

# 亚麻 FAD 基因家族的生物信息学鉴定分析

侯静静, 赵利, 王斌

(甘肃省农业科学院作物研究所, 甘肃 兰州 730070)

**摘要:** 不饱和脂肪酸为人体提供基本代谢所必需的能量, 须从膳食中补充。脂肪酸去饱和酶 FAD(fatty acid desaturase)是植物不饱和脂肪酸合成途径中的关键酶, 植物体内脂肪酸的各组分比例和不饱和度与 FAD 的去饱和和作用息息相关。为探究亚麻 FAD 基因家族的表达与进化, 为在亚麻高品质育种中的应用提供理论依据。运用生物信息学方法对亚麻全基因组 FAD 基因家族的 43 个 *LuFADs* 基因进行分析。结果显示, 该家族成员编码的蛋白质大小为 152~453 个氨基酸, 大部分为碱性不稳定亲水蛋白。与拟南芥 FADs 蛋白序列构建系统发育树, 可分为 4 个主要亚家族:  $\Delta 12/\omega-3$  去饱和酶、“前端”去饱和酶、 $\Delta 7/\Delta 9$  去饱和酶和 SAD 去饱和酶。保守结构域和外显子-内含子结构分析得出, 同一亚组中的家族成员具有较为相似的基因结构。染色体定位分析呈随机性分布。亚细胞定位预测得出, 叶绿体上的家族成员最多。启动子顺式作用元件分析发现, 该家族成员中抗氧化反应元件(ARE)数量最多。

**关键词:** 亚麻; FAD 基因家族; 生物信息学分析

**中图分类号:** S565.9

**文献标志码:** A

**文章编号:** 2097-2172(2023)03-0246-08

**doi:** 10.3969/j.issn.2097-2172.2023.03.011

## Identification and Bioinformatics Analysis of FAD Gene Family in *Linum usitatissimum* L.

HOU Jingjing, ZHAO Li, WANG Bin

(Institute of Crops, Gansu Academy of Agricultural Sciences, Lanzhou Gansu 730070, China)

**Abstract:** Unsaturated fatty acids provide the human body with the energy necessary for basic metabolism and must be supplemented from the diet. FAD (fatty acid desaturase) is a key enzyme in the biosynthesis of unsaturated fatty acids in plants. The proportion and unsaturation of fatty acids in plants are closely related to the desaturation of FAD. In order to explore the expression and evolution of flax (*Linum usitatissimum* L.) FAD gene family and to provide theoretical basis for its application in high-quality flax breeding, in this study, bioinformatics method was used to analyze 43 *LuFADs* genes of flax. Results showed that the size of proteins encoded by the members of this family were between 152 to 453 amino acids, most of which were alkaline unstable hydrophilic proteins. The phylogenetic tree was constructed with the FADs protein sequence of *Arabidopsis thaliana*, which could be divided into four main subfamilies:  $\Delta 12/\omega-3$  desaturase, 'front-end' desaturase,  $\Delta 7/\Delta 9$  desaturase and SAD desaturase. The analysis of conserved domain and exon-intron structure showed that the family members in the same subgroup had relatively similar gene structure. The chromosome localization analysis showed random distribution, and the subcellular localization predicted that the most family members were on the chloroplast. Analysis of promoter cis-acting elements showed that the number of antioxidant response elements (ARE) were the largest among this family.

**Key words:** *Linum usitatissimum* L.; FAD gene family; Bioinformatics analysis

亚麻(*Linum usitatissimum* L.)是一年生自交草本植物,在我国甘肃、内蒙古、山西、宁夏和河北等地广泛种植<sup>[1]</sup>。亚麻可分为油用、纤维用和油纤兼用 3 种类型,其中油用亚麻俗称胡麻,胡麻籽含油率为 40%左右,具有较高的食用营养价值<sup>[2]</sup>。胡麻籽油的脂肪酸组成包括 2 种脂肪酸(棕

桐酸及硬脂酸)和 3 种不饱和脂肪酸(油酸、亚油酸及  $\alpha$ -亚麻酸),这 5 种脂肪酸占到胡麻油粗脂肪含量的 99%,其中  $\alpha$ -亚麻酸达到 54%左右<sup>[3]</sup>。 $\alpha$ -亚麻酸是人体必需脂肪酸,属于  $\omega-3$  系不饱和脂肪酸,具有降血脂、抗血栓肿瘤、健脑明目和改善肌肤等多种保健功能<sup>[4]</sup>。

**收稿日期:** 2022-11-28

**基金项目:** 国家自然科学基金(32160497、32160499); 国家特色油料产业技术体系(CARS-14-1-03)。

**作者简介:** 侯静静(1994—),女,甘肃兰州人,研究实习员,硕士,主要从事胡麻遗传育种研究。Email: 528725294@qq.com。

**通信作者:** 赵利(1973—),女,陕西武功人,研究员,博士,主要从事胡麻种质资源创新及遗传育种研究。Email: zhyj0801@126.com。

不饱和脂肪酸的合成以饱和脂肪酸硬脂酸为底物, 在脂肪酸延长酶(fatty acid elongase, FAE)和脂肪酸去饱和酶 FAD(fatty acid desaturase)的作用下, 通过一系列的脱氢、延长、去饱和作用形成<sup>[5]</sup>。其中 FAD 是植物不饱和脂肪酸合成的关键酶, 植物体内脂肪酸各组分的比例及其不饱和度与 FAD 的去饱和作用息息相关<sup>[6]</sup>。FAD 家族主要在质体和内质网中催化油酸生成亚油酸、亚油酸生成亚麻酸<sup>[7]</sup>。依据辅因子和亚细胞定位的不同, FAD 可划分为膜结合和可溶性 2 种类型。膜结合 FAD 包括  $\omega$ -6 FAD(催化油酸形成亚油酸)和  $\omega$ -3 FAD(催化亚油酸形成 $\alpha$ -亚麻酸), 可溶性 FAD 包括  $\Delta$ 4 FAD(催化二十碳五烯酸 EPA 生成十二碳六烯酸 DHA)、 $\Delta$ 6 FAD(催化亚油酸生成 $\gamma$ -亚麻酸)及  $\Delta$ 9 SAD(催化硬脂酸生成油酸)<sup>[8]</sup>。学者们对芝麻、油菜、向日葵、油茶等油料作物的 FAD 基因家族进行了生物信息学分析<sup>[9-12]</sup>, 在其他作物中, 例如藜麦、荞麦、茄子等植物中也有该基因家族分析相关报道<sup>[13-15]</sup>。相关研究发现, FAD 与细胞膜的稳定性和流动性也相关, 进而对植株抗逆性产生影响<sup>[16]</sup>。亚麻 FAD 基因家族研究基于转录组数据的分析<sup>[17]</sup>, 但随着高通量测序技术的发展以及亚麻全基因组数据的更新, 对于脂肪酸合成途径相关基因家族的分析有待进一步完善。我们运用生物信息学方法对亚麻全基因组的 FAD 基因家族进行分析, 包括蛋白理化性质分析、系统进化分析、保守结构域及基因结构分析、染色体定位分析、亚细胞定位预测和启动子顺式作用元件分析, 以期为其在亚麻高品质育种中的应用提供理论依据。

## 1 材料与方法

### 1.1 序列来源

从 NCBI 上获得亚麻 (version 2.0)15 条染色体的基因序列(登录号为 CP027619-CP027633)<sup>[18]</sup>, 从 phytozome 数据库(<https://phytozome.jgi.doe.gov/pz/portal.html#>)中下载蛋白质序列。

### 1.2 研究方法

1.2.1 LuFADs 基因家族成员筛选与鉴定 使用以下 2 种方法鉴定亚麻 FAD 蛋白基因。第 1 种是基于拟南芥基因组 (Version10.0, <https://www.arabidopsis.org/>)中的 27 个参考序列, 通过 BLASTP 来

鉴定亚麻基因组中的 FAD 蛋白, E 值设置为  $1.0E-10$ 。第 2 种是从 Pfam 蛋白家族数据库 (<http://pfam.xfam.org/>)下载与 FA\_Desaturase (PF00487)、FA\_Desaturase 2 (PF03405)和 TMEM189 (PF10520)结构域对应的隐马尔可夫模型 (HMM)文件作为 query, 使用 HMMER v3.3 (<http://eddylab.org/software/hmmer/hmmer-3.3.tar.gz>)中的 hmmsearch 工具来查找亚麻蛋白数据库中所有可能的 FAD 蛋白序列, E 值  $< 1e^{-5}$ 。

1.2.2 亚麻 LuFADs 基因的生物信息学分析 利用 ClustalW 对亚麻、拟南芥 FAD 蛋白序列进行多重序列比对, 接着在 MEGA11 软件中使用邻接法进行系统进化树构建。使用 Bioperl 工具(<https://bioperl.org/>)计算亚麻 FAD 家族成员的氨基酸数量、蛋白理论等电点 (PI)和分子量大小 (Mw)等指标<sup>[19]</sup>。利用本地 MEME 工具(版本 4.11.2, <http://alternatememe-suite.org/tools/meme>)搜索亚麻 FAD 家族成员序列中的保守基序。通过在线软件 (MG2C, [http://mg2c.iask.in/mg2c\\_v2.1/](http://mg2c.iask.in/mg2c_v2.1/))和 (GSDS, <http://gsds.cbi.pku.edu.cn/>)分别构建亚麻 FAD 家族成员的染色体位置图和外显子-内含子结构图。使用 CELLO (<http://cello.life.nc-tu.edu.tw/>)进行 LuFAD 家族成员亚细胞定位预测。采用 Perl 脚本提取亚麻 FAD 家族成员基因编码区上游 1 500 bp 的序列, 并利用在线数据库 PlantCARE(<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)对该基因家族成员中的顺式作用元件进行鉴定<sup>[20]</sup>。

## 2 结果与分析

### 2.1 LuFADs 基因家族成员鉴定及蛋白理化性质分析

综合 2 种方法的筛选结果, 运用 SMART (<http://smart.embl.de/>)和 NCBI-CDD (<https://www.ncbi.nlm.nih.gov/cdd/>)数据库进一步验证筛选结果中是否具有 FAD 结构域组成, 去除有错误、短小 ( $< 100$  aa)和没有 FAD 结构域的蛋白质序列, 最终鉴定了 43 个 LuFADs 基因, 并依据它们在 15 条染色体位置信息和系统发育分析结果进行命名(表 1、图 1)。编码的蛋白质范围为 152 ~ 453 个氨基酸。LuFADs 的理论分子量为 17.243 9 (LuFAD6-3) ~ 51.246 8 kD(LuSLD2), 等电点 (pI)为 6.07 ~ 10.26, LuSTAD (1~4)系列的 pI  $< 7.00$ , 呈酸性, 其余 39

个成员 pI 均大于 7, 呈碱性。LuFADs 基因蛋白的不稳定性指数范围为 76.96~91.61, 均为不稳定蛋白。亲水性平均值 (GRAVY) 为正数的有 5 个, 分别为 LuSLD (1~4) 系列和 LuFAD6-4, 表现为不亲水性, 其余 38 个家族成员均为亲水蛋白(表1)。

2.2 LuFADs 家族成员系统进化分析

运用邻接法将 LuFADs 蛋白序列(43)与拟南芥 FADs 蛋白序列(27)构建系统发育树(图 1), 可分为 4 个主要亚族: Δ12/ω-3 去饱和酶亚族、“前

端”去饱和酶亚族、Δ7/Δ9 去饱和酶亚族和 SAD 去饱和酶亚族。

Δ12/ω-3 去饱和酶亚家族包括 3 个分支: FAD6、FAD3 和 FAD6C。其中 FAD6 分支包括拟南芥的 AtFAD6 和亚麻的 LuFAD6 (1-13); FAD3 分支为膜结合 ω-3FAD 系列基因, 包括拟南芥的 AtFAD3、AtFAD3C、AtFAD3D 和亚麻的 LuFAD3C、LuFAD3D、LuFAD3-1、LuFAD3-2、LuFAD3-3、LuFAD3-4; FAD6C 分支包括拟南芥的 AtFAD6C

表 1 LuFADs 基因家族成员鉴定

基因名称	基因ID	染色体位置	氨基酸数量 /个	等电点	分子量大小 /kD	不稳定 系数	亲水性 平均值
LuFAD6-1	Lus10004175	Chr6:15636287-15637435(+)	383	8.37	43.645 8	82.96	-0.097
LuFAD6-2	Lus10021051	Chr8:5480883-5482031(+)	383	8.37	43.665 8	81.68	-0.100
LuFAD6-3	Lus10012008	Chr1:5295671-5296126(+)	152	9.53	17.243 9	79.47	-0.041
LuFAD6-4	Lus10004178	Chr6:15620008-15621871(-)	393	9.04	45.296 1	91.30	0.045
LuFAD6-5	Lus10021045	Chr8:5450383-5451492(+)	370	8.32	42.920 1	89.81	-0.057
LuFAD6-6	Lus10004177	Chr6:15626046-15627170(+)	375	7.64	43.303 4	88.93	-0.079
LuFAD6-7	Lus10021049	Chr8:5468309-5469463(+)	385	7.64	44.391 6	88.39	-0.067
LuFAD6-8	Lus10004176	Chr6:15628973-15630091(-)	373	8.45	42.831 9	87.28	-0.073
LuFAD6-9	Lus10021050	Chr8:5471361-5472503(+)	381	8.46	43.791 0	84.68	-0.059
LuFAD6-10	Lus10004180	Chr6:15611968-15613113(-)	382	8.17	44.839 2	80.31	-0.147
LuFAD6-11	Lus10021047	Chr8:5457127-5458272(-)	382	8.35	44.638 0	79.00	-0.138
LuFAD6-12	Lus10004181	Chr6:15609199-15610344(+)	382	8.34	44.868 5	80.31	-0.116
LuFAD6-13	Lus10021046	Chr8:5454415-5455539(-)	375	8.75	44.094 5	77.86	-0.125
LuFAD3-4	Lus10040660	Chr3:6170950-6173029(+)	393	8.56	44.714 1	85.05	-0.099
LuFAD3-3	Lus10018245	Chr12:4934348-4936386(-)	387	8.81	44.013 3	87.38	-0.073
LuFAD3-2	Lus10038321	Chr7:16089395-16092602(+)	393	9.31	44.765 2	81.30	-0.149
LuFAD3-1	Lus10036184	Chr12:1035256-1038244(-)	392	9.31	44.651 0	81.28	-0.147
LuFAD3D	Lus10005039	Chr1:8589844-8591694(-)	374	9.10	42.336 1	84.72	-0.213
LuFAD3C	Lus10027809	Chr1:4658469-4660574(+)	453	9.22	50.902 9	83.52	-0.248
LuFAD6C-1	Lus10035831	Chr14:14675565-14678231(+)	442	9.03	50.895 4	87.17	-0.142
LuFAD6C-2	Lus10036613	Chr11:7014248-7016887(+)	442	9.17	50.806 4	86.96	-0.123
LuDES2	Lus10011189	Chr15:1131775-1133796(+)	339	8.37	39.181 0	80.18	-0.148
LuDES1	Lus10018485	Chr8:23110465-23112504(+)	339	7.97	39.225 0	79.02	-0.142
LuSLD1	Lus10031102	Chr9:6750393-6751739(-)	449	8.66	51.159 6	90.29	0.051
LuSLD2	Lus10035482	Chr7:6125545-6126891(+)	449	8.78	51.246 8	89.42	0.039
LuSLD3	Lus10036408	Chr11:7964191-7965537(+)	449	8.58	50.937 3	91.61	0.039
LuSLD4	Lus10041083	Chr15:9247799-9249064(+)	422	9.20	47.619 6	86.82	0.021
LuADS1	Lus10006180	Chr4:7383607-7385255(+)	331	9.11	37.545 9	84.21	-0.092
LuADS2	Lus10041043	Chr15:9046976-9048747(-)	332	8.79	37.707 9	78.37	-0.138
LuADS3	Lus10041042	Chr15:9042739-9044594(+)	332	8.64	37.685 9	79.21	-0.110
LuADS4	Lus10009469	Chr14:273441-274910(-)	324	9.50	37.386 6	79.13	-0.226
LuADS5	Lus10001279	Chr1:29035163-29036635(-)	323	9.82	37.517 6	76.96	-0.273
LuADS6	Lus10019726	Chr13:4426515-4428612(-)	293	7.92	34.226 9	77.12	-0.259
LuADS7	Lus10009470	Chr14:267158-269467(+)	281	9.93	32.765 5	82.25	-0.149
LuADS8	Lus10017773	Chr6:6410686-6411935(+)	316	8.70	36.555 4	77.43	-0.231
LuADS9	Lus10023128	Chr3:21588055-21589871(-)	386	10.26	44.072 2	77.79	-0.198
LuADS10	Lus10011488	Chr15:3266147-3270693(-)	380	9.98	43.389 7	83.64	-0.115
LuFAD4-1	Lus10037808	Chr12:5694965-5695819(-)	285	8.18	31.727 1	82.78	-0.093
LuFAD4-2	Lus10017089	Chr3:24302510-24303367(-)	286	8.39	31.986 3	84.53	-0.075
LuSTAD1	Lus10018926	Chr1:28379153-28381248(+)	395	6.71	44.695 9	78.98	-0.419
LuSTAD2	Lus10028627	Chr14:1106499-1108392(+)	369	6.07	41.669 5	83.23	-0.361
LuSTAD3	Lus10027486	Chr10:2257483-2259997(-)	397	6.08	44.856 7	77.63	-0.443
LuSTAD4	Lus10039241	Chr11:17364817-17367335(-)	397	6.08	44.900 8	77.63	-0.444

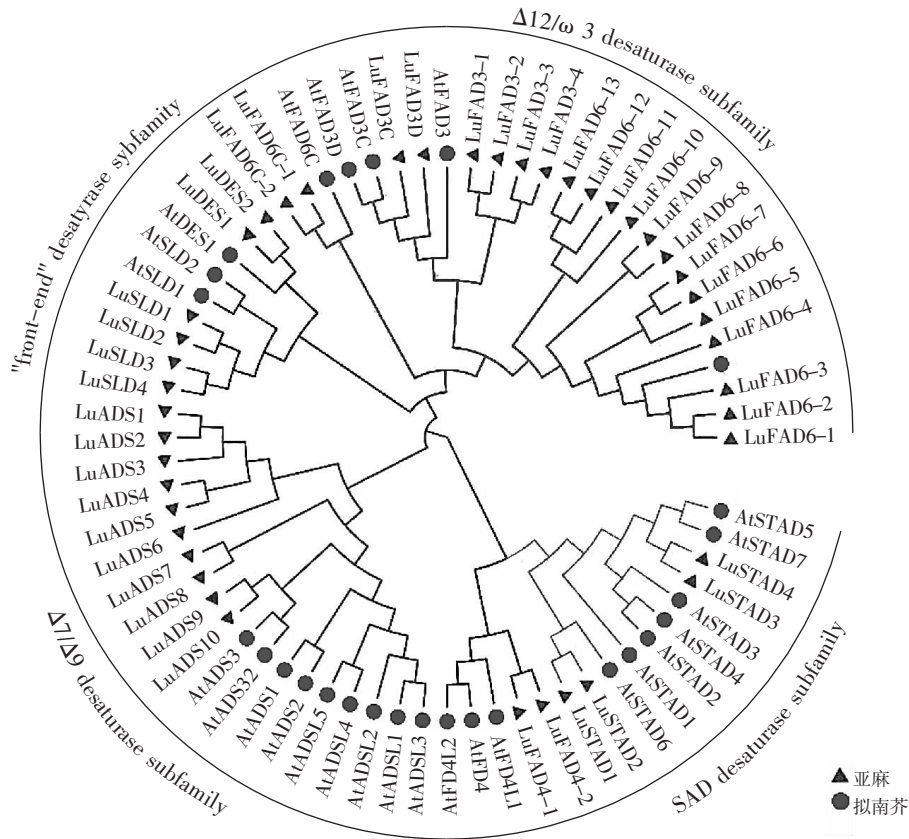


图 1 亚麻 LuFADs 和拟南芥 AtFADs 系统发育树

和亚麻的 *LuFAD6C-1*、*LuFAD6C-2*。其余 3 个亚族中, 亚麻基因成员和拟南芥基因成员均表现较高的聚集性, 特别是在  $\Delta 7/\Delta 9$  去饱和酶亚族中, 亚麻基因 *LuADS* (1-10) 和拟南芥 *AtADS* 系列基因分别各自聚为 1 个分支。

2.3 亚麻 FAD 家族保守结构域和基因结构分析

对亚麻 LuFADs 基因家族进行保守结构域分析 (表 2, 图 2) 可知, 同一亚族中保守结构域的分布

较为相似, LuADS 亚组中均包含 motif-1、motif-3 和 motif-7; LuSTAD 亚组中均只含 motif-10; LuDES 和 LuSLD 亚组中含有 motif-2; LuFAD6 亚组中, 除 *LuFAD6-3* 外其他基因均含有 motif-1、motif-2、motif-4、motif-5、motif-6、motif-8 和 motif-9; *LuFAD4-1* 和 *LuFAD4-2* 不含有 motif。

对 LuFADs 基因家族进行外显子-内含子结构分析 (图 2) 显示, 在 LuFAD6 亚组中, 只有 *Lu-*

表 2 保守基序 motif 序列分布情况

保守基序名称	具体序列分布情况
motif-1	FGICFSYHRNLAHHSFKVPKWLEYFCA YCGA HALQGD P M T W V S N H R F H H Q Y T D T E R D P H S P I E G F W H S H M S W I F D S G S I R E K C G K R D N V G D L K S Q A F Y R W
motif-2	NSAAHVWGSQPWKTN D L S R N N W V V A I L T F G E G W H N N H H A F K Y S A R Q G L E W W Q I D I T W
motif-3	NKVFHNITDTHVAHHLFSTIPHYHAMEATKAIKPV L G E Y Y Q
motif-4	MITEDALPTYZTMJNTLDGVRDETGASQPWAJWTRAWTAENRHGDLLRKYLYLSGRVDMRQIEKTIQYLI G S G M D P G T E N P Y L G F I Y T S F Q E R A T F I
motif-5	HGNTARLAKDGGDPKLARICGTIAADEKRHETAYTKIVEKLF E I D P D G A V L A L A D M M R K K I S M P A H L M Y D G E D P K L F D N Y S S V A Q R J G V Y T A K D Y A D I L E
motif-6	PYSKPPFTLADJKKAIPPHCFKRSJPRSFA Y L A Y D L T V I T I L Y H I A T T Y F H L L P K P L S
motif-7	DDVVGFIHLSALLVPYFSWKHSRRHHANTGSLERDES F V P K K K D E J R W H F K Y
motif-8	VAWPLYWAAQGC M F V A V W V L A H E C G H H A F
motif-9	HTHPKLPHYDSSEWDWLRGALTTVDRDYG
motif-10	GFASHFYPMSPYINDRERFGIFLSDV G M L A M W F A L Y K L A L A K G V G W V V C V Y G V P L L L Q N G L F V T I T Y L H

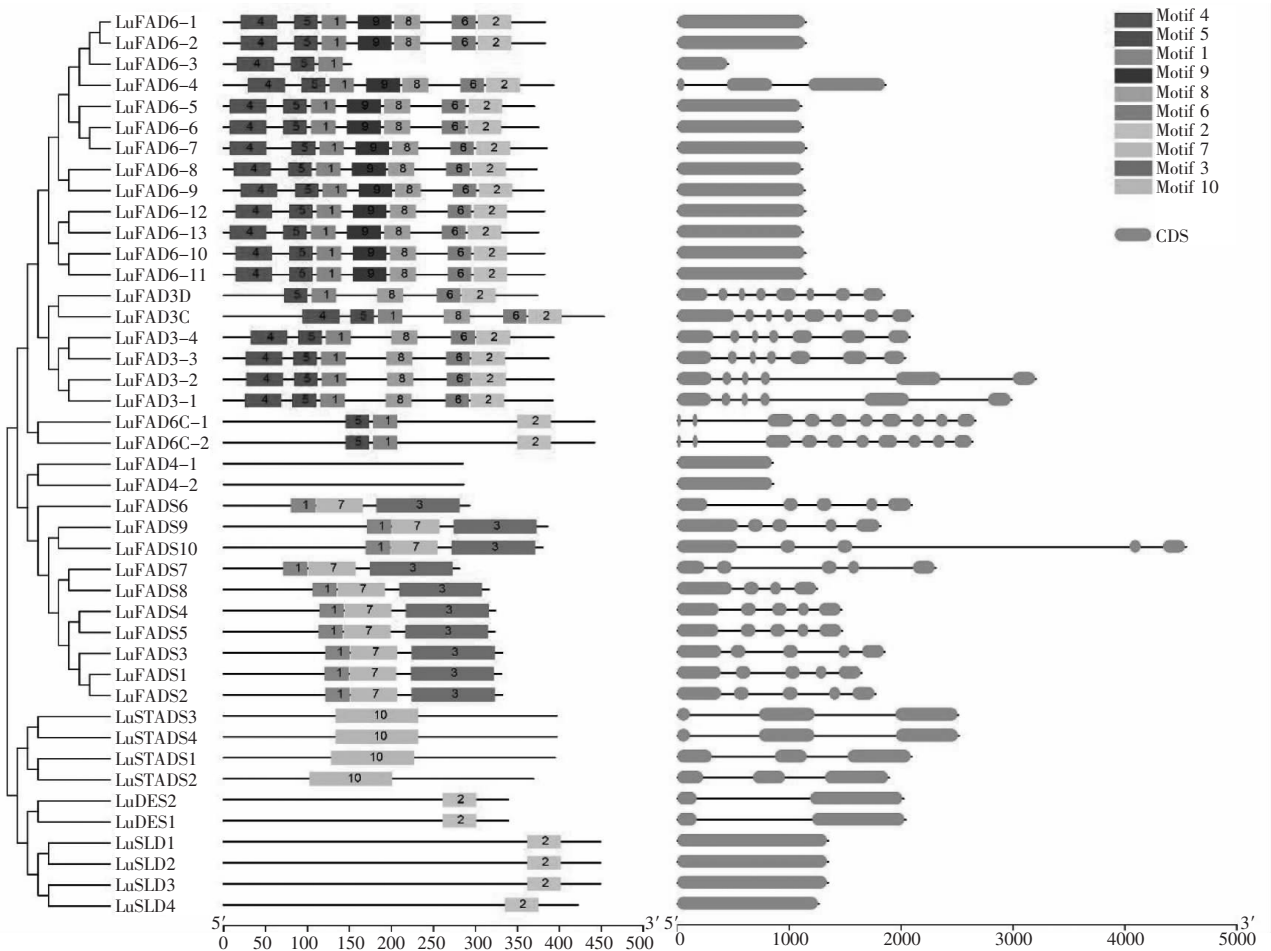


图2 亚麻FAD家族成员的系统进化、保守motif 和基因结构分析

*FAD6-4* 含有 2 个内含子，其余 12 个基因均只有外显子；*LuFAD4-1*、*LuFAD4-2* 和 *LuSLD* 系列没有内含子，剩余基因有不同数量(1~7, 9)内含子。*LuSLD* (1~3)的 CDS 区最长，大于 1 000 bp；*LuADS10* 的基因长度最长，约为 4 500 bp，*LuFAD6-3* 基因最短，不足 500 bp。

综合家族系统进化可以看出，同一亚组中的基因具有较为相似的结构。LuSTAD 亚组中均含有 2 个内含子。LuADS 亚组中，除 *LuADS8* 的显子数量为 4 外，其余均为 5。*LuFAD6C-1* 和 *LuFAD6C-2* 内含子最多，具有 9 个。*LuADS10* 的内含子序列最长。

#### 2.4 LuFADs 基因家族成员染色体定位和亚细胞预测

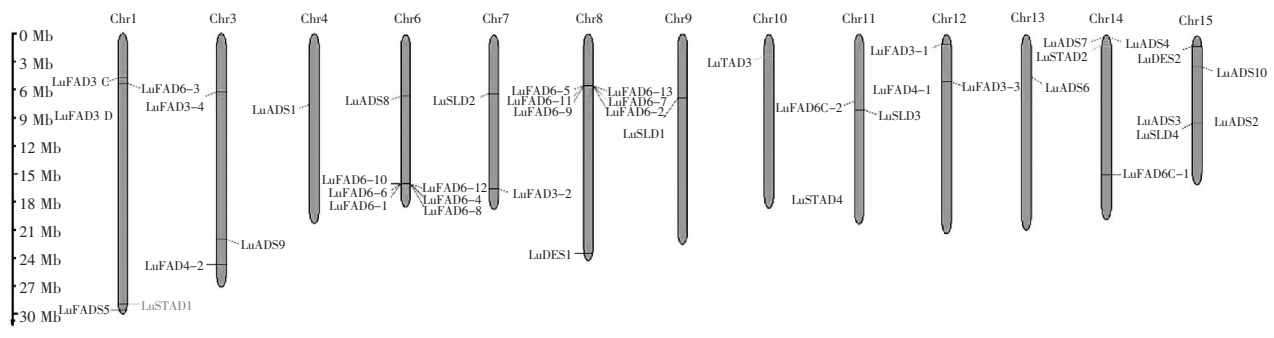
由图 3 可知，LuFADs 基因家族随机分布于除第 2 号和 5 号染色体(Chr2、Chr5)之外的所有亚麻染色体上，分布广泛。Chr4、Chr9、Chr10 和 Chr13 上各自有 1 个基因成员，Chr7 上有 2 个基

因，Chr3、Chr11 和 Chr12 上各分布着 3 个基因，Chr14 上有 4 个基因，Chr1 和 Chr15 各为 5 个基因。Chr6 和 Chr8 的 LuFADs 家族成员最多，均为 7 个，其中在 Chr6 上 *LuFAD6-1*、*LuFAD6-4*、*LuFAD6-6*、*LuFAD6-8*、*LuFAD6-10* 和 *LuFAD6-12* 距离较近，在 Chr8 染色体上 *LuFAD6-2*、*LuFAD6-5*、*LuFAD6-7*、*LuFAD6-9*、*LuFAD6-11* 和 *LuFAD6-13* 的距离较为接近。

对 LuFADs 家族成员进行亚细胞定位预测(表 3)可知，定位结果分为 4 类，包括叶绿体、细胞质、细胞核和线粒体。定位在叶绿体上的家族成员最多，共 32 个基因。细胞质上 7 个基因成员，细胞核上 3 个基因，线粒体上 1 个基因。其中 *LuFAD6*、*LuFAD4* 和 *LuSTAD* 成员的基因均定位在叶绿体上。

#### 2.5 LuFAD 启动子顺式作用元件分析

对 LuFADs 进行启动子顺式作用元件分析(图

图 3 *LuFADs* 基因家族成员染色体分布情况表 3 *LuFADs* 的亚细胞定位

基因名称	位置	得分	基因名称	位置	得分
<i>LuFAD6-1</i>	叶绿体	0.995 6	<i>LuDES1</i>	细胞质	0.602 3
<i>LuFAD6-2</i>	叶绿体	0.995 6	<i>LuSLD1</i>	细胞核	0.746 4
<i>LuFAD6-3</i>	叶绿体	0.994 8	<i>LuSLD2</i>	细胞核	0.746 4
<i>LuFAD6-4</i>	叶绿体	0.999 7	<i>LuSLD3</i>	细胞质	0.969 9
<i>LuFAD6-5</i>	叶绿体	0.999 2	<i>LuSLD4</i>	叶绿体	0.944 8
<i>LuFAD6-6</i>	叶绿体	0.998 3	<i>LuADS1</i>	叶绿体	0.864
<i>LuFAD6-7</i>	叶绿体	0.998 3	<i>LuADS2</i>	细胞质	0.964 2
<i>LuFAD6-8</i>	叶绿体	0.996 7	<i>LuADS3</i>	细胞质	0.940 4
<i>LuFAD6-9</i>	叶绿体	0.999 9	<i>LuADS4</i>	细胞质	0.602 3
<i>LuFAD6-10</i>	叶绿体	0.996	<i>LuADS5</i>	细胞质	0.602 3
<i>LuFAD6-11</i>	叶绿体	0.996 7	<i>LuADS6</i>	线粒体	0.665 4
<i>LuFAD6-12</i>	叶绿体	0.997 4	<i>LuADS7</i>	叶绿体	0.978 4
<i>LuFAD6-13</i>	叶绿体	0.998 3	<i>LuADS8</i>	细胞核	0.582 5
<i>LuFAD3-4</i>	细胞质	0.940 4	<i>LuADS9</i>	叶绿体	0.999
<i>LuFAD3-3</i>	叶绿体	0.562 8	<i>LuADS10</i>	叶绿体	0.998 3
<i>LuFAD3-2</i>	叶绿体	0.991 3	<i>LuFAD4-1</i>	叶绿体	0.829 4
<i>LuFAD3-1</i>	叶绿体	0.897 3	<i>LuFAD4-2</i>	叶绿体	0.829 4
<i>LuFAD3D</i>	叶绿体	0.999 8	<i>LuSTAD1</i>	叶绿体	0.999 5
<i>LuFAD3C</i>	叶绿体	0.991 1	<i>LuSTAD2</i>	叶绿体	0.995 1
<i>LuFAD6C-1</i>	叶绿体	0.996 6	<i>LuSTAD3</i>	叶绿体	0.968 7
<i>LuFAD6C-2</i>	叶绿体	0.980 4	<i>LuSTAD4</i>	叶绿体	0.968 7
<i>LuDES2</i>	叶绿体	0.639 8			

4)可知,该基因家族含有的元件可分为4类,光响应元件(G-box、Box 4、GATA-motif、I-box、TCT-motif、Sp1、TCCC-motif、GT1-motif)、激素响应元件(TGACG-motif、CGTCA-motif、ABRE、TGA-element、AuxRR-core、TCA-element、P-box、GARE-motif、TATC-box)、胁迫响应元件(MBS、TC-rich repeats、LTR、ARE、GC-motif)和生长发育相关元件(A-box、O2-site、CCAAT-box)。其中抗氧化反应元件 ARE 数量最多(89),其次是 G-box(77), Box 4(74)排第3, TGACG-motif 和 CGTCA-motif 在基因家族中的含量分布数量相同(72)。基因 *LuFAD3-3* 含元件数最多,为35个; *LuADS3* 和 *LuADS7* 的元件数最少,均为6个。

### 3 讨论与结论

近年来, FAD 基因家族在多种植物物种中得

到鉴定验证,其家族成员数量不尽相同,如甜荞麦中10个<sup>[14]</sup>,大豆中29个<sup>[21]</sup>,陆地棉中41个<sup>[22]</sup>。本研究分析鉴定了亚麻43个 *LuFADs* 基因家族成员, *gDNA* 长度为456~4547 bp,编码的氨基酸为152~453个。等电点(pI)6.07~10.26。除 *LuSLD* (1~4)系列和 *LuFAD6-4*,其余38个家族成员均为亲水蛋白。有研究者对茄子的 FAD 家族序列进行理化性质分析,表明分子量与成员所含氨基酸数量成正比,大部分家族成员为亲水性蛋白,这与本研究结果相似<sup>[15]</sup>。系统发育树分析聚为4类,同一亚族中保守结构域的分布较为相似,在  $\Delta 7/\Delta 9$  去饱和酶亚族中均含 motif-1、motif-3 和 motif-7,在 SAD 去饱和酶亚族中均只含 motif-10;在“前端”去饱和酶亚族中只含 motif-2,在  $\Delta 12/\omega-3$  去饱和酶亚族中 *LuFAD6*、*LuFAD6C*、*LuFAD3C*

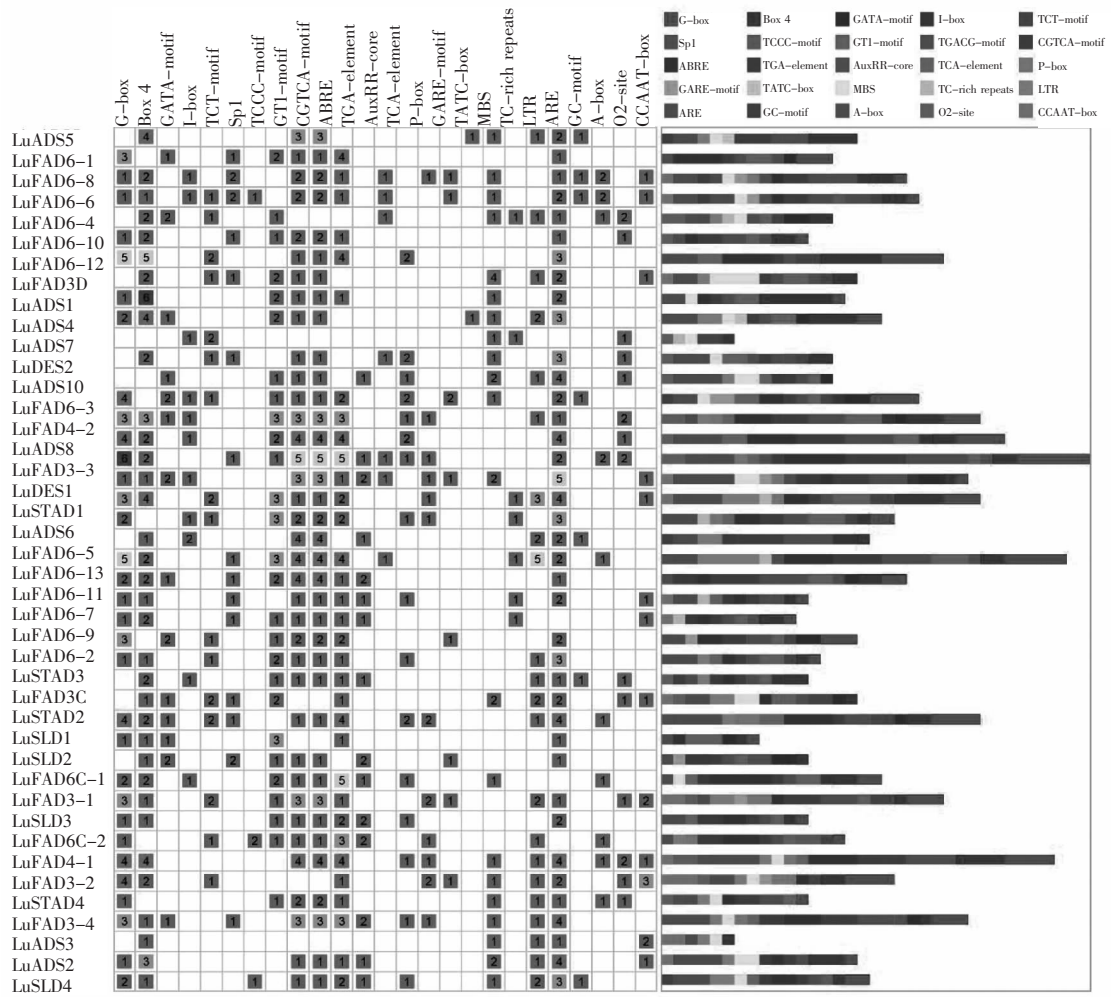


图 4 LuFAD 基因家族成员启动子顺式作用元件分布情况

和 *LuFAD3D* 各分支的保守结构域相似。Liu 等 [23] 用最大似然法将核桃 FADs 蛋白序列与拟南芥 FADs 蛋白序列一起构建系统发育树，分析 JrFADs 基因家族，也分为 4 个亚家族，与本研究相一致。

对基因外显子 - 内含子结构分析表明，同一亚组中的基因具有较为相似的基因结构。Xue 等 [10] 对油菜 FAD 基因家族进行基因结构分析，各个 FAD 亚家族的外显子 - 内含子结构高度保守。Liu 等 [23] 分析 JrFADs 基因家族，结果表明同一亚家族中的基因具有相似的基因结构。Cheng 等 [24] 对香蕉同一亚科中的 MaFAD 家族成员进行分析，显示出相似的内含子 / 外显子结构和内含子相，并且它们编码的蛋白质由相似的基序组成。

染色体定位和亚细胞预测发现，LuFADs 基因家族随机地分布在除 Chr2 和 Chr5 之外的所有亚麻染色体上。亚细胞定位在叶绿体上的家族成员最多，共有 32 个基因。Xue 等 [10] 的研究表明，

油菜 3 个芸薹属物种中 FAB2、FAD4、FAD6 和 FAD8 蛋白质的所有成员均位于叶绿体中。启动子顺式作用元件分析得出，抗氧化反应元件 ARE 的数量最多，为 89 个，基因 *LuFAD3-3* 所含元件数最多。Yasemin 等 [25] 对向日葵的 FAD 基因家族的研究表明，主要定位在叶绿体和内质网膜中，其中 FAB2 家族成员位于叶绿体中，可能与 FAD 在细胞不同部位脂肪酸去饱和中的作用有关。综合分析向日葵、芝麻、油菜、可可树和亚麻的 FAD 基因家族，在氧化还原过程、脂肪酸生物合成过程或脂质代谢过程中发挥作用，这与本研究结果相似。

参考文献:

[1] 伍叶娜, 潘根, 姜慧, 等. 亚麻全基因组关联分析研究进展[J]. 中国麻业科学, 2022, 44(2): 109-118.

[2] 张辉, 贾霄云, 高凤云, 等. 胡麻 [M]. 北京: 中

- 国农业科学技术出版社, 2021.
- [3] 廖振林, 李倩滢, 陈俊杰, 等. 亚麻籽油组分的功能活性研究进展[J]. 现代食品科技, 2021, 37(11): 379–389; 337.
- [4] 唐佳芮, 魏冰, 石珊珊.  $\alpha$ -亚麻酸分离纯化技术研究进展[J]. 中国油脂, 2020, 45(8): 28–31; 38.
- [5] LEE K R, CHEN G Q, KIM H U. Current progress towards the metabolic engineering of plant seed oil for hydroxy fatty acids production[J]. Plant Cell Reports, 2015, 34(4): 603–15.
- [6] POVKHOVA L V, MELNIKOVA N V, ROZHMINA T A, et al. Genes associated with the flax plant type (oil or fiber) identified based on genome and transcriptome sequencing data[J]. Plants, 2021, 10(12): 2616.
- [7] CHELLAMUTHU M, KUMARESAN K, SUBRAMANIAN S. Increase in  $\alpha$ -linolenic acid content by simultaneous expression of fatty acid metabolism genes in Sesame (*Sesamum indicum* L.) [J]. Physiology and Molecular Biology of Plants, 2022, 28(3): 559–572.
- [8] 吴端, 王力军, 杨仕梅, 等. 植物种子  $\alpha$ -亚麻酸形成及调控机理研究进展[J]. 植物遗传资源学报, 2020, 21(1): 49–62.
- [9] WANG L H, YU S, TONG C B, et al. Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis[J]. Genome Biology, 2014, 15(2): 39–46.
- [10] XUE Y, CHEN B, WANG R, et al. Genome-wide survey and characterization of fatty acid desaturase gene family in brassica napus and its parental species [J]. Applied Biochemistry and Biotechnology, 2018, 184(2): 582–598.
- [11] 于海峰, 韩平安, 李美娜, 等. 基于 RNA-Seq 技术的向日葵油酸形成的转录组学分析[J]. 中国油料作物学报, 2018, 40(6): 769–776.
- [12] 江南, 谭晓风, 张琳, 等. 基于 RNA-Seq 的油桐种子  $\alpha$ -亚麻酸代谢途径及相关基因分析[J]. 林业科学, 2014, 50(8): 68–75.
- [13] 刘江, 王育川, 董艳辉, 等. 藜麦 *FAD2* 基因鉴定及生物信息学分析[J]. 分子植物育种, 2022, 20(3): 715–721.
- [14] 赵训超, 魏玉磊, 丁冬, 等. 甜荞麦脂肪酸脱氢酶基因 (*FeFAD*) 家族的鉴定与分析[J]. 东北农业科学, 2021, 46(1): 36–41.
- [15] 朱宗文, 张爱冬, 吴雪霞, 等. 生物信息学鉴定分析茄子脂肪酸去饱和酶 (FAD) 基因家族 [J/OL]. 分子植物育种: 1–15 (2021–05–12) [2022–05–18]. <http://kns.cnki.net/kcms/detail/46.1068.S.20210512.0826.002.html>
- [16] DOMINGUEZ T, HERNANDEZ M L, PENNYCOOKE J C, et al. Increasing  $\omega$ -3 desaturase expression in tomato results in altered aroma profile and enhanced resistance to cold stress[J]. Plant Physiology, 2010, 153(2): 655–665.
- [17] YOU F M, LI P C, KUMAR S, et al. Genome-wide identification and characterization of the gene families controlling fatty acid biosynthesis in flax (*Linum usitatissimum* L.) [J]. Journal of Proteomics & Bioinformatics, 2014, 7(10): 310.
- [18] YOU F M, XIAO J, LI P, et al. Chromosome-scale pseudomolecules refined by optical, physical and genetic maps in flax[J]. Plant Journal, 2018, 95(2): 371–384.
- [19] 刘畅畅, 黄敏. 玉米 *LEC1* 基因家族的鉴定与生物信息学分析[J]. 甘肃农业科技, 2022, 53(1): 31–37.
- [20] 张国琴, 葛玉彬, 张正英. 高粱耐冷基因全基因组鉴定及进化分析[J]. 甘肃农业科技, 2020(12): 37–41.
- [21] CHI X Y, YANG Q L, LU Y D, et al. Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*) [J]. Plant Molecular Biology Reporter, 2011, 29(4): 769–783.
- [22] FENG J, DONG Y, LIU W, et al. Genome-wide identification of membrane-bound fatty acid desaturase genes in *Gossypium hirsutum* and their expressions during abiotic stress[J]. Sci Rep, 2017, 1(7): 1–12.
- [23] LIU K, ZHAO S, WANG S, et al. Identification and analysis of the FAD gene family in walnuts (*Juglans regia* L.) based on transcriptome data[J]. BMC Genomics, 2020, 21(1): 299–303.
- [24] CHENG C, LIU F, SUN X, et al. Genome-wide identification of FAD gene family and their contributions to the temperature stresses and mutualistic and parasitic fungi colonization responses in banana[J]. International Journal of Biological Macromolecules, 2022, 204: 661–676.
- [25] YASEMIN C A, NECDET M U, MEHMET C B, et al. Comparative identification and evolutionary relationship of fatty acid desaturase (FAD) genes in some oil crops: the sunflower model for evaluation of gene expression pattern under drought stress[J]. Biotechnology & Biotechnological Equipment, 2018, 32(4): 846–857.