 765 54. D'Auria JC, Pichersky E, Schaub A, Hansel A, Gershenzon J. Characterization of a BAHD
766 acyltransferase responsible for producing the green leaf volatile (Z)-3-hexen-1-yl acetate in
767 Arabidopsis thaliana. *Plant J.* 2007;49(2):194-207.
- 768 55. Kunishima M, Yamauchi Y, Mizutani M, Kuse M, Takikawa H, Sugimoto Y. Identification
769 of (Z)-3:(E)-2-Hexenal Isomerases Essential to the Production of the Leaf Aldehyde in Plants. *J Biol*
770 *Chem.* 2016;291(27):14023-33.
- 771 56. Spyropoulou EA, Dekker HL, Steemers L, van Maarseveen JH, de Koster CG, Haring MA, et
772 al. Identification and Characterization of (3Z):(2E)-Hexenal Isomerases from Cucumber. *Front Plant*
773 *Sci.* 2017;8:1342.

774 57. Grassa CJ, Weiblen GD, Wenger JP, Dabney C, Poplawski SG, Timothy Motley S, et al. A
775 new Cannabis genome assembly associates elevated cannabidiol (CBD) with hemp introgressed into
776 marijuana. *New Phytol.* 2021;230(4):1665-79.

777 58. Rangwala SH, Kuznetsov A, Ananiev V, Asztalos A, Borodin E, Evgeniev V, et al.
778 Accessing NCBI data using the NCBI Sequence Viewer and Genome Data Viewer (GDV). *Genome*
779 *Res.* 2021;31(1):159-69.

780 59. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the
781 conserved domain database in 2020. *Nucleic Acids Res.* 2020;48(D1):D265-D8.

782 60. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al.
783 CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.*
784 2011;39(Database issue):D225-9.

785 61. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs
786 in bipolymers. 1994.

787 62. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res.*
788 2015;43(W1):W39-49.

789 63. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc:
790 prediction of protein subcellular localization using deep learning. *Bioinformatics.* 2017;33(21):3387-
791 95.

792 64. Sperschneider J, Catanzariti AM, DeBoer K, Petre B, Gardiner DM, Singh KB, et al.
793 LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell.
794 *Sci Rep.* 2017;7(1):44598.

795 65. Sahu SS, Loaiza CD, Kaundal R. Plant-mSubP: a computational framework for the prediction
796 of single-and multi-target protein subcellular localization using integrated machine-learning
797 approaches. *AoB Plants.* 2020;12(3):plz068.

798 66. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A,
799 et al. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance.*
800 2019;2(5).

801 67. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment,
802 interactive sequence choice and visualization. *Brief Bioinform.* 2019;20(4):1160-6.

803 68. Kuraku S, Zmasek CM, Nishimura O, Katoh K. aLeaves facilitates on-demand exploration of
804 metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity.
805 *Nucleic Acids Res.* 2013;41(Web Server issue):W22-8.

806 69. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative
807 Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant.* 2020;13(8):1194-202.

808 70. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a
809 database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter
810 sequences. *Nucleic Acids Res.* 2002;30(1):325-7.

811 71. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection
812 and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40(7):e49.

813 72. Braich S, Baillie RC, Jewell LS, Spangenberg GC, Cogan NOI. Generation of a
814 Comprehensive Transcriptome Atlas and Transcriptome Dynamics in Medicinal Cannabis. *Sci Rep.*
815 2019;9(1):16583.

816 73. Zager JJ, Lange I, Srividya N, Smith A, Lange BM. Gene Networks Underlying Cannabinoid
817 and Terpenoid Accumulation in Cannabis. *Plant Physiol.* 2019;180(4):1877-97.

818 74. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
819 *Bioinformatics.* 2018;34(17):i884-i90.

820 75. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware
821 quantification of transcript expression. *Nat Methods.* 2017;14(4):417-9.

822 76. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis.
823 *BMC Bioinformatics.* 2008;9(1):559.

824 77. Greenfest-Allen E, Cartailier J-P, Magnuson MA, Stoeckert CJ. iterativeWGCNA: iterative
825 refinement to improve module detection from WGCNA co-expression networks. *bioRxiv.*
826 2017:234062.

827 78. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software
828 environment for integrated models of biomolecular interaction networks. *Genome Res.*
829 2003;13(11):2498-504.

830 79. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server
831 for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*
832 2019;47(W1):W191-W8.

833 80. Van Bel M, Silvestri F, Weitz EM, Kreft L, Botzki A, Coppens F, et al. PLAZA 5.0:
834 extending the scope and power of comparative and functional genomics in plants. *Nucleic Acids Res.*
835 2022;50(D1):D1468-D74.

836 81. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein*
837 *Sci.* 2019;28(11):1947-51.

838 82. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating
839 viruses and cellular organisms. *Nucleic Acids Res.* 2021;49(D1):D545-D51.

840 83. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*
841 2000;28(1):27-30.

842 84. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool
843 for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25-9.

844 85. Gene Ontology C. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.*
845 2021;49(D1):D325-D34.

846 86. Majorek KA, Dunin-Horkawicz S, Steczkiewicz K, Muszewska A, Nowotny M, Ginalski K,
847 et al. The RNase H-like superfamily: new members, comparative structural analysis and evolutionary
848 classification. *Nucleic Acids Res.* 2014;42(7):4160-79.

849 87. Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse
850 transcriptase sequences. *EMBO J.* 1990;9(10):3353-62.

851 88. Farmaki T, Sanmartin M, Jimenez P, Paneque M, Sanz C, Vancanneyt G, et al. Differential
852 distribution of the lipoxygenase pathway enzymes within potato chloroplasts. *J Exp Bot.*
853 2007;58(3):555-68.

854 89. Mikulska-Ruminska K, Shrivastava I, Krieger J, Zhang S, Li H, Bayir H, et al.
855 Characterization of Differential Dynamics, Specificity, and Allosteric of Lipoxygenase Family
856 Members. *J Chem Inf Model.* 2019;59(5):2496-508.

- 857 90. Wasternack C, Feussner I. The Oxylipin Pathways: Biochemistry and Function. *Annu Rev*
858 *Plant Biol.* 2018;69:363-86.
- 859 91. Toporkova YY, Smirnova EO, Iljina TM, Mukhtarova LS, Gorina SS, Grechkin AN. The
860 CYP74B and CYP74D divinyl ether synthases possess a side hydroperoxide lyase and epoxyalcohol
861 synthase activities that are enhanced by the site-directed mutagenesis. *Phytochemistry.*
862 2020;179:112512.
- 863 92. Li JW, Guang YL, Yang YX, Zhou Y. Identification and Expression Analysis of Two allene
864 oxide cyclase (AOC) Genes in Watermelon. *Agriculture-Basel.* 2019;9(10):225.
- 865 93. Beynon ER, Symons ZC, Jackson RG, Lorenz A, Rylott EL, Bruce NC. The role of
866 oxophytodienoate reductases in the detoxification of the explosive 2,4,6-trinitrotoluene by
867 *Arabidopsis*. *Plant Physiol.* 2009;151(1):253-61.
- 868 94. Liu X, Hou X. Antagonistic Regulation of ABA and GA in Metabolism and Signaling
869 Pathways. *Front Plant Sci.* 2018;9:251.
- 870 95. Park S, Liu W. 12-oxo-phytodienoic acid: a fuse and/or switch of plant growth and defense
871 responses? *Frontiers in Plant Science.* 2021:1687.
- 872 96. van Bakel H, Stout JM, Cote AG, Tallon CM, Sharpe AG, Hughes TR, et al. The draft
873 genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 2011;12(10):R102.
- 874 97. Aldridge DC, Galt S, Giles D, Turner WB. Metabolites of *Lasiodiplodia-Theobromae*. *J*
875 *Chem Soc C.* 1971(9):1623-&.
- 876 98. Huang H, Liu B, Liu L, Song S. Jasmonate action in plant growth and development. *J Exp*
877 *Bot.* 2017;68(6):1349-59.
- 878 99. Howe GA, Major IT, Koo AJ. Modularity in Jasmonate Signaling for Multistress Resilience.
879 *Annu Rev Plant Biol.* 2018;69:387-415.
- 880 100. Heitz T, Smirnova E, Marquis V, Poirier L. Metabolic Control within the Jasmonate
881 Biochemical Pathway. *Plant Cell Physiol.* 2019;60(12):2621-8.
- 882 101. Wasternack C, Hause B. Jasmonates: biosynthesis, perception, signal transduction and action
883 in plant stress response, growth and development. An update to the 2007 review in *Annals of Botany.*
884 *Ann Bot.* 2013;111(6):1021-58.
- 885 102. Upadhyay RK, Edelman M, Mattoo AK. Identification, Phylogeny, and Comparative
886 Expression of the Lipoxygenase Gene Family of the Aquatic Duckweed, *Spirodela polyrrhiza*, during
887 Growth and in Response to Methyl Jasmonate and Salt. *Int J Mol Sci.* 2020;21(24):9527.
- 888 103. Zhang B, Chen K, Bowen J, Allan A, Espley R, Karunairetnam S, et al. Differential
889 expression within the LOX gene family in ripening kiwifruit. *J Exp Bot.* 2006;57(14):3825-36.
- 890 104. Bannenberg G, Martinez M, Hamberg M, Castresana C. Diversity of the enzymatic activity in
891 the lipoxygenase gene family of *Arabidopsis thaliana*. *Lipids.* 2009;44(2):85-95.
- 892 105. Umate P. Genome-wide analysis of lipoxygenase gene family in *Arabidopsis* and rice. *Plant*
893 *Signal Behav.* 2011;6(3):335-8.
- 894 106. Chen Z, Chen D, Chu W, Zhu D, Yan H, Xiang Y. Retention and Molecular Evolution of
895 Lipoxygenase Genes in Modern Rosid Plants. *Front Genet.* 2016;7:176.

- 896 107. Wang J, Hu T, Wang W, Hu H, Wei Q, Wei X, et al. Bioinformatics Analysis of the
897 Lipoyxygenase Gene Family in Radish (*Raphanus sativus*) and Functional Characterization in
898 Response to Abiotic and Biotic Stresses. *Int J Mol Sci.* 2019;20(23):6095.
- 899 108. Yan C, Jia K, Zhang J, Xiao Z, Sha X, Gao J, et al. Genome-wide identification and
900 expression pattern analysis of lipoyxygenase gene family in turnip (*Brassica rapa* L. subsp. *rapa*).
901 *PeerJ.* 2022;10:e13746.
- 902 109. Yang XY, Jiang WJ, Yu HJ. The expression profiling of the lipoyxygenase (LOX) family
903 genes during fruit development, abiotic stress and hormonal treatments in cucumber (*Cucumis*
904 *sativus* L.). *Int J Mol Sci.* 2012;13(2):2481-500.
- 905 110. Liu S, Liu X, Jiang L. Genome-wide identification, phylogeny and expression analysis of the
906 lipoyxygenase gene family in cucumber. *Genet Mol Res.* 2011;10(4):2613-36.
- 907 111. Zhang C, Jin YZ, Liu JY, Tang YF, Cao SX, Qi HY. The phylogeny and expression profiles
908 of the lipoyxygenase (LOX) family genes in the melon (*Cucumis melo* L.) genome. *Scientia*
909 *Horticulturae.* 2014;170:94-102.
- 910 112. Liu JP, Zhou Y, Li JW, Wang F, Yang YX. Comprehensive Genomic Characterization and
911 Expression Analysis of the Lipoyxygenase Gene Family in Watermelon under Hormonal Treatments.
912 *Agriculture-Basel.* 2020;10(10):429.
- 913 113. Mou Y, Sun Q, Yuan C, Zhao X, Wang J, Yan C, et al. Identification of the LOX Gene
914 Family in Peanut and Functional Characterization of AhLOX29 in Drought Tolerance. *Front Plant*
915 *Sci.* 2022;13:832785.
- 916 114. Shin JH, Van K, Kim DH, Kim KD, Jang YE, Choi BS, et al. The lipoyxygenase gene family:
917 a genomic fossil of shared polyploidy between *Glycine max* and *Medicago truncatula*. *BMC Plant*
918 *Biol.* 2008;8(1):133.
- 919 115. Shaban M, Ahmed MM, Sun H, Ullah A, Zhu L. Genome-wide identification of lipoyxygenase
920 gene family in cotton and functional characterization in response to abiotic stresses. *BMC Genomics.*
921 2018;19(1):599.
- 922 116. Liu F, Li H, Wu J, Wang B, Tian N, Liu J, et al. Genome-wide identification and expression
923 pattern analysis of lipoyxygenase gene family in banana. *Sci Rep.* 2021;11(1):9948.
- 924 117. Ogunola OF, Hawkins LK, Mylroie E, Kolomiets MV, Borrego E, Tang JD, et al.
925 Characterization of the maize lipoyxygenase gene family in relation to aflatoxin accumulation
926 resistance. *PLoS One.* 2017;12(7):e0181265.
- 927 118. Zhang Q, Zhao Y, Zhang J, Li X, Ma F, Duan M, et al. The Responses of the Lipoyxygenase
928 Gene Family to Salt and Drought Stress in Foxtail Millet (*Setaria italica*). *Life (Basel).*
929 2021;11(11):1169.
- 930 119. Shrestha K, Pant S, Huang Y. Genome-wide identification and classification of Lipoyxygenase
931 gene family and their roles in sorghum-aphid interaction. *Plant Mol Biol.* 2021;105(4-5):527-41.
- 932 120. Ji QQ, Zhang TY, Zhang D, Lv SM, Tan AJ. Genome-wide identification and expression
933 analysis of lipoyxygenase genes in Tartary buckwheat. *Biotechnology & Biotechnological Equipment.*
934 2020;34(1):273-86.
- 935 121. Li JD, Chen LC, Chen P, Li QC, Yang QQ, Zhang Y, et al. Genome-wide identification and
936 expression of the lipoyxygenase gene family in jujube (*Ziziphus jujuba*) in response to phytoplasma
937 infection. *Journal of Plant Biochemistry and Biotechnology.* 2022;31(1):139-53.

- 938 122. Vogt J, Schiller D, Ulrich D, Schwab W, Dunemann F. Identification of lipoxygenase (LOX)
939 genes putatively involved in fruit flavour formation in apple (*Malus x domestica*). *Tree Genetics &*
940 *Genomes*. 2013;9(6):1493-511.
- 941 123. Guo SL, Song ZZ, Ma RJ, Yang Y, Yu ML. Genome-wide identification and expression
942 analysis of the lipoxygenase gene family during peach fruit ripening under different postharvest
943 treatments. *Acta Physiologiae Plantarum*. 2017;39(5):111.
- 944 124. Li M, Li L, Dunwell JM, Qiao X, Liu X, Zhang S. Characterization of the lipoxygenase
945 (LOX) gene family in the Chinese white pear (*Pyrus bretschneideri*) and comparison with other
946 members of the Rosaceae. *BMC Genomics*. 2014;15(1):444.
- 947 125. Chen Z, Chen X, Yan H, Li W, Li Y, Cai R, et al. The Lipoxygenase Gene Family in Poplar:
948 Identification, Classification, and Expression in Response to MeJA Treatment. *PLoS One*.
949 2015;10(4):e0125526.
- 950 126. Sarde SJ, Kumar A, Remme RN, Dicke M. Genome-wide identification, classification and
951 expression of lipoxygenase gene family in pepper. *Plant Mol Biol*. 2018;98(4-5):375-87.
- 952 127. Upadhyay RK, Handa AK, Mattoo AK. Transcript Abundance Patterns of 9- and 13-
953 Lipoxygenase Subfamily Gene Members in Response to Abiotic Stresses (Heat, Cold, Drought or
954 Salt) in Tomato (*Solanum lycopersicum* L.) Highlights Member-Specific Dynamics Relevant to Each
955 Stress. *Genes (Basel)*. 2019;10(9):683.
- 956 128. Upadhyay RK, Mattoo AK. Genome-wide identification of tomato (*Solanum lycopersicum*
957 L.) lipoxygenases coupled with expression profiles during plant development and in response to
958 methyl-jasmonate and wounding. *J Plant Physiol*. 2018;231:318-28.
- 959 129. Zhu J, Wang X, Guo L, Xu Q, Zhao S, Li F, et al. Characterization and Alternative Splicing
960 Profiles of the Lipoxygenase Gene Family in Tea Plant (*Camellia sinensis*). *Plant Cell Physiol*.
961 2018;59(9):1765-81.
- 962 130. Podolyan A, White J, Jordan B, Winefield C. Identification of the lipoxygenase gene family
963 from *Vitis vinifera* and biochemical characterisation of two 13-lipoxygenases expressed in grape
964 berries of Sauvignon Blanc. *Functional Plant Biology*. 2010;37(8):767-84.
- 965 131. Mochizuki S, Sugimoto K, Koeduka T, Matsui K. Arabidopsis lipoxygenase 2 is essential for
966 formation of green leaf volatiles and five-carbon volatiles. *FEBS letters*. 2016;590(7):1017-27.
- 967 132. Conneely LJ, Mauleon R, Mieog J, Barkla BJ, Kretschmar T. Characterization of the
968 *Cannabis sativa* glandular trichome proteome. *PLoS One*. 2021;16(4):e0242633.
- 969 133. Watts S, McElroy M, Migicovsky Z, Maassen H, van Velzen R, Myles S. Cannabis labelling
970 is associated with genetic variation in terpene synthase genes. *Nat Plants*. 2021;7(10):1330-4.
- 971 134. Rice S, Koziel JA. Characterizing the Smell of Marijuana by Odor Impact of Volatile
972 Compounds: An Application of Simultaneous Chemical and Sensory Analysis. *PLoS One*.
973 2015;10(12):e0144160.
- 974 135. Zhou Y, Guang YL, Li JW, Wang F, Ahammed GJ, Yang YX. The CYP74 Gene Family in
975 Watermelon: Genome-Wide Identification and Expression Profiling Under Hormonal Stress and
976 Root-Knot Nematode Infection. *Agronomy-Basel*. 2019;9(12):872.
- 977 136. De Domenico S, Taurino M, Gallo A, Poltronieri P, Pastor V, Flors V, et al. Oxylipin
978 dynamics in *Medicago truncatula* in response to salt and wounding stresses. *Physiol Plant*.
979 2019;165(2):198-208.

- 980 137. Xiong R, He T, Wang Y, Liu S, Gao Y, Yan H, et al. Genome and transcriptome analysis to
981 understand the role diversification of cytochrome P450 gene under excess nitrogen treatment. BMC
982 Plant Biol. 2021;21(1):447.
- 983 138. Froehlich JE, Itoh A, Howe GA. Tomato allene oxide synthase and fatty acid hydroperoxide
984 lyase, two cytochrome P450s involved in oxylipin metabolism, are targeted to different membranes
985 of chloroplast envelope. Plant Physiol. 2001;125(1):306-17.
- 986 139. Toporkova YY, Askarova EK, Gorina SS, Ogorodnikova AV, Mukhtarova LS, Grechkin AN.
987 Epoxyalcohol synthase activity of the CYP74B enzymes of higher plants. Biochim Biophys Acta
988 Mol Cell Biol Lipids. 2020;1865(9):158743.
- 989 140. Toporkova YY, Gogolev YV, Mukhtarova LS, Grechkin AN. Determinants governing the
990 CYP74 catalysis: conversion of allene oxide synthase into hydroperoxide lyase by site-directed
991 mutagenesis. FEBS Lett. 2008;582(23-24):3423-8.
- 992 141. Lee DS, Nioche P, Hamberg M, Raman CS. Structural insights into the evolutionary paths of
993 oxylipin biosynthetic enzymes. Nature. 2008;455(7211):363-8.
- 994 142. Wu Q, Wu J, Sun H, Zhang D, Yu D. Sequence and expression divergence of the AOC gene
995 family in soybean: insights into functional diversity for stress responses. Biotechnol Lett.
996 2011;33(7):1351-9.
- 997 143. Stenzel I, Otto M, Delker C, Kirmse N, Schmidt D, Miersch O, et al. ALLENE OXIDE
998 CYCLASE (AOC) gene family members of Arabidopsis thaliana: tissue- and organ-specific
999 promoter activities and in vivo heteromerization. J Exp Bot. 2012;63(17):6125-38.
- 1000 144. Otto M, Naumann C, Brandt W, Wasternack C, Hause B. Activity Regulation by
1001 Heteromerization of Arabidopsis Allene Oxide Cyclase Family Members. Plants (Basel).
1002 2016;5(1):3.
- 1003 145. Matsui H, Nakamura G, Ishiga Y, Toshima H, Inagaki Y, Toyoda K, et al. Structure and
1004 expression of 12-oxophytodienoate reductase (subgroup I) genes in pea, and characterization of the
1005 oxidoreductase activities of their recombinant products. Mol Genet Genomics. 2004;271(1):1-10.
- 1006 146. Guang Y, Luo S, Ahammed GJ, Xiao X, Li J, Zhou Y, et al. The OPR gene family in
1007 watermelon: Genome-wide identification and expression profiling under hormone treatments and
1008 root-knot nematode infection. Plant Biol (Stuttg). 2021;23 Suppl 1:80-8.
- 1009 147. Zhang J, Simmons C, Yalpani N, Crane V, Wilkinson H, Kolomiets M. Genomic analysis of
1010 the 12-oxo-phytodienoic acid reductase gene family of Zea mays. Plant Mol Biol. 2005;59(2):323-
1011 43.
- 1012 148. Liu S, Sun R, Zhang X, Feng Z, Wei F, Zhao L, et al. Genome-Wide Analysis of OPR Family
1013 Genes in Cotton Identified a Role for GhOPR9 in Verticillium dahliae Resistance. Genes (Basel).
1014 2020;11(10):1134.
- 1015 149. Li W, Zhou F, Liu B, Feng D, He Y, Qi K, et al. Comparative characterization, expression
1016 pattern and function analysis of the 12-oxo-phytodienoic acid reductase gene family in rice. Plant
1017 Cell Rep. 2011;30(6):981-95.
- 1018 150. Mou Y, Liu Y, Tian S, Guo Q, Wang C, Wen S. Genome-Wide Identification and
1019 Characterization of the OPR Gene Family in Wheat (Triticum aestivum L.). Int J Mol Sci.
1020 2019;20(8):1914.

151. Sobajima H, Takeda M, Sugimori M, Kobashi N, Kiribuchi K, Cho EM, et al. Cloning and characterization of a jasmonic acid-responsive gene encoding 12-oxophytodienoic acid reductase in suspension-cultured rice cells. *Planta*. 2003;216(4):692-8.
152. Farmer EE, Davoine C. Reactive electrophile species. *Curr Opin Plant Biol*. 2007;10(4):380-6.
153. Li W, Liu B, Yu L, Feng D, Wang H, Wang J. Phylogenetic analysis, structural evolution and functional divergence of the 12-oxo-phytyldienoate acid reductase gene family in plants. *BMC Evol Biol*. 2009;9(1):90.
154. Balcke GU, Bennewitz S, Bergau N, Athmer B, Henning A, Majovsky P, et al. Multi-Omics of Tomato Glandular Trichomes Reveals Distinct Features of Central Carbon Metabolism Supporting High Productivity of Specialized Metabolites. *Plant Cell*. 2017;29(5):960-83.
155. Crombie WM. Fatty Acids in Chloroplasts and Leaves. *Journal of Experimental Botany*. 1958;9(26):254-61.
156. Poincelot RP. Lipid and Fatty Acid composition of chloroplast envelope membranes from species with differing net photosynthesis. *Plant Physiol*. 1976;58(4):595-8.
157. Engelberth J, Engelberth M. Variability in the Capacity to Produce Damage-Induced Aldehyde Green Leaf Volatiles among Different Plant Species Provides Novel Insights into Biosynthetic Diversity. *Plants (Basel)*. 2020;9(2):213.
158. Mano J, Belles-Boix E, Babiychuk E, Inze D, Torii Y, Hiraoka E, et al. Protection against photooxidative injury of tobacco leaves by 2-alkenal reductase. Detoxification of lipid peroxide-derived reactive carbonyls. *Plant Physiol*. 2005;139(4):1773-83.
159. Maynard D, Kumar V, Sproß J, Dietz K-J. 12-oxophytodienoic acid reductase 3 (OPR3) functions as NADPH-dependent α , β -ketoalkene reductase in detoxification and monodehydroascorbate reductase in redox homeostasis. *Plant and Cell Physiology*. 2020;61(3):584-95.
160. Tola AJ, Jaballi A, Germain H, Missihoun TD. Recent Development on Plant Aldehyde Dehydrogenase Enzymes and Their Functions in Plant Development and Stress Signaling. *Genes (Basel)*. 2020;12(1):51.
161. Li S, Hu PC, Malmstadt N. Imaging molecular transport across lipid bilayers. *Biophys J*. 2011;101(3):700-8.
162. Piovesana S, Aita SE, Cannazza G, Capriotti AL, Cavaliere C, Cerrato A, et al. In-depth cannabis fatty acid profiling by ultra-high performance liquid chromatography coupled to high resolution mass spectrometry. *Talanta*. 2021;228:122249.
163. Pathi KM, Rink P, Budhagatapalli N, Betz R, Saado I, Hiekel S, et al. Engineering Smut Resistance in Maize by Site-Directed Mutagenesis of LIPOXYGENASE 3. *Front Plant Sci*. 2020;11:543895.
164. Abbaraju HKR, Gupta R, Appenzeller LM, Fallis LP, Hazebroek J, Zhu G, et al. A vegetative storage protein improves drought tolerance in maize. *Plant Biotechnol J*. 2022;20(2):374-89.
165. Wang J, Kuang HQ, Zhang ZH, Yang YQ, Yan L, Zhang MC, et al. Generation of seed lipoxygenase-free soybean using CRISPR-Cas9. *Crop Journal*. 2020;8(3):432-9.

1063 **10 Supporting information**

1064 **S1 Fig.** Regions of duplicated LOX genes on *C. sativa* Chromosomes 2 and 9.

1065 **S2 Fig.** Regions of duplicated *OPR*, *AOS*, and *AOC* genes on *C. sativa* Chromosomes 1, 8, and X.

1066 **S3 Fig.** Expression of oxylipin biosynthetic genes across four developmental stages in *C. sativa*
1067 flowers and trichomes.

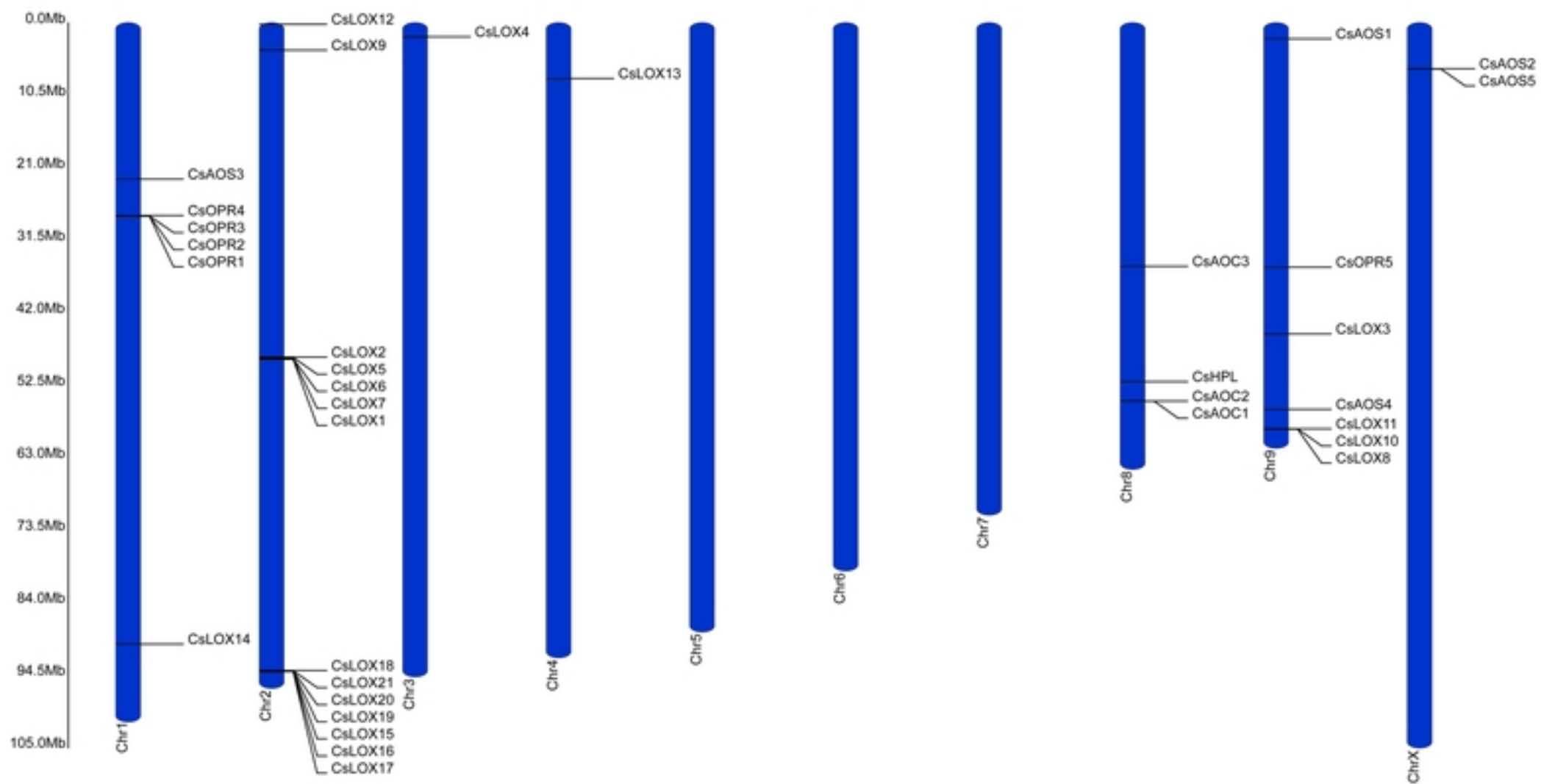


Fig1

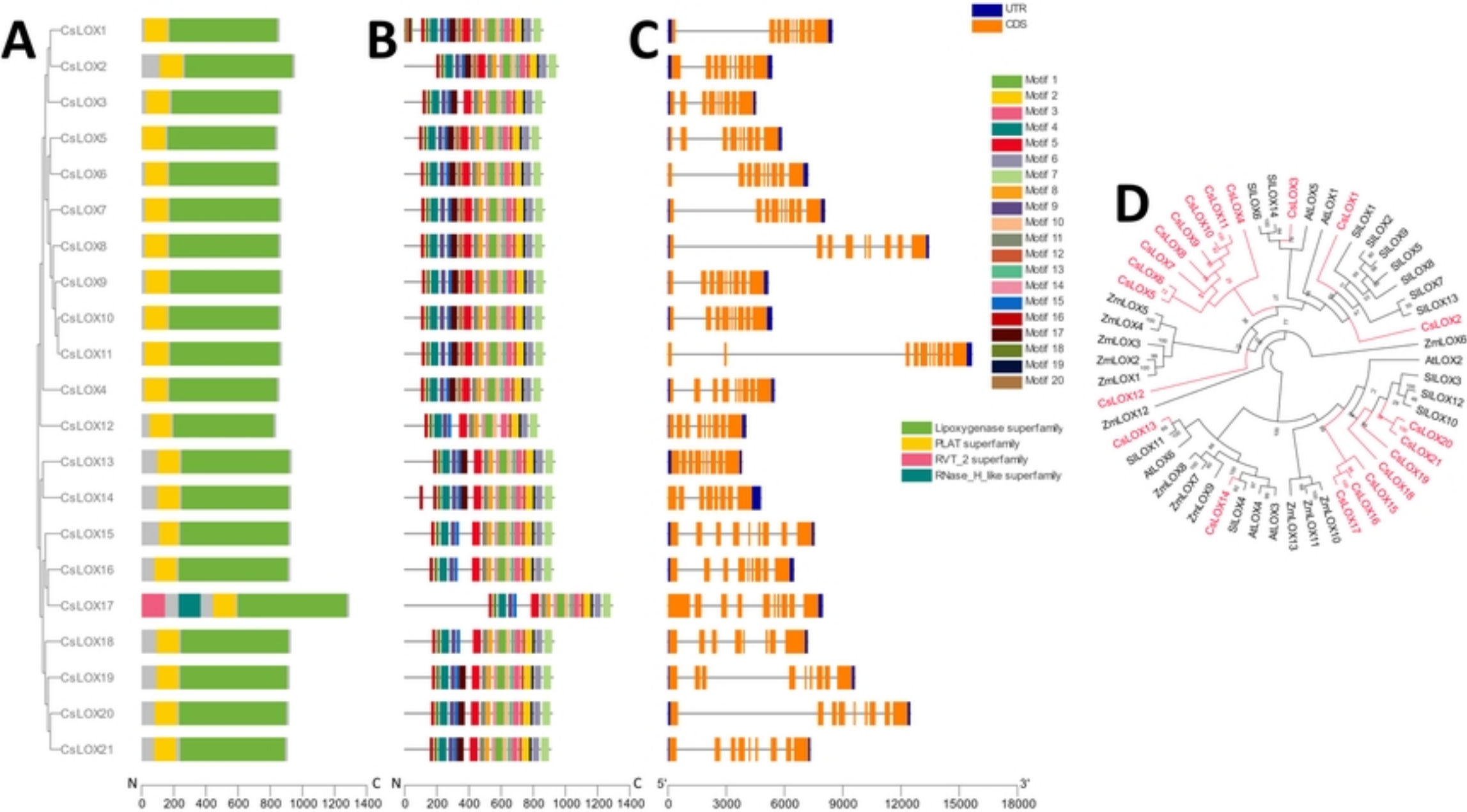


Fig2

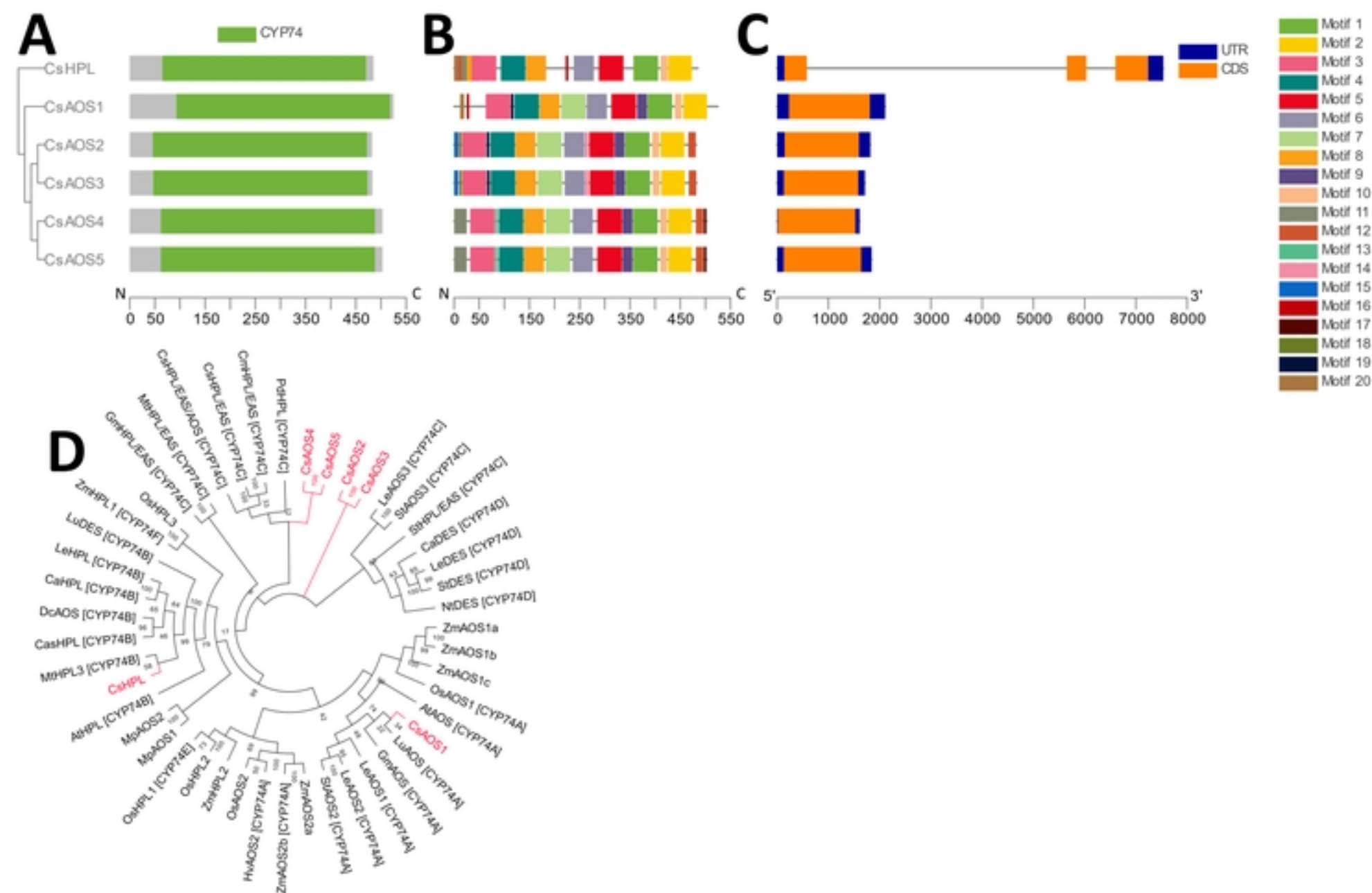


Fig3

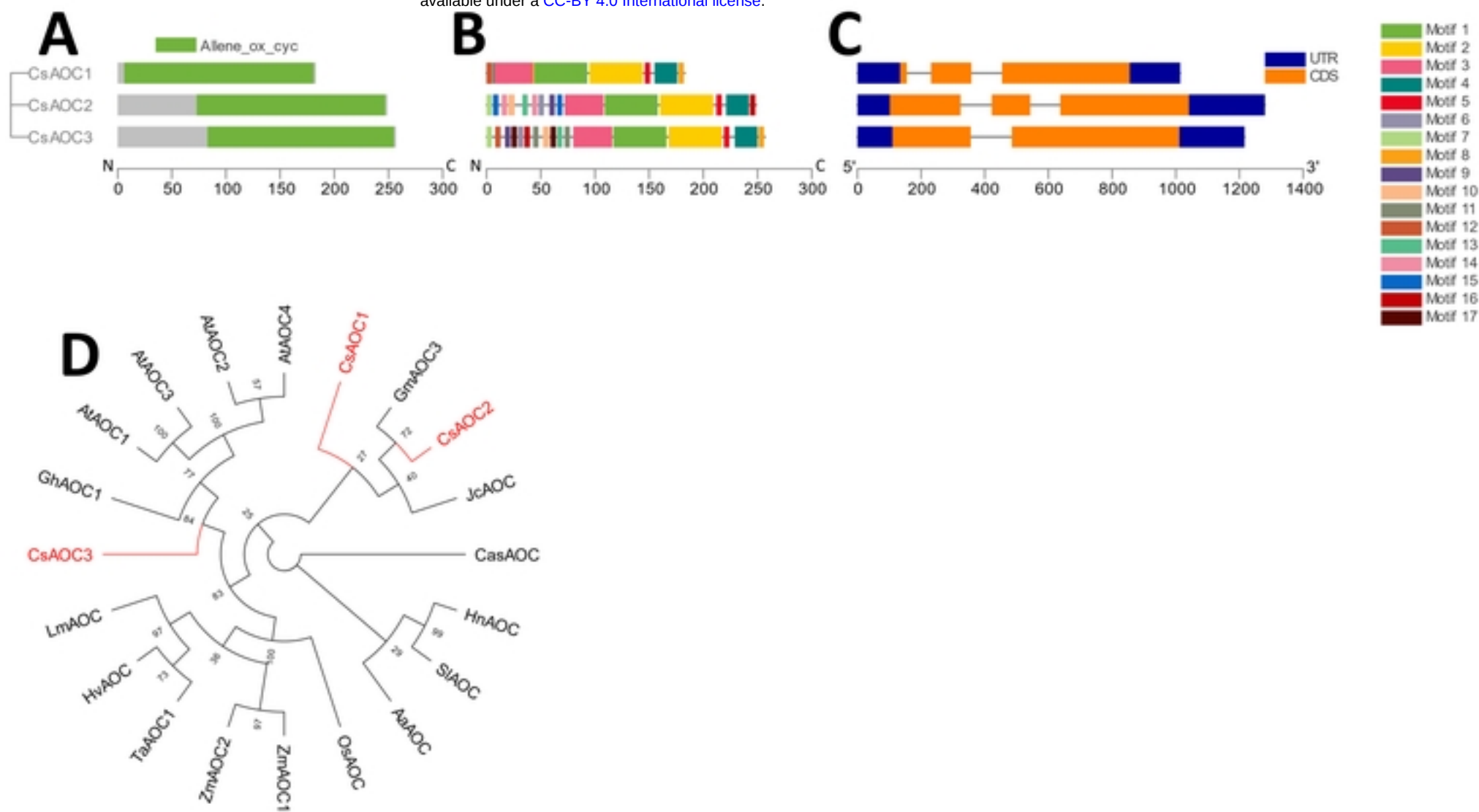


Fig4

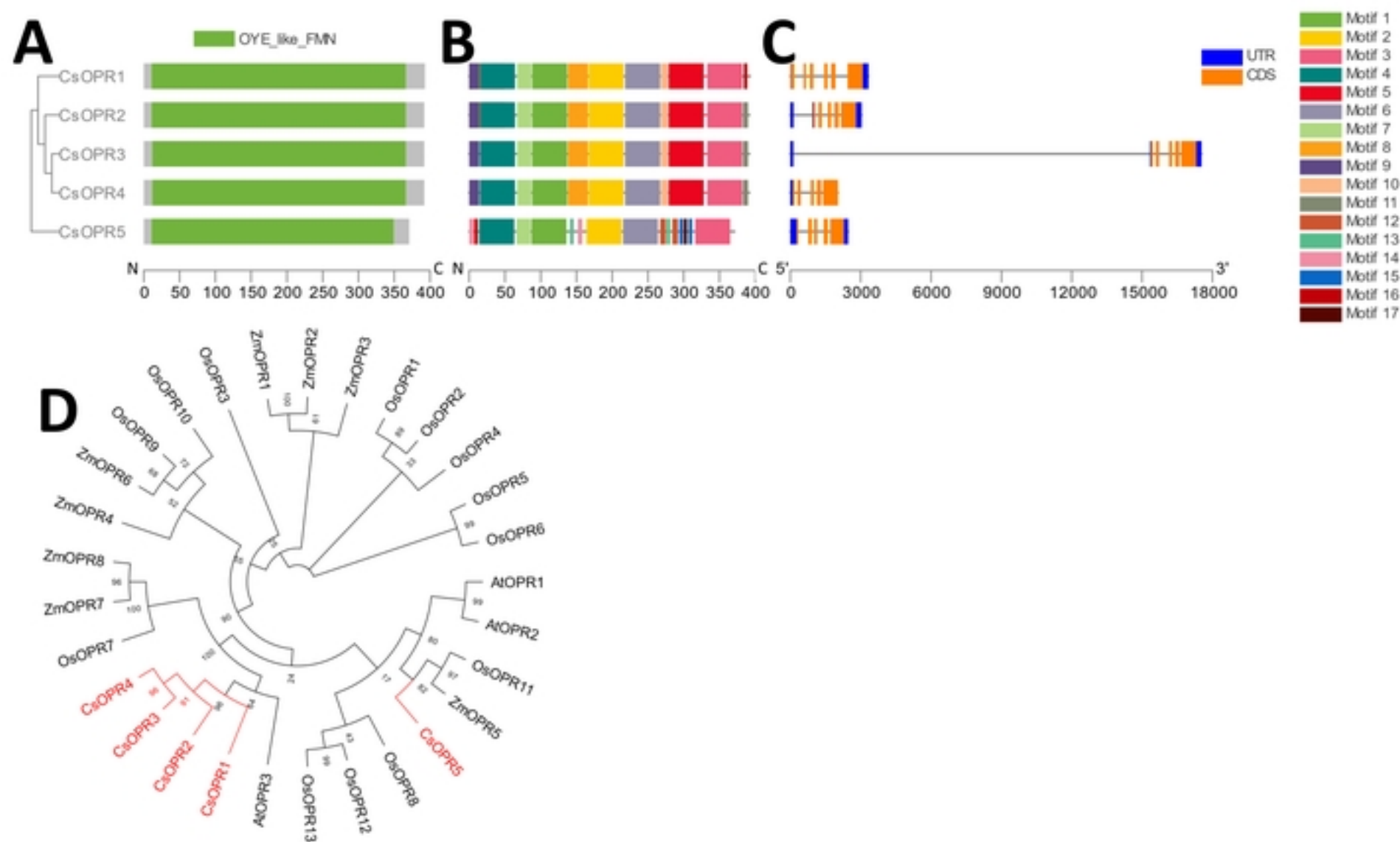


Fig5

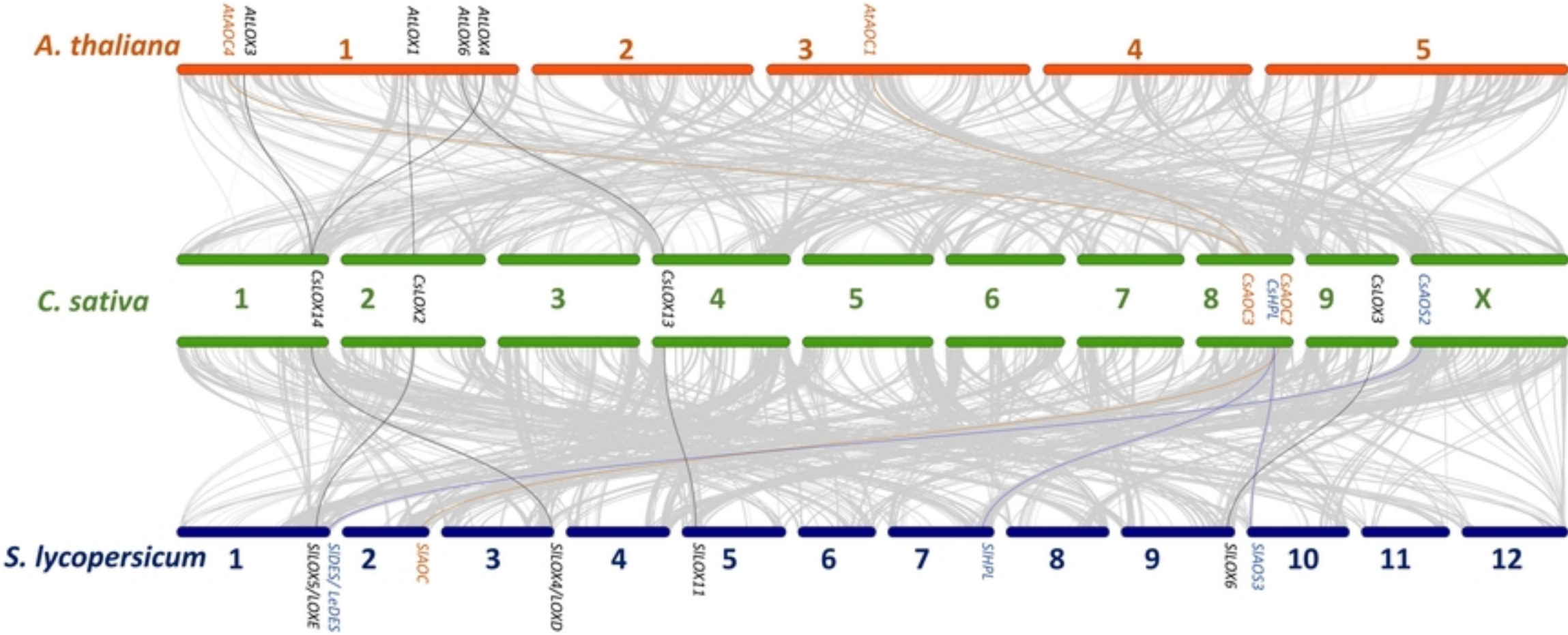


Fig6

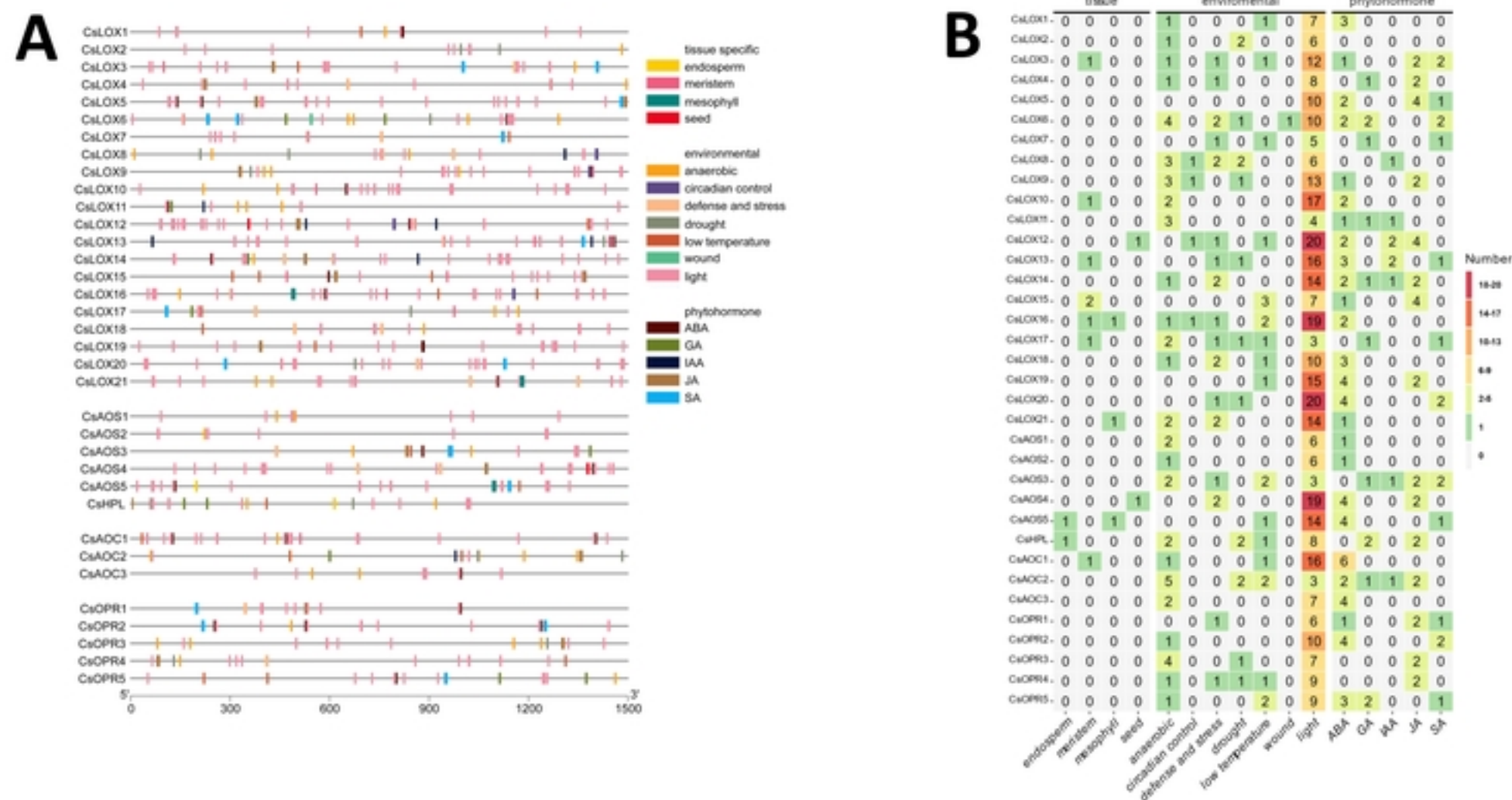


Fig7

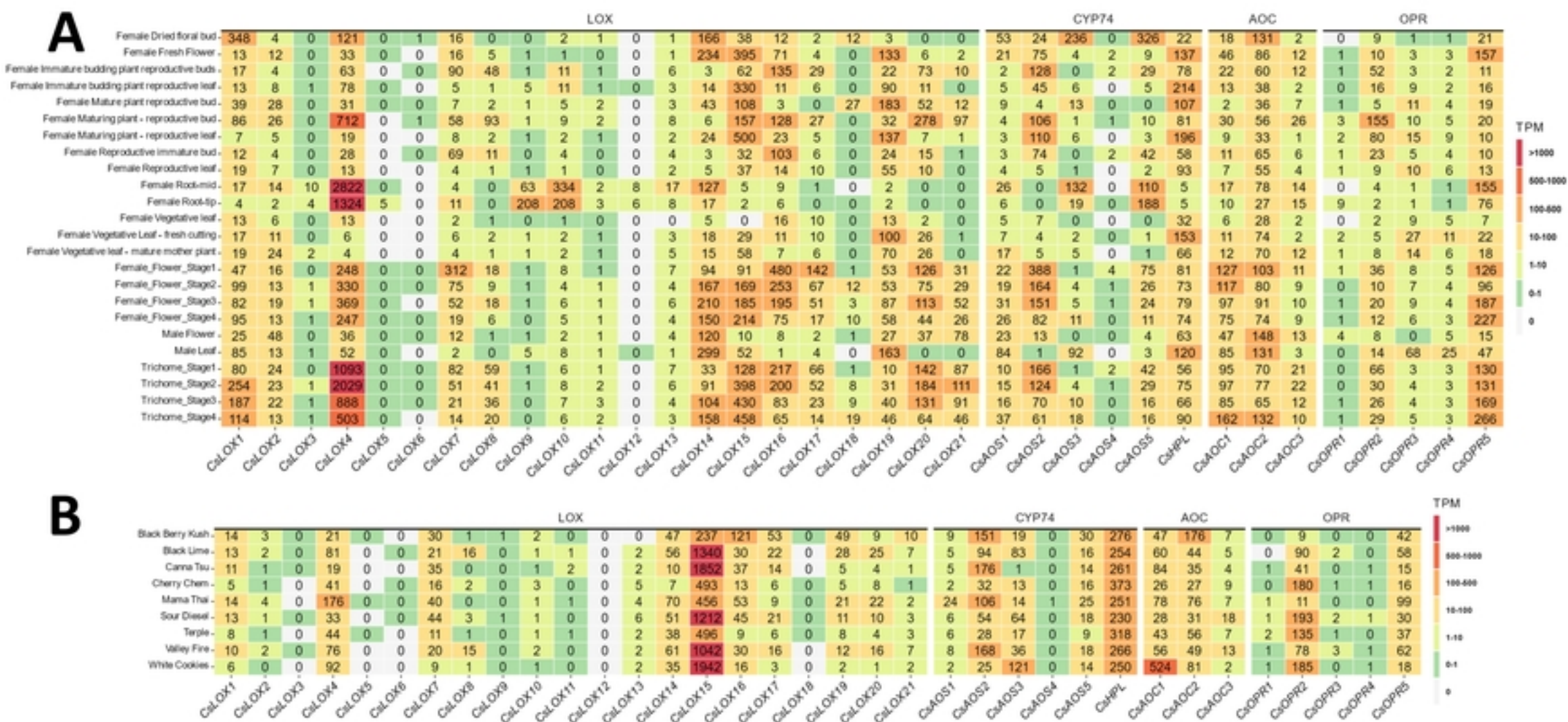
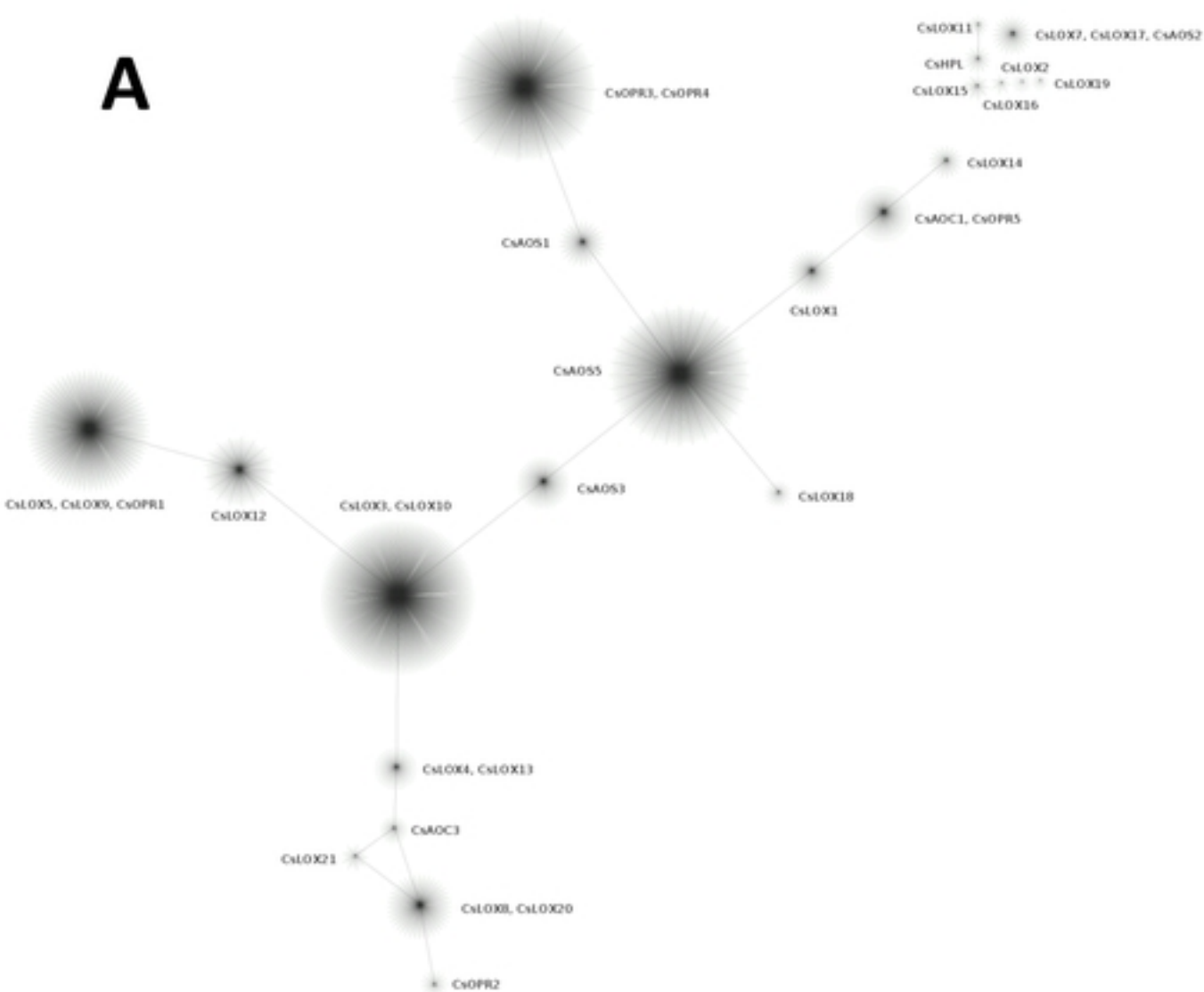


Fig8

bioRxiv preprint doi: <https://doi.org/10.1101/2022.07.30.502131>; this version posted August 1, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](#).



GO:MF		stats
Term name	Term ID	P(w)
protein binding	GO:0005515	6.941×10^{-10}
binding	GO:0005488	5.960×10^{-10}
DNA-binding transcription factor activity	GO:0003700	7.914×10^{-10}
transcription regulator activity	GO:0140110	6.003×10^{-11}
ion binding	GO:0043167	4.063×10^{-10}
sequence-specific DNA binding	GO:0043565	8.620×10^{-10}
transcription regulatory region nucleic acid binding	GO:0001067	3.085×10^{-9}
transcription cis-regulatory region binding	GO:0000976	1.065×10^{-9}
double-stranded DNA binding	GO:0003690	1.872×10^{-9}
catalytic activity	GO:0003824	3.175×10^{-9}

GO BP		stats
Term name	Term ID	Pval
cellular process	GO:0009987	3.223×10^{-105}
response to stimulus	GO:0050896	3.198×10^{-93}
anatomical structure development	GO:0048856	7.749×10^{-92}
developmental process	GO:0032502	1.121×10^{-91}
response to chemical	GO:0042221	2.421×10^{-85}
multicellular organismal process	GO:0032501	2.695×10^{-78}
multicellular organism development	GO:0007275	4.805×10^{-79}
system development	GO:0048731	4.980×10^{-65}
response to oxygen-containing compound	GO:1901700	1.418×10^{-64}
metabolic process	GO:0008152	1.950×10^{-63}

GO:CC		stats
Term name	Term ID	Pw()
intracellular anatomical structure	GO:0005622	6.955×10^{-90}
intracellular membrane-bounded organelle	GO:0043231	1.399×10^{-73}
membrane-bounded organelle	GO:0043227	3.430×10^{-72}
intracellular organelle	GO:0043229	1.510×10^{-66}
organelle	GO:0043226	1.019×10^{-64}
cytoplasm	GO:0005737	2.294×10^{-56}
cellular anatomical entity	GO:0110165	6.390×10^{-47}
cell periphery	GO:0071944	1.104×10^{-34}
plasma membrane	GO:0005886	4.158×10^{-34}
plant-type vacuole	GO:0000325	1.300×10^{-32}

KEGG	start
Term name	Term ID
Biosynthesis of secondary metabolites	KEGG:01110
Metabolic pathways	KEGG:01100
alpha-Unsaturated acid metabolism	KEGG:00590
Peroxisome	KEGG:04146
Valine, leucine and isoleucine degradation	KEGG:00280

Fig9

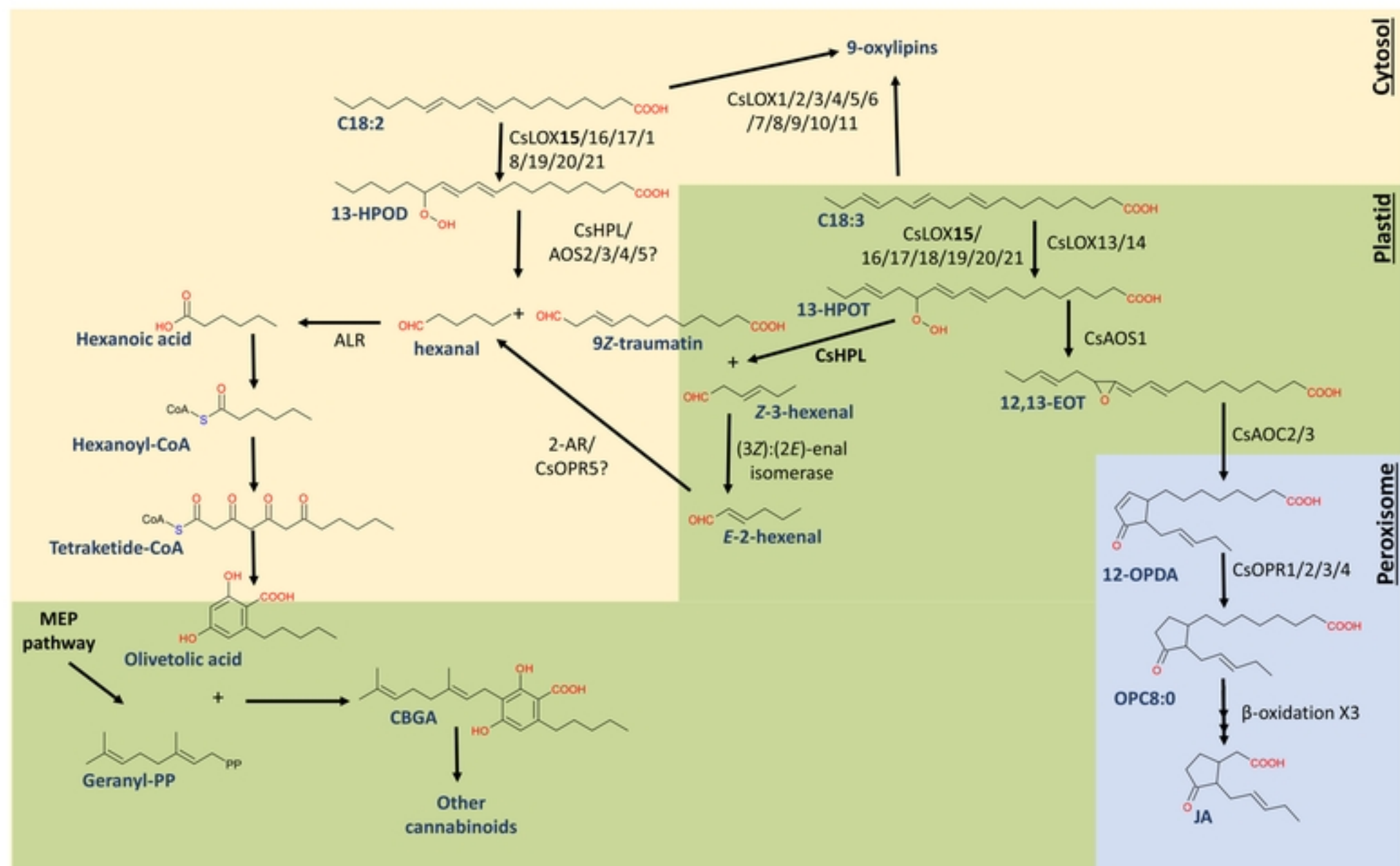


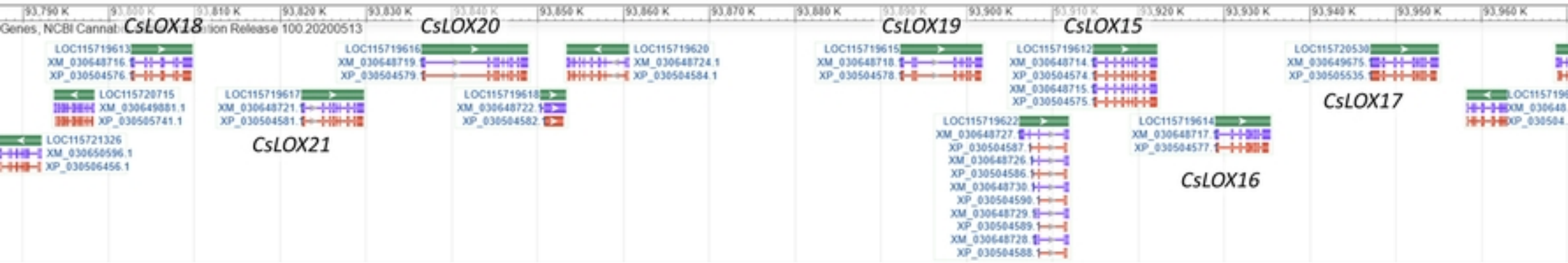
Fig10

bioRxiv preprint doi: <https://doi.org/10.1101/2022.07.30.502131>; this version posted August 1, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](#).

Chromosome 2



Chromosome 2

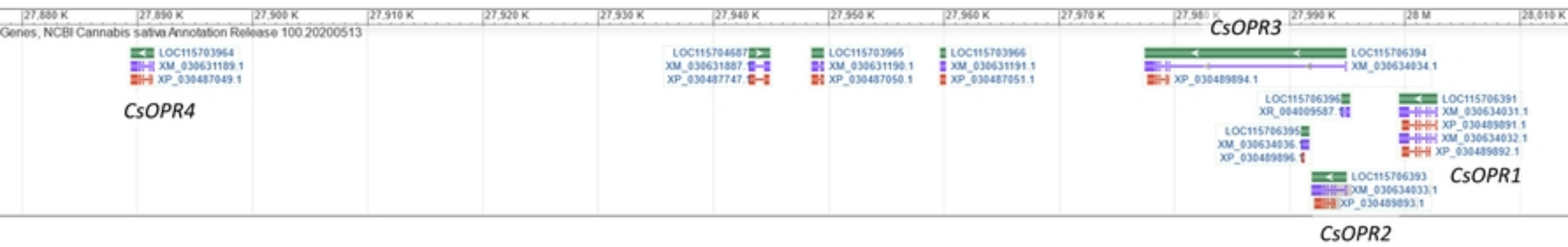


Chromosome 9



FigS1

Chromosome 1



Chromosome 8



Chromosome X



