

微藻中二酰基甘油酰基转移酶基因(DGAT) 结构及进化分析

杨 晓 梁成伟

青岛科技大学海洋科学与生物工程学院, 青岛 266042

摘要 从数据库中获取微藻及高等植物中的二酰基甘油酰基转移酶基因(DGAT),对 4 种酶通过基因和蛋白质结构的比较及系统发育树的构建进行多样性和进化程度分析。为了进一步分析微藻中 DGAT 的进化过程,对其进行正选择位点的检验,推测其在进化过程中受到的选择压力。结果显示:微藻中不同类型 DGAT 的基因结构及蛋白质结构差异较大,微藻与其他高等植物相比,同一种类型的 DGAT 的蛋白质结构和基因结构也有较大差别。同时发现微藻中 DGAT 同其他高等植物中的 DGAT 有不同的进化过程。选择压力分析结果显示,微藻及高等植物的 DGAT2 和 WS/DGAT 存在可信度较高的正选择位点,DGAT1 和 DGAT3 未发现正选择位点。对微藻中 DGAT 的进化研究有助于破译其在油脂积累中的作用,推动可再生能源的研究。

关键词 微藻;二酰基甘油酰基转移酶;基因结构;正选择位点;系统发育

中图分类号 Q 939.9 **文献标识码** A **文章编号** 1000-2421(2018)06-0030-010

能源是人类生存与发展的生物基础,传统的能源有水、木材、石油、天然气和煤炭等^[1]。但是随着经济迅速发展,石油、煤炭等过度消耗,能源面临枯竭的危机,许多国家把研究的重点转移到可再生生物能源的开发利用上,以此替代化石能源^[2]。微藻是第 3 代生物能源的潜在资源,它们有着以自养方式生产、不占用农业土地资源、可以在海水或废水中培养等优点,与陆地高等植物相比,微藻光合能力强,许多微藻脂质含量超过其干质量的 50%,因此,被看作当今最具有开发前途的能源之一^[3]。小球藻含油率 8%~32%,三角褐指藻含油率为 20%~30%,微拟球藻含油率 31%~68%^[4-6],莱茵衣藻也是开发可再生柴油的重点研究对象。Xiao 等^[7]发现 *Nannochloropsis oceanic* IMET1 通过半连续培养油脂含量高达自身干质量的 69%。Moheimani 等^[8]通过跑道池长期半连续培养颗石藻(*Pleurochrysis carterae*)后测得生物质产率为 0.19 g/(L·d),含油率占 30%。Chiu 等^[9]通过柱式反应器培养微拟球藻(*Nannochloropsis oceanica*)同样得到较高的产油率。在微藻及高等植物的 TAG 生物合成途径中,DGAT 是二脂酰甘油向三

脂酰甘油转换的关键酶,是植物油脂储存积累的关键限速酶。一些研究已经证明植物中 DGAT 在控制脂肪酸转变为甘油三酯的质量和数量上扮演着重要的角色。微拟球藻高产油藻 DGAT 基因转录本的绝对丰度明显高于其他低产油微藻^[10]。张晓舟^[11]通过诱变使得莱茵衣藻突变菌株油脂含量显著提高,实时荧光定量 PCR 结果显示突变菌株的 DGAT1、DGAT2、DGAT3 基因转录水平极显著提高。最近,在对欧洲油菜(*Brassica napus*)DGAT1 的研究中证明抑制 DGAT1 基因的表达可以减少种子中油含量并且抑制发芽率,还会导致严重的发育异常^[12]。此外,大豆(*Glycine max*)中 DGAT2 不同程度的表达使种子油含量不同程度地增加^[13]。正反向基因研究结果揭示了在一些植物中 DGAT 的突变将直接影响油含量。根据细胞、结构和亚细胞定位可将 DGAT 分为 4 种不同的类型:DGAT1、DGAT2、DGAT3(CytoDGAT)和 WS/DGAT。一些高等植物中的 DGAT1 和 DGAT2 已得到广泛的研究。Turchetto-Zolet 等^[14]通过系统发育树、基因结构和表达分析揭示高等植物 DGAT 的多样性和进化过程,发现 DGAT1、DGAT2、DGAT3 和 WS/

收稿日期:2018-01-14

基金项目:国家自然科学基金项目(31770393)

杨 晓,硕士研究生。研究方向:藻类生物学。E-mail: 511653224@qq.com

通信作者:梁成伟,博士,副教授。研究方向:藻类生物学。E-mail: zpyangxiao1993@163.com

DGAT 是有很大差异的基因并在植物中有不同的来源,它们在不同器官和组织中有不同的表达模式。Fei 等^[15]发现绿藻的 DGAT3 对低温和磷缺乏有着显著的反应,并通过诱变研究 DGAT3 的启动子原件。Wei 等^[16]研究发现微拟球藻 DGAT1 中的 NoDGAT1A 可以通过过表达而促进 TAG 的积累,推动了微藻遗传工程的研究。

尽管 DGAT 被认为是油脂合成途径中重要的限速酶,但是其在微藻中的研究还不是很广泛,本研究选取几种具有代表性的高产油脂微藻,从微观的角度,分析微藻中 DGAT 蛋白质结构、基因结构及进化关系,并同高等植物的 DGAT 进行比较,旨在为探明 DGAT 在微藻高产油脂方面的作用提供理论依据。

1 材料与方法

以 Turchetto-Zolet 等^[14]报道的来自玉米(*Zea mays*, NP_001288553.1)、麻风树(*Jatropha curcas*, NP_001292973.1)、黏红酵母(*Rhodotorula glutinis*, ABC41546.1)及拟南芥(*Arabidopsis*

thaliana, AED96346.2)中的 DGAT 为目标序列;在 NCBI 和 Phytozome 数据库中通过同源性搜索,获得微拟球藻(*Nannochloropsis oceanica*)、莱茵衣藻(*Chlamydomonas reinhardtii*)、团藻(*Volvox carteri* f. *nagariensis*)、三角褐指藻(*Phaeodactylum tricornutum*)及小球藻(*Chlorella variabilis*)这 5 种微藻和 27 种高等植物的 DGAT 编码区序列和蛋白质序列,并得到相应的 DGAT 的基因组序列,在微藻中未找到 WS/DGAT 序列,同时原始单细胞微藻如蓝藻中迄今为止未发现 DGAT,故不纳入研究范围。本研究中所用到的 DGAT 序列所属的物种名、基因 ID、序列号及蛋白质长度见表 1。

1.1 序列比对和进化分析

微藻及高等植物中 DGAT 的编码区序列和蛋白质序列在分子进化遗传分析软件中通过 ClustalW 采用默认设置进行比对(MEGA version 7.0)^[17],多重比对需要在最后的分析中手动对齐,采用 MEGA 7.0 软件中的最大似然法(maximum likelihood, ML)进行蛋白质序列比对后进化模型的构建。

表 1 本研究中所用到的 DGAT 序列所属物种名、基因 ID、序列号及蛋白质长度

Table 1 Species, gene name, accession numbers and protein length of DGAT sequences retrieved in this study

| 物种 Species | 分类术语 Taxa terminologies | 基因 ID Gene ID | 数据库 Database | 序列号 Access | 长度(aa) Length |
|----------------------------------|----------------------------|------------------|-----------------|--------------------|------------------|
| <i>Chlorella vulgaris</i> | Cvu | DGAT1 | NCBI | ALP13865.1 | 460 |
| | | DGAT1 | NCBI | ALP13863.1 | 460 |
| <i>Chlorella variabilis</i> | Cva | DGAT1 | NCBI | XP_005842809.1 | 445 |
| <i>Solanum pennellii</i> | Spe | DGAT1 | NCBI | XP_015059289.1 | 516 |
| <i>Solanum lycopersicum</i> | Sly | DGAT2 | NCBI | XP_010317058.1 | 304 |
| | | WS/DGAT | Phytozome | Solyc01g011430.2.1 | 506 |
| <i>Medicago truncatula</i> | Mtr | DGAT1 | NCBI | XP_003595231.2 | 539 |
| | | DGAT2 | NCBI | XP_003612436.1 | 321 |
| | | DGAT3 | NCBI | XP_003609890.1 | 341 |
| | | WS/DGAT | Phytozome | Mtr_XP_003588522.1 | 483 |
| <i>Zea mays</i> | Zma | DGAT1 | NCBI | NP_001288553.1 | 494 |
| <i>Perilla frutescens</i> | Pfr | DGAT1 | NCBI | AAG23696.1 | 534 |
| <i>Xanthoceras sorbifolium</i> | Xso | DGAT1 | NCBI | AGO32048.1 | 518 |
| | | DGAT2 | NCBI | AGO32049.1 | 333 |
| <i>Sesamum indicum</i> | Sin | DGAT1 | NCBI | NP_001291334.1 | 543 |
| | | DGAT2 | NCBI | XP_011098009.1 | 335 |
| <i>Olea europaea</i> | Oeu | DGAT1 | NCBI | AAS01606.1 | 532 |
| <i>Nicotiana tabacum</i> | Nta | DGAT1 | NCBI | NP_001313014.1 | 532 |
| | | DGAT2 | NCBI | NP_001312219.1 | 332 |
| <i>Lotus japonicus</i> | Lja | DGAT1 | NCBI | AAW51456.1 | 511 |
| <i>Jatropha curcas</i> | Jcu | DGAT1 | NCBI | NP_001292973.1 | 352 |
| <i>Punica granatum</i> | Pgr | DGAT1 | NCBI | AFX71924.1 | 540 |
| <i>Volvox carteri</i> | Vca | DGAT1 | NCBI | XP_002957418.1 | 464 |
| | | DGAT2 | NCBI | XP_002947215.1 | 325 |
| | | DGAT3 | NCBI | XP_002946808.1 | 370 |
| <i>Chlamydomonas reinhardtii</i> | Cre | DGAT1 | NCBI | XP_001692593.1 | 387 |
| | | DGAT2 | NCBI | XP_001694904.1 | 324 |
| | | DGAT2 | NCBI | XP_001691447.1 | 346 |
| | | DGAT3 | NCBI | XP_001691342.1 | 346 |

续表 1 Continued Table 1

| 物种 Species | 分类术语 Taxa terminologies | 基因 ID Gene ID | 数据库 Database | 序列号 Access | 长度(aa) Length |
|---------------------------------|----------------------------|------------------|-----------------|-------------------|------------------|
| <i>Gonium pectorale</i> | Gpe | DGAT1 | NCBI | KXZ53651.1 | 277 |
| | | DGAT3 | NCBI | KXZ54505.1 | 314 |
| <i>Vernicia fordii</i> | Vfo | DGAT3 | NCBI | AGL81309.1 | 278 |
| <i>Ricinus communis</i> | Rco | DGAT2 | NCBI | NP_001310616.1 | 340 |
| | | WS/DGAT | Phytozome | Sco_29742.m001420 | 506 |
| <i>Solanum tuberosum</i> | Stu | DGAT2 | NCBI | XP_006365015.1 | 333 |
| <i>Nannochloropsis oceanica</i> | Noc | DGAT2 | NCBI | AQT19709.1 | 363 |
| | | DGAT2 | NCBI | EWM23187.1 | 360 |
| <i>Phaeodactylum tricornutu</i> | Ptr | DGAT2 | NCBI | AFM37314.1 | 329 |
| | | DGAT2 | NCBI | AFQ23659.1 | 359 |
| <i>Emiliania huxleyi</i> | Ehu | DGAT2 | NCBI | XP_005760702.1 | 334 |
| <i>Camelina sativa</i> | Csa | DGAT3 | NCBI | AQM52372.1 | 285 |
| | | WS/DGAT | NCBI | XP_019091450.1 | 469 |
| <i>Rhodotorula glutinis</i> | Rgl | DGAT3 | NCBI | ABC41546.1 | 221 |
| <i>Glycine max</i> | Gma | DGAT3 | NCBI | XP_003542403.1 | 327 |
| | | WS/DGAT | Phytozome | Glyma.09G196400.1 | 480 |
| <i>Cicer arietinum</i> | Car | DGAT3 | NCBI | XP_012573346.1 | 315 |
| <i>Vigna angularis</i> | Van | DGAT3 | NCBI | XP_017423000.1 | 330 |
| <i>Phaseolus vulgaris</i> | Pvu | DGAT3 | NCBI | XP_007154487.1 | 328 |
| <i>Vitis vinifera</i> | Vvi | DGAT3 | NCBI | XP_002269582.1 | 362 |
| <i>Arabidopsis thaliana</i> | Ath | DGAT3 | Phytozome | AT1G48300.1 | 285 |
| | | WS/DGAT | Phytozome | AED91807.1 | 480 |
| | | WS/DGAT | Phytozome | AED92282.1 | 488 |
| | | WS/DGAT | Phytozome | AAU90066.1 | 481 |
| | | WS/DGAT | Phytozome | AAT69156.1 | 206 |
| <i>Raphanus sativus</i> | Rsa | WS/DGAT | NCBI | XP_018475406.1 | 481 |
| | | WS/DGAT | NCBI | JAU65157.1 | 483 |
| <i>Noccaea caerulea</i> | Nca | WS/DGAT | NCBI | JAU65157.1 | 483 |
| <i>Brassica napus</i> | Bna | WS/DGAT | NCBI | XP_013674531.1 | 476 |
| <i>Brassica rapa</i> | Bra | WS/DGAT | NCBI | XP_009131282.1 | 479 |
| <i>Manihot esculenta</i> | Mes | WS/DGAT | Phytozome | Mes_OAY49015.1 | 515 |
| <i>Arabidopsis lyrata</i> | Aly | WS/DGAT | Phytozome | AL7G45900.t1 | 500 |

1.2 基因和蛋白质结构分析

对 *DGAT1*、*DGAT2*、*DGAT3* 及 *WS/DGAT* 基因结构的比较,主要通过 GSDS2.0^[18] 服务器分析 *DGAT* 编码区序列及对应基因组序列。除此之外,还通过跨膜预测服务器 TMHMM-2.0 和 SMART^[19] 数据库对 *DGAT* 蛋白质序列进行跨膜结构预测。

1.3 正选择位点检验

采用 PAML 软件包估算微藻及高等植物的 *DGAT* 在进化中受到的选择压力^[20]。在蛋白质水平上,可以用 *dN* 和 *dS* 的比值 ω 来表示选择压力,进而判断自然选择对同义突变的固定是起推动作用还是阻碍作用^[21]。如果 $\omega=1$ 即同义替换率与非同义替换率相等,则被认为是中性进化。如果 $\omega<1$,即 $dN<dS$,则代表净化选择;与之相反,若 $\omega>1$,即 $dN>dS$,代表正选择。 $\omega>1$ 的显著性可被当作蛋白质发生适应性进化的证据^[22]。采用位点模型、分支模型进行分析。

1) 位点特异性模型。通过位点模型(site model)研究位点特异性选择情况,采用以下 6 个最大似

然模型: $M0$ (单一 ω)、 $M1a$ (近中性)、 $M2a$ (正选择)、 $M3$ (离散)、 $M7(\beta)$ 和 $M8(\beta \& \omega)$ ^[23]。

2) 分支特异性模型。在 *DGAT1*、*DGAT2* 及 *DGAT3* 的树文件中单独标出藻类分支依次为 a、b、c。通过分支位点模型研究藻类与高等植物选择压力的不同,由于未找到微藻的 *WS/DGAT* 序列,故在此对其不作分析。设置 model=0, Nsite=0 与 $M0$ 模型进行比较,并通过 LRTs 是否显著判断可信度^[18],后验概率通过 PAML 中贝叶斯检验(BEB)计算得到^[24]。

2 结果与分析

2.1 序列比对和进化分析

为研究微藻中 4 种 *DGAT* 类型的系统发育关系,分析微藻与其他高等植物 *DGAT* 的进化关系,通过 NCBI 及 Phytozome 数据库以 Turchetto-Zolet 等^[14] 报道的部分序列作为目标序列进行同源性搜索,获得了 5 种微藻及 27 种高等植物的 *DGAT* 序列,通过 ML 法构建系统发育树,结果如图 1 所

示。从树的拓扑结构可以看出4种DGAT基本单独成支,且在DGAT1、DGAT2及DGAT3分支中均观察到藻类属于单独的分支。值得注意的是,小球藻的DGAT1单独成支,与玉米、蒺藜状苜蓿(*Medicago truncatula*)等其他高等植物的DGAT1有共同的起源,相较于其他高等植物DGAT1,团藻与莱茵衣藻的DGAT1与其自身DGAT2亲缘关系更近,与其他微藻及高等植物DGAT1有共同起源。三角褐指藻、微拟球藻、团藻及莱茵衣藻的DGAT2有明显的独立分支,这与Turchetto-Zolet^[14]等观察到的结果相符。莱茵衣藻和小球藻的DGAT3也有单独的分支,与拟南芥、大豆等高等植物有共同的起源,结果与DGAT2相似。WS/DGAT序列,不仅与DGAT1序列关系相近,且有共同的起源。通过分析系统发育树可知,不仅微藻与其他高等植物的DGAT1进化路线不同,不同微藻中的DGAT1进化路线也不同。此外,DGAT3在DGAT1、DGAT2及WS/DGAT进化之前就已经存在了,提示我们微藻及高等植物的DGAT在进化过程中可能首先进化出了DGAT3,随着时间推移再相继进化出WS/DGAT、DGAT1及DGAT2。系统发育树所表现出来的微藻单独成支,说明在漫长的进化过程中,微藻可能由于环境等因素形成独立进化路线,因此,微藻中DGAT的生物功能与高等植物中DGAT的功能可能存在差别。

2.2 基因和蛋白质结构分析

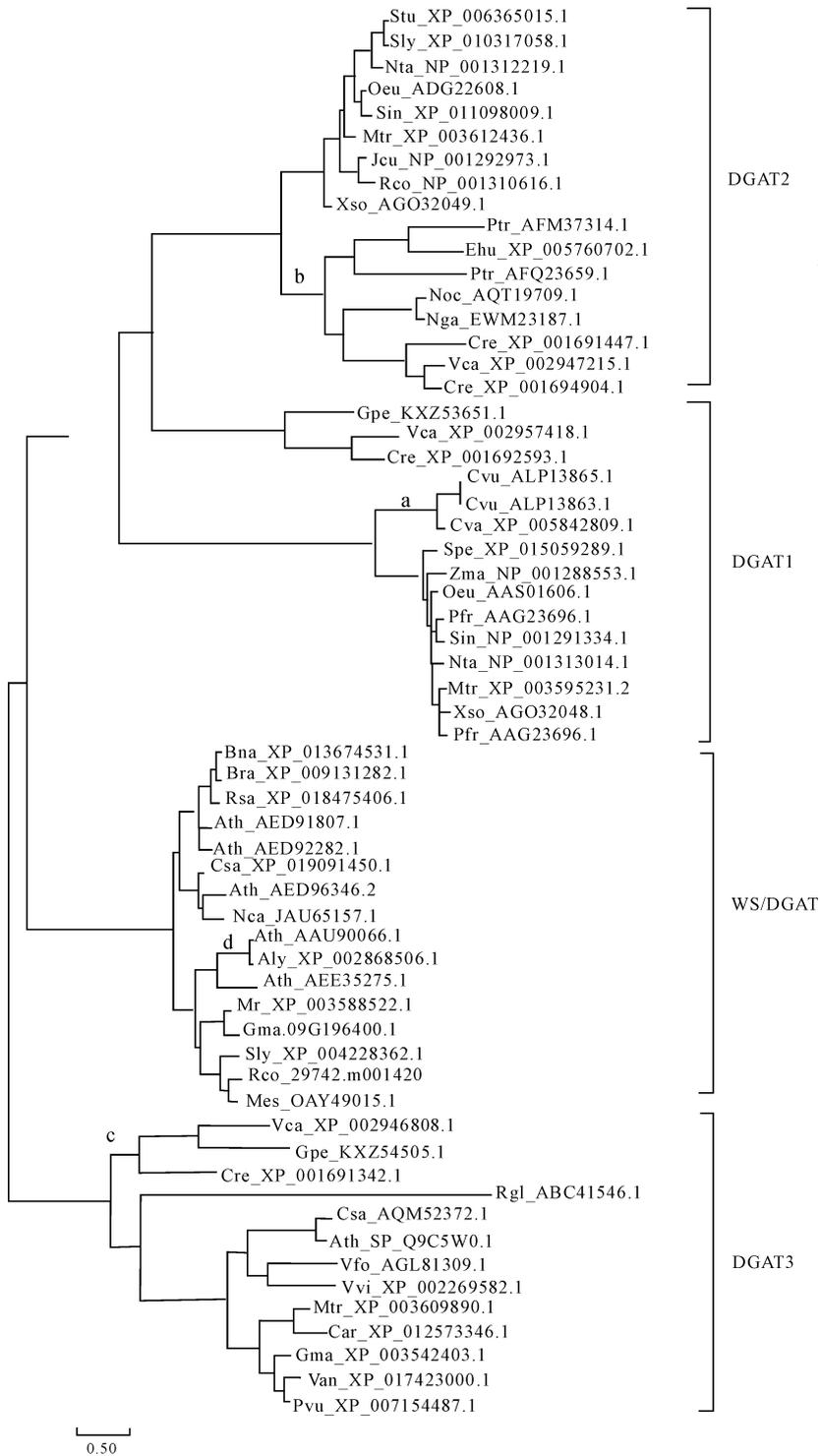
1) 蛋白质跨膜域分析。从NCBI及Phytozome数据库中得到5种微藻和27种高等植物的蛋白质序列,通过跨膜预测服务器TMHMM-2.0和SMART数据库来分析其蛋白质结构,结果如图2所示。可以看出不同DGAT类型有明显的结构模式的区分。其中,大部分高等植物的DGAT1都有8~9个跨膜域,而团藻和莱茵衣藻的DGAT1跨膜域只有5个和3个,系统发育分析中团藻和莱茵衣藻的DGAT1与DGAT2亲缘关系较近,而大多数植物的DGAT2的跨膜域数量在1~4个,与系统发育树结果一致。三角褐指藻的DGAT2有2~4个跨膜域,莱茵衣藻有2~3个跨膜域,微拟球藻及团藻的DGAT2的跨膜域数量分别为1个和2个,与其他高等植物跨膜域数量相差不大,系统发育树提示的这4种微藻与其他高等植物进化路线不同这一

点没有从跨膜域数量上体现出来。微藻及高等植物的DGAT3均未检测到跨膜域。有趣的是,一些与拟南芥同源的WS/DGAT蛋白质序列有2到3个跨膜区,其他则没有。微藻中DGAT跨膜域的分析结果与系统发育树得到的结果基本一致,显示出微藻与其他高等植物相比经历了不同的进化过程,而4种DGAT类型也经历了不同的进化过程。

2) 基因结构分析。将从NCBI及Phytozome数据库中得到的4种DGAT类型的编码区序列及对应的基因组序列,通过GSDS服务器分析微藻及其他高等植物DGAT的基因结构,结果如图3所示,分析可知不同DGAT的基因结构有明显的区别。微藻及高等植物的DGAT1的内含子较多,数量在7~15个,其中变异型小球藻、团藻及莱茵衣藻的内含子为7~8个,明显低于其他高等植物,这也与我们通过系统发育树所得到的微藻有独立分支的结果相符。团藻的DGAT2有6个内含子,莱茵衣藻2个不同种的DGAT2分别有8个和12个内含子,大多数高等植物的内含子为8个。提示我们不仅微藻的内含子与其他高等植物不同,同一种微藻的DGAT也有不同的内含子数量。本研究中除了团藻含有3个内含子,大多物种的DGAT3基因含有1个内含子,表明在植物进化过程中内含子发生了获得/缺失事件。大多数物种的WS/DGAT基因有6个内含子。4种DGAT中微藻与其他高等植物内含子差异较大,提示我们在进化过程中微藻及高等植物发生了不同程度的分化。内含子的缺失和获得是DGAT基因进化的模式,通过分析4种DGAT的内含子数量似乎进一步验证了系统发育树得出的结果,随着微藻及高等植物分化的进行,先分化出的DGAT3内含子数量最少,WS/DGAT其次,DGAT1和DGAT2数量差距不大。

2.3 正选择位点检验

1) 位点特异模型。通过PAML软件包,采用M0(单一 ω)、M1a(近中性)、M2a(正选择)、M3(离散)、M7(β)、及M8(β & ω)6个模型分析的结果如表2所示。4种DGAT类型中只有DGAT2和WS/DGAT的M8模型中发现 $\omega > 1$ 的正选择位点,大部分位点在进化过程中受到负选择作用。其中,DGAT2中M8模型的4M及5A检测到可能存在正选择位点,经过LRT检验对M8与M7进行比



a,b,c 分别为微藻在 DGAT1、DGAT2 及 DGAT3 的分支。a, b, c are the branches of DGAT1, DGAT2, and DGAT3 of microalgae, respectively.

图 1 植物中 DGAT1、DGAT2、DGAT3 及 WS/DGAT 蛋白质序列的进化树

Fig.1 Phylogenetic relationship among plant DGAT1, DGAT2, DGAT3, WS/DGAT protein sequence

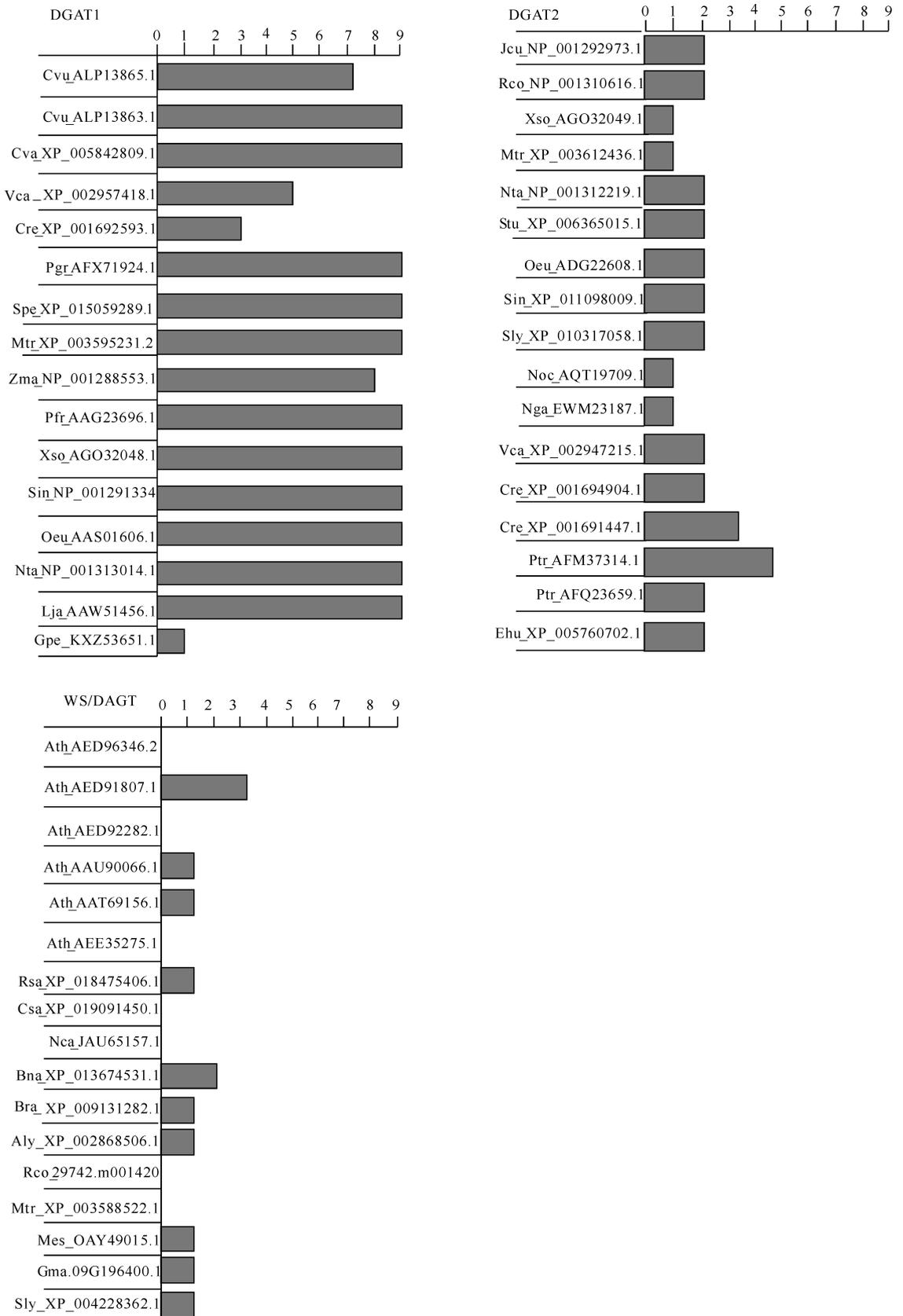


图 2 植物中 DGAT1、DGAT2、DGAT3 及 WS/DGAT 跨膜域的预测

Fig.2 Predicted transmembrane domain for DGAT1, DGAT2, DGAT3 and WS/DGAT from plants

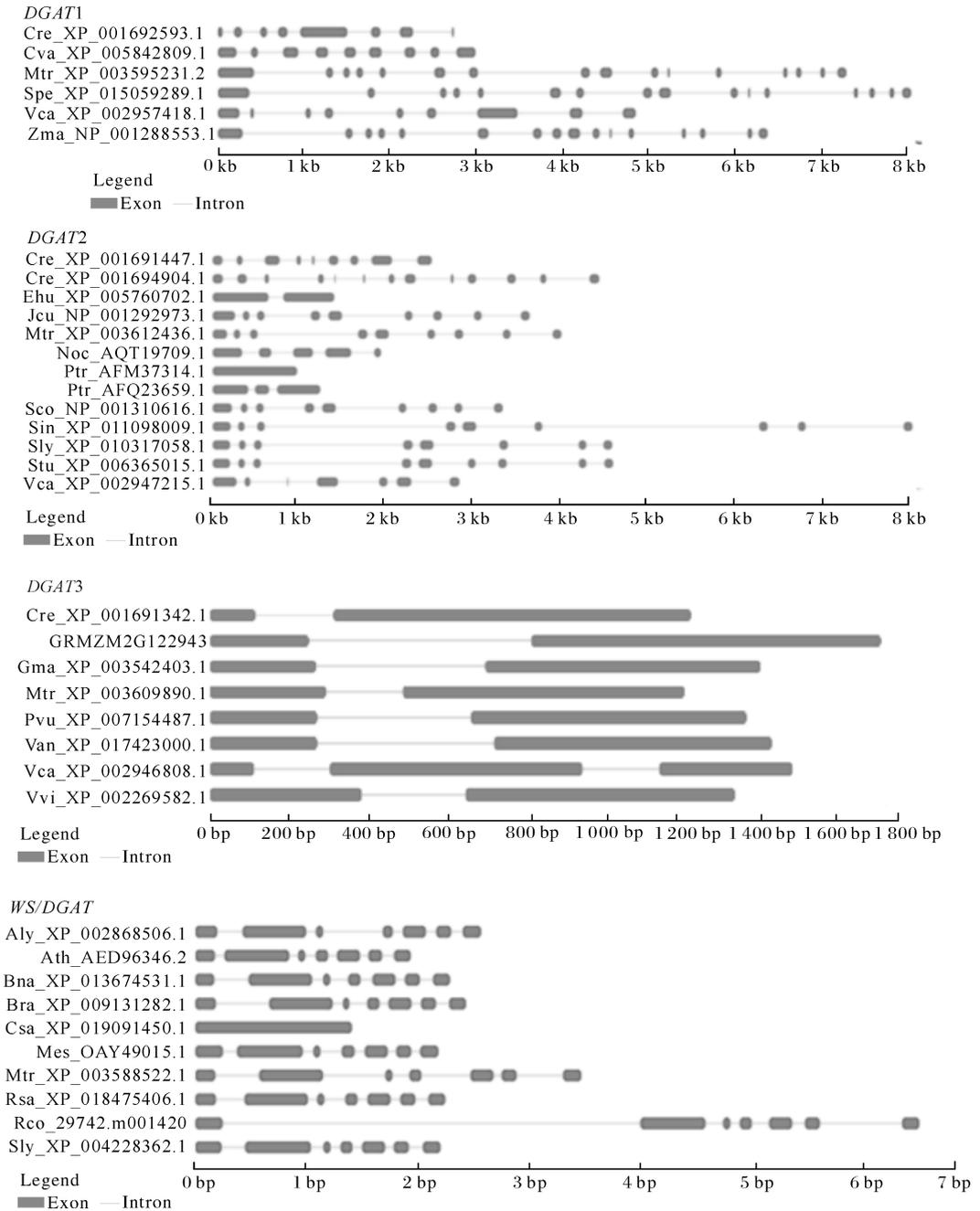


图 3 植物中 *DGAT1*、*DGAT2*、*DGAT3*、*WS/DGAT* 基因结构的分析

Fig.3 The analysis of gene structure in *DGAT1*, *DGAT2*, *DGAT3*, *WS/DGAT* from plants

较发现前者显著优于后者,故结果具有一定的可信度。同样,在 *WS/DGAT* 中的 M8 模型中,可能存在 5 个正选择位点,经过 M8 与 M7 的比较得到的结果具有显著性,同样具有较高的可信度。综合 *DGAT1*、*DGAT2*、*DGAT3* 及 *WS/DGAT* 来看,负选择对 *DGAT* 基因的进化起到了重要的作用。

2) 分支特异模型。在系统发育树(图 1)中单独标注微藻分支依次为 a、b、c,由于未找到微藻的

WS/DGAT 序列,故不对其进行分支特异模型的分析。根据分支特异模型,设置 $model=0$, $N_{site}=0$ 与 M0 模型进行比较,并通过 LRTs 显著性判断结果的可信度,结果如表 3 所示,微藻分支用方框单独标注。

由于分支模型中参数较多,所以有可能会造成 ω 值不精确。然而,该模型的优势在于,缺少序列信息时,估测 ω 值时不会受到影响。分支特异模型对

微藻中 3 种 DGAT 类型的参数估计如表 3 所示, DGAT1 中微藻分支 $\omega=0.086\ 26$, 而其他高等植物分支 $\omega=0.277\ 04$, M0 模型总体的 $\omega=0.087\ 02$ 。DGAT2 中微藻分支和高等植物分支 ω 分别为 0.005 22 和 0.165 50, M0 模型 $\omega=0.163\ 49$ 。DGAT3 中微藻分支 $\omega=0.011\ 57$, 高等植物分支 $\omega=0.245\ 40$, medel=0 时 $\omega=0.238\ 04$ 。藻类分支与高等植物分支有不同的 ω 值, 且 $\omega < 1$, 提示我们藻类与其他高等植物有不同的进化过程。通过分析

进化过程中的选择压力, 微藻与其他高等植物有不同的选择压力, 但又差别不大, 可以以此来进一步分析 DGAT 在微藻中的生物功能。

3 讨 论

本研究通过 NCBI 和 Phytozome 数据库同源性搜索获得微藻中的 3 种 DGAT 类型, 未发现微藻的 WS/DGAT 序列, 采用最大似然性算法对微藻及高等植物的 DGAT 构建进化树, 得到分支较为明显的

表 2 位点特异模型对 4 种 DGAT 基因的参数估计

Table 2 Parameter estimates for the four DGATs genes under site-specific model

| 基因类型 Type | 模型 Model | lnL | 参数估计值 Estimates of parameters | | | P 值 P value | 正向选择位点 Positive sites |
|--------------|-------------|-----------------|---|---|--|----------------|--|
| DGAT1 | M3 | -11 545.779 710 | $P_0=0.280\ 83$ $\omega_0=0.041\ 50$ | $P_1=0.603\ 92$ $\omega_1=0.178\ 69$ | $P_2=0.115\ 26$ $\omega_2=0.512\ 95$ | | □ |
| | M0 | -11 700.040 900 | | | | | Not allowed |
| | M2a | -11 650.177 670 | $P_0=0.906\ 09$ $\omega_0=0.148\ 35$ | $P_1=0.003\ 63$ $\omega_1=1.000\ 00$ | $P_2=0.090\ 28$ $\omega_2=1.000\ 00$ | | □ |
| | M1a | -11 650.177 670 | $P_0=0.906\ 09$ $\omega_0=0.148\ 35$ | $P_1=0.093\ 91$ $\omega_1=1.000\ 00$ | | | Not allowed |
| | M8 | -11 548.435 290 | $P_0=0.993\ 21$ $(P_1=0.006\ 79)$ | $P=1.143\ 15$ $\omega=3.278\ 90$ | $q=5.263\ 97$ | | 4 M** , 5 A** |
| | M7 | -11 552.954 300 | $P=1.113\ 43$ | $q=4.872\ 29$ | | 0.002645 | Not allowed |
| | M3 | -3 795.099 496 | $P_0=0.338\ 22$ $\omega_0=0.043\ 58$ | $P_1=0.539\ 47$ $\omega_1=0.346\ 82$ | $P_2=0.122\ 30$ $\omega_2=0.735\ 13$ | | □ |
| | M0 | -3 854.774 459 | $\omega_0=0.2380\ 4$ | | | | Not allowed |
| DGAT2 | M2a | -3 822.990 554 | $P_0=0.537\ 87$ $\omega_0=0.139\ 96$ | $P_1=0.045\ 84$ $\omega_1=1.000\ 00$ | $P_2=0.416\ 29$ $\omega_2=1.000\ 00$ | | □ |
| | M1a | -3 822.990 554 | $P_0=0.537\ 87$ $\omega_0=0.139\ 96$ | $P_1=0.462\ 13$ $\omega_1=1.000\ 00$ | | | Not allowed |
| | M8 | -3 796.435 199 | $P_0=0.999\ 99$ $(P_1=0.000\ 01)$ | $P=0.696\ 94$ $\omega=19.648\ 63$ | $q=1.583\ 04$ | | |
| | M7 | -3 796.434 07 | $P=0.696\ 94$ | $q=1.583\ 04$ | | | Not allowed |
| | M3 | -3 795.099 496 | $P_0=0.338\ 22$ $\omega_0=0.043\ 58$ | $P_1=0.539\ 47$ $\omega_1=0.346\ 82$ | $P_2=0.122\ 30$ $\omega_2=0.735\ 13$ | | □ |
| | M0 | -3 854.774 459 | $\omega_0=0.238\ 04$ | | | | Not allowed |
| | M2a | -3 822.990 554 | $P_0=0.537\ 87$ $\omega_0=0.139\ 96$ | $P_1=0.045\ 84$ $\omega_1=1.000\ 00$ | $P_2=0.416\ 29$ $\omega_2=1.000\ 00$ | | □ |
| | M1a | -3 822.990 554 | $P_0=0.537\ 87$ $\omega_0=0.139\ 96$ | $P_1=0.462\ 13$ $\omega_1=1.000\ 00$ | | | Not allowed |
| DGAT3 | M8 | -3 796.435 199 | $P_0=0.999\ 99$ $(P_1=0.000\ 01)$ | $P=0.696\ 94$ $\omega=19.648\ 63$ | $q=1.583\ 04$ | | |
| | M7 | -3 796.434 070 | $P=0.696\ 94$ | $q=1.583\ 04$ | | | Not allowed |
| | M3 | -13 458.582 279 | $P_0=0.524\ 78$ $\omega_0=0.097\ 75$ | $P_1=0.475\ 22$ $\omega_1=0.424\ 60$ | $P_2=0.000\ 00$ $\omega_2=71.057\ 04$ | | □ |
| | M0 | -13 619.203 977 | $\omega_0=0.228\ 32$ | | | | Not allowed |
| | M2a | -13 510.017 419 | $P_0=0.822\ 08$ $\omega_0=0.188\ 89$ | $P_1=0.161\ 15$ $\omega_1=1.000\ 00$ | $P_2=0.016\ 77$ $\omega_2=1.000\ 00$ | | □ |
| | M1a | -13 510.017 419 | $P_0=0.822\ 08$ $\omega_0=0.188\ 89$ | $P_1=0.177\ 92$ $\omega_1=1.000\ 00$ | | | Not allowed |
| | M8 | -13 443.133 331 | $P_0=0.951\ 74$ $(P_1=0.048\ 26)$ | $P=1.077\ 07$ $\omega=1.000\ 00$ | $q=3.440\ 92$ | | 2 E** , 166 E** , 197 S** , 264 T** , |
| | M7 | -13 446.609 198 | $P=0.932\ 70$ | $q=2.517\ 87$ | | 0.008 374 | 415 A** |
| WS/DGAT | | | | | | | Not allowed |

注: 在 95% 和 99% 后验概率下推测出的正选择位点分别用 * 和 ** 标出。Note : The positively selected sites were in * and ** under 95% and 99% posterior probability values, respectively .

表 3 分支特异模型对 4 种 DGAT 类型的参数估计

Table 3 Parameter estimates for the DGAT genes under branch-specific model

| 类型 Type | 模型 Model | lnL | 参数估计值 Estimates of parameters | | P 值 P value |
|------------|-------------|-----------------|----------------------------------|----------------------|----------------|
| DGAT1 | Model 2 | -7 433.800 756 | $\omega_1=0.086\ 26$ | $\omega_2=0.277\ 04$ | 0.392 559 242 |
| | Model 0 | -7 434.166 256 | $\omega=0.087\ 02$ | | |
| DGAT2 | Model 2 | -11 698.044 070 | $\omega_1=0.165\ 50$ | $\omega_2=0.005\ 22$ | 0.045 672 081 |
| | Model 0 | -11 700.040 900 | $\omega=0.163\ 49$ | | |
| DGAT3 | Model 2 | -3 852.040 818 | $\omega_1=0.245\ 40$ | $\omega_2=0.011\ 57$ | 0.019 375 736 |
| | Model 0 | -3 854.774 459 | $\omega=0.238\ 04$ | | |

注:黑框内为微藻分支数据。Note: The microalgae branch data marked by black box.

拓扑结构。小球藻的 DGAT1 单独成支,与其他高等植物的 DGAT1 有共同起源,与 Turchetto-Zolet 等^[14] 研究结果一致。团藻及莱茵衣藻的 DGAT1 与 DGAT2 亲缘关系较近。团藻、莱茵衣藻、微拟球藻及三角褐指藻的 DGAT2 亲缘关系较近且有独立的分支,与其他高等植物有共同起源,郝敬云等^[25] 通过构建系统发育树对莱茵衣藻 DGAT 蛋白进行鉴定和分析,莱茵衣藻 DGAT2 与三角褐指藻 DGAT2 亲缘关系非常近且与其他高等植物有共同起源,与本研究结果一致。团藻和莱茵衣藻的 DGAT3 同样观察到独立分支,提示我们微藻与高等植物有独立的进化路线。

在微藻的 DGAT 的蛋白质结构分析中,除了 DGAT3 中未发现跨膜域外,其他 3 种 DGAT 的蛋白质结构均呈现出与系统发育分析一致的结果,即团藻、莱茵衣藻、微拟球藻、小球藻及三角褐指藻的跨膜域数量与其他高等植物有较大差别。在基因结构分析中,微藻中不同 DGAT 类型内含子数目有较大差异,同时微藻与其他高等植物同一种 DGAT 类型中内含子数目也有一定差异。郭宁等^[26] 在对棉花与模式植物 DGAT1 基因的聚类与基因结构分析中发现,莱茵衣藻的 DGAT1 内含子数量与棉花等高等植物内含子数量有较大差距。通过系统发育树得知的最先分化出的 DGAT3 的内含子数量最少,随着分化的进行,内含子数量逐渐增多。

在利用不同模型分析 DGAT 不同位点及不同分支的选择压力中,通过位点模型观察到 DGAT2 及 WS/DGAT 均在 M8 模型下发现较少的正选择位点,且通过 LRT 检验发现结果可信度较高,提示我们 4 种 DGAT 类型在进化过程中受到较强的负选择作用。通过分支模型发现藻类与其他高等植物的 ω 值不同,说明微藻和高等植物有不同的进化过程。

通过环境选择等因素造就了微藻与其他高等植物不同的 DGAT 结构和进化,而 DGAT 又是合成

三酰基甘油的最后一步,这为后续微藻产油脂基因工程改造提供了理论依据,为陶高微藻油脂提供了可能的策略。

参 考 文 献

- [1] 翟映雪,梁成伟,张伟,等.微藻油脂的研究进展[J].食品工业科技,2014,35(17):370-374.
- [2] SHARMA K K, SCHUHMANN H, SCHENK P M, et al. High lipid induction in microalgae for biodiesel production[J]. Energies, 2012, 5(5): 1532-1553.
- [3] WLJFFELS R H, BARBOSA M J. An outlook on microalgal biofuels[J]. Science, 2010, 329 (5993): 796-799.
- [4] CHISTI Y. Biodiesel from microalgae[J]. Biotechnology advances, 2007, 25(3): 294-306.
- [5] TRAN N H, BARTLETT J R, KANNANGARA G S K, et al. Catalytic upgrading of biorefinery oil from microalgae[J]. Fuel, 2010, 89(2): 265-274.
- [6] BROWN A C, KNIGHTS B A, CONWAY E. Hydrocarbon content and its relationship to physiological state in green algae *Botryococcus braunii*[J]. Phytochemistry, 1969, 8(3): 543-547.
- [7] XIAO Y, ZHANG J, CUI J, et al. Metabolic profiles of *Nannochloropsis oceanic* IMET1 under nitrogen-deficiency stress[J]. Bioresource Technol, 2013, 130: 731-738.
- [8] MOHEIMANI N, RBOROWITZKA M A. The long-term culture of the coccolithophore *Pleurochrysis carterae* (Haptophyta) in outdoor raceway ponds[J]. Journal of applied phycology, 2006, 18(6): 703-712.
- [9] CHIU S Y, KAO C Y, TSAI M T, et al. Lipid accumulation and CO₂ utilization of *Nannochloropsis oculata* in response to CO₂ aeration[J]. Bioresource technology, 2009, 100: 833-838.
- [10] WANG D M, NING K, LI J, et al. *Nannochloropsis* genomes reveal evolution of microalgal oleaginous traits[J]. PLoS genetics, 2014, 10(1): 2-13.
- [11] 张晓舟. 莱茵衣藻高产油脂藻株的诱变筛选及其油脂代谢关键酶基因的转录分析[D]. 青岛: 中国海洋大学, 2013.
- [12] LOCK Y Y, SNYDER C L, ZHU W, et al. Antisense suppression of type 1 diacylglycerol acyltransferase adversely affects plant development in *Brassica napus* [J]. Physiol Plant, 2009, 137(1): 61-71.
- [13] LARDIZABAL K, EFFERTZ R, LEVERING C, et al. Express-

- sion of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean[J]. *Plant Physiol*, 148(1): 89-96.
- [14] TURCHETTO-ZOLET A C, MARASCHIN F S, DE MORAIS G L, et al. Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis[J]. *BioMed central academic evolutionary biology*, 2011, 11: <http://www.biomedcentral.com/147r2148/11/263>.
- [15] FEI X, LI P, LI X, et al. Low-temperature-and phosphate deficiency-responsive elements control DGAT3 expression in *Chlamydomonas reinhardtii*[J]. *Journal of eukaryotic microbiology*, 2017, 13: 214-238.
- [16] WEI H, SHI Y, MA X, et al. A type-I diacylglycerol acyltransferase modulates triacylglycerol biosynthesis and fatty acid composition in the oleaginous microalga, *Nannochloropsis oceanica*[J]. *Biotechnol Biofuels*, 2017, 5(10): 174-184.
- [17] KUMAR S, STECHER G, TAMURA K. Molecular evolutionary genetics analysis (MEGA) Version 7.0 for bigger datasets [J]. *Molecular biology and evolution*, 2016, 33(7): 1870-1874.
- [18] HU B, JIN J P, GUO A Y, et al. GSDS 2.0: an upgraded gene feature visualization server[J]. *Bioinformatics*, 2015, 31(8): 1296-1297.
- [19] LETUNIC I, COPLEY R R, SCHMIDT S, et al. SMART 4.0: towards genomic data integration[J]. *Nucleic Acids Res*, 2004, 32(1): 142-144.
- [20] YANG Z. PAML: a program package for phylogenetic analysis by maximum likelihood[J]. *Computer applications in the biosciences*, 1997, 13: 555-556.
- [21] 陈洁, 张丽君, 王艇. 蕨类植物 rbcL 基因正选择和负选择位点的鉴定[J]. *西北植物学报*, 2009, 29(12): 2391-2400.
- [22] YANG Z, BIELAWSKI J P. Statistic method for detecting molecular adaptation[J]. *Trends ecology evolution*, 2000, 15(12): 496-503.
- [23] 梁成伟, 程江峰, 苏忠亮, 等. 八氢番茄红素脱氢酶(PDS)基因分子进化特征[J]. *青岛科技大学学报(自然科学报)*, 2009, 30(5): 408-411.
- [24] 李丽莎, 李祥龙, 周荣艳, 等. 酪氨酸相关蛋白酶 1 基因编码区系统发育分析和正选择位点检测[J]. *中国畜牧兽医*, 2015, 42(9): 2427-2435.
- [25] 郝敬云, 周广航, 邵雪梅, 等. 莱茵衣藻 DGAT2 基因家族的鉴定与功能分析[J]. *分子植物育种*, 2016, 14(9): 2343-2352.
- [26] 郭宁, 胡利宗, 解恒昌, 等. 棉花与模式植物 DGAT1 基因的鉴定与比较分析[J]. *广东农业科学*, 2015, 15: 98-104.

Structure and evolution of diacylglycerol acyltransferase (DGAT) in microalgae

YANG Xiao LIANG Chengwei

College of Marine Science and Biological Engineering, Qingdao University of Science and Technology, Qingdao 266042, China

Abstract Microalgae is the third generation of potential bio-energy resources, and the presence of grease in microalgae is mainly triacylglycerol. Diacylglycerol acyltransferase (DGAT) is a key rate-limiting enzyme in the control of TAG synthesis. DGAT is divided into DGAT1, DGAT2, DGAT3 and WS/DGAT. The DGAT of microalgae and higher plants was obtained from the database. The diversity and evolution of the four DGAT types and gene and protein structures was analyzed with phylogenetic methods. In order to further analyze the evolution of DGAT in microalgae, the positive selection sites were selected to estimate the selection pressure in the evolutionary process. Result showed that the genetic structure and protein structure of different DGAT types in microalgae were different. In the same DGAT type, the gene structure and protein structure of microalgae was different from that of other higher plants. It was found that microalgae did have different evolutionary processes with other higher plants. There was a positive selection sites of DGAT2 and WS/DGAT. The positive selection sites of DGAT1 and DGAT3 were not found. It will help us to decipher its functional role in plants and promote the research of renewable energy.

Keywords microalgae; diacylglycerol acyltransferase; gene structure; positive selection site; phylogeny