

# 皮状丝孢酵母脂肪酸 CoA 连接酶 (Facl) 基因的克隆与分析

尚常花 王忠铭 袁振宏 朱顺妮 秦磊 吕鹏梅  
(中国科学院可再生能源重点实验室 中国科学院广州能源研究所, 广州 510640)

**摘要** 从皮状丝孢酵母(*Trichosporon cutaneum*)中克隆了 *Facl* 基因及其 5' - 上游序列, 并对 *Facl* 基因及其 5' - 上游序列进行了分析。根据前期的基因组文库测序结果, 获得了 898 bp DNA 序列。以此为基础, 2 次使用染色体步移(Genome walking)方法, 分别获得了 854 bp 和 1 236 bp 5'-DNA 序列。根据获得的 5 和 3 端序列, 合成引物用于扩增 *Facl* 基因的全长 DNA 和 cDNA 序列。分析发现, 皮状丝孢酵母 *Facl* 基因 cDNA 包含一个 1 662 bp 的开放读码框, DNA 不含内含子序列。同时获得了 482 bp 的 5' - 上游序列。在 5' - 上游序列区域存在一系列预测的顺式作用元件。相似性分析的结果表明 *T. cutaneum* 和 *Rhodococcus erythropolis* 的 *Facl* 氨基酸序列相似性最高。该研究旨在为后继的 *Facl* 基因的功能和表达的研究、油脂代谢的遗传改造奠定基础。

**关键词** 皮状丝孢酵母 *Facl* 基因 5' - 上游序列 序列分析

**中图分类号**: Q789 **文献标识码**: A **文章编号**: 1003 - 0174(2015)01 - 0071 - 05

随着化石能源即将耗竭以及生态环境日益恶化, 可再生能源受到全世界的青睐。作为可再生能源的一员, 生物柴油因其具备可生物降解、清洁、无毒和可再生等优点, 引起人们的广泛关注<sup>[1]</sup>, 同时, 与传统的石化柴油相比较, 生物柴油具有相似的化学特性<sup>[2]</sup>。目前, 生物柴油主要是利用植物油制备, 然而大规模发展植物来源的生物柴油将需要占用大量土地以及导致食用油短缺。在这种情况下, 微生物油脂成为生物柴油的可持续原料。

产油微生物因具备可产生大量油脂、生长迅速及所产生油脂的脂肪酸组成与植物油相似等特点, 被认为是一种可选择的生物柴油原料<sup>[3-4]</sup>。已经发现许多微生物可以积累油脂, 包括细菌、酵母、微藻和丝状真菌。

皮状丝孢酵母(*Trichosporon cutaneum*)是一种可以同时利用葡萄糖和木糖的产油酵母<sup>[5]</sup>。*T. cutaneum* 不仅可以利用葡萄糖和木糖, 还可以利用各种生物物质的水解液, 包括玉米芯的酸水解液<sup>[6-7]</sup>和大米草的水解液<sup>[8]</sup>等, 因此, *T. cutaneum* 用于制备生物柴油具备广泛的应用前景。但是, 关于 *T. cutaneum* 油脂合成途径调控的研究还非常少, 限制了利用基因工

程技术对 *T. cutaneum* 产油性能进行改造。

脂肪酸 CoA 连接酶(*Facl*)可以催化脂肪酸、ATP 和 CoA 合成脂酰 CoA。它通过将游离脂肪酸活化为它们的 CoA 硫酯, 在不同的代谢和调控过程中发挥着重要作用, 包括油脂生物合成、脂肪酸降解、中间代谢、基因表达、三酰甘油掺入、信号转导和基因表达调控<sup>[9-10]</sup>。*Facl* 催化的反应通过两步进行。首先, 脂肪酸和 ATP 反应生成脂酰 - AMP 和焦磷酸。其次, 脂酰 - AMP 和 CoA 反应, 生成脂酰 CoA 和 AMP。根据碳链的长度, 脂肪酸可以被划分为 3 种类型: 短链(C1 - C4), 中链(C5 - C16)和长链(大于 C16)。与 *Facl* 在细胞代谢中的多种功能相对应, 许多微生物含有可利用不同长度碳链脂肪酸的 *Facl* 同工酶<sup>[11-12]</sup>。

鉴于 *Facl* 在油脂代谢调控中的重要功能, 本研究从 *T. cutaneum* 中克隆了 *Facl* 基因并对其进行了分析。该工作为以后 *Facl* 基因的功能分析、超量表达研究、抑制表达研究及对 *T. cutaneum* 油脂代谢的遗传调控奠定了基础。

## 1 材料与方法

### 1.1 材料

#### 1.1.1 菌株

皮状丝孢酵母 2.571 购买自中国普通微生物菌

基金项目: 国家科技支撑计划(2011BAD14B03), 国家自然科学基金(31100189)

收稿日期: 2013 - 09 - 27

作者简介: 尚常花, 男, 1980 年出生, 助理研究员, 微生物学

通讯作者: 袁振宏, 男, 1953 年出生, 研究员, 生物物质技术

种保藏管理中心。大肠杆菌 (*E. coli*) DH5 $\alpha$  由本实验室保存。

### 1.1.2 主要试剂和仪器

Taq 酶、pMD20-T 载体、T<sub>4</sub> DNA Ligase、Genome walking kit、琼脂糖凝胶回收试剂盒、RNA 提取试剂 RNAiso Plus reagent、反转录酶 PrimeScript Reverse Transcriptase; 大连宝生物公司; PCR 仪 Mycycler; Bio-Rad 公司; 离心机 Centrifuge 5417R; Eppendorf 公司; 凝胶成像系统 Tanon-1600; 上海天能科技有限公司; 电泳系统 DYY-7C; 北京市六一仪器厂。

### 1.1.3 引物序列

引物由生工生物工程(上海)股份有限公司合成。引物根据 Genome walking kit 试剂盒的使用说明设计。引物序列见表 1。

表 1 皮状丝孢酵母 *Facl* 基因扩增使用的引物

引物	序列(5'-3')
Facl(5sp1)	GCAGCCACTCGCCGAAGGGAT
Facl(5sp2)	CGACTGCGATGGCGGTGCTGC
Facl-25sp1	ACCTGCTCCTTGTAGTCGTTCT
Facl-25sp2	TGAGTAGATAACGCAGTGGGCT
Facl-N	ATGACTGCCCGCAGATCCG
Facl-C	TCATACGGTGATCTCGGCCCG
Facl(qs)	CGGATTGATGCCTACGATCCT
Facl(qa)	AGTCATGGGATGGCCTTTCCT

### 1.1.4 生物信息学分析

同源基因通过 NCBI Blast 搜索被发现 (<http://www.ncbi.nlm.nih.gov/BLAST/>)。潜在的 motif 通过 ScanProsite Results Viewer 工具被发现 (<http://cn.expasy.org/prosite/>)。等电点 pI 和分子质量通过 Compute pI/Mw 工具被计算出 ([http://www.expasy.ch/tools/pi\\_tool.html](http://www.expasy.ch/tools/pi_tool.html))。序列比对使用 Clustal W 软件。进化树的构建利用 MEGA 4 软件。功能域的搜索利用 NCBI 网站的 NCBI CD Search 在线工具。

## 1.2 方法

### 1.2.1 总 DNA 和总 RNA 的提取

12 000 r/min 离心 2 min 收集对数生长期的酵母细胞,然后按照 CTAB 法提取基因组 DNA<sup>[13]</sup>。处于对数生长期的酵母细胞经离心后,去除上清,向细胞沉淀中加入 RNAiso Plus reagent,按照试剂盒的说明书提取总 RNA。取 1  $\mu$ g RNA 使用 oligo(dT)<sub>17</sub> 引物和反转录酶 PrimeScript Reverse Transcriptase 进行反转录反应生成 cDNA。

### 1.2.2 *Facl* 基因的 5' Genome walking

本实验室在前期的 *T. cutaneum* 基因组文库测序中获得了 898 bp 的 *Facl* DNA 片段。根据已

获得的 *T. cutaneum Facl* 基因的部分 DNA 序列,设计用于 5' Genome walking 试验的引物 Facl(5sp1) 和 Facl(5sp2),按试剂盒的说明书进行 PCR,以获取基因的 5'-端 DNA 序列。由于第 1 次 5' Genome walking 没有发现起始密码子。因此根据第 1 次 5' Genome walking 的结果,设计出用于第 2 次 5' Genome walking 的引物 Facl-25sp1 和 Facl-25sp2。PCR 产物回收纯化、连接、转化与重组子检测均按常规的分子生物学试验方法进行<sup>[14-15]</sup>。阳性重组子送上海英骏生物技术公司测序。

### 1.2.3 *Facl* 基因全长及 5'-上游序列的获得和分析

将得到的 *T. cutaneum Facl* 基因的部分 DNA 序列与 2 次 5' Genome walking 扩增获得的 DNA 序列进行拼接。根据拼接序列,设计出扩增 *Facl* 基因 DNA 和 cDNA 全长的引物 (*Facl-N* 和 *Facl-C*) 以及扩增 5'-上游序列的引物 *Facl(qs)* 和 *Facl(qa)*,来验证拼接序列的正确性。使用 NCBI Blastx 软件进行 *Facl* 基因的同源性比对。使用 PlantCARE 软件对 5'-上游序列中潜在的顺式作用元件进行分析。

## 2 结果与讨论

### 2.1 *Facl* 部分 DNA 序列的获得

前期的 *T. cutaneum* 基因组文库测序获得了 898 bp DNA。利用 Blast x 软件分析获得的序列。结果显示;*T. cutaneum* 的 898 bp DNA 序列翻译为氨基酸序列后与 *Rhodococcus erythropolis* 的 *Facl* 氨基酸序列相似性最高,为 99%。与 *Rhodococcus opacus*、*Rhodococcus imtechensis* 和 *Rhodococcus sp. DK17* 的 *Facl* 氨基酸序列相似性分别为 87%、87% 和 86%。上述结果表明所获得的部分序列是 *Facl* 基因的序列。

### 2.2 *Facl* 基因的 5' Genome walking 结果与分析

根据已获得的 *Facl* 基因的部分序列,用 5' Genome walking 的方法扩增得到了 *Facl* 的 5' 端 DNA。第 1 次使用 *Facl(5sp1)* 和 *Facl(5sp2)* 进行的 5' Genome walking 获得了 854 bp 序列,该序列与 898 bp 核心序列重叠,表明获得了正确的 DNA 5' 端序列。但是,经过 Blastx 比对,发现得到的 DNA 5' 端序列不够完整,尚未扩增到起始密码子区域。因此,根据获得的 854 bp 序列设计引物 *Facl-25sp1* 和 *Facl-25sp2* 进行了第 2 次 5' Genome walking,获得了 1236 bp DNA 序列,该序列与第 1 次 5' Genome walking 的结果重叠,表明获得了正确的第 2 次 5' Genome walk-

ing 结果。分析获得的 DNA 5'端序列时,发现了位于起始密码子上游长度为 482 bp 的 5' - 上游序列。

### 2.3 Facl 基因全长序列和 5' - 上游序列的获得

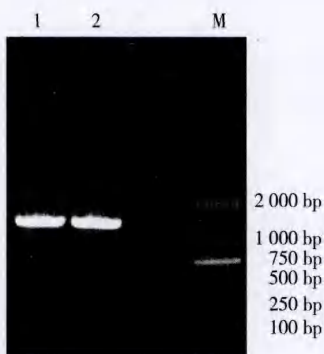
利用设计的 Facl 基因的 DNA 全长扩增引物以及 5' - 上游序列的扩增引物进行验证,表明 5' - 上游序列及拼接的序列均正确。T. cutaneum Facl 开放读码框 (ORF) 为 1662 bp (图 1 和图 2)。使用 BioEdit 软件对全长 DNA 和 cDNA 序列进行比较,发现 T. cutaneum Facl 基因不含内含子。

```

1 ATGACTGCCCCCGAGATCCGAGTTCCTGACGACAGGGCGGTCTCACGCTCACCC
GTCCAGGTGGGCCATTTCGAGATCGTCCGAAGAAGAGGTTCTCGGCACCGCTATGCCCGT
GATGAAGAACCCTGGTCAAGCCATCGGTGAACATCTCGGCATCTCGCGCTGGGGG
GACGGCGAATCTCGTCCACAGCGATCTGGAATCTCGTCCACCGAGCAGCCGATT
CCGTGCGCGCGTGGCACTGCTGCGGGGCAAGTACGGAGTCCGCAAGGGCGACC
GTGTGGGCATCTCGTCCCAACACTCCCGAATGGGTATGACGTTCTTCGCGACGCA
GGCGCTCGGCGGATCGTTCGCTCAACGGTGTGGTGGTTCGCGCGAGGTCACC
CACGGATCGAAGTCTCACGTCGCAAGATTCTGGTGTGGACGCGCAAGCGGCGCC
GCCCTCCCGAATGCGCGGAGACATACGCTCTCCATCGAGGAGGATCTGCCCCG
CGCTCATCGCGAATATGACAGGCTCCGACGTGCGGACCGACATCGACGAGGACGA
TCCCCCCGTATCTCGTTCACCAAGTGTACCAAGCGGCGCCCAAGGGCGCACTCCAC
TCGCACCGCAACGTGATGGCGTGTCTGACTACCAACAAATGACAGCGATCGGG
CAGCATTACCGGCGGTGAGTACGACGTTCCAAAGCGAGCCACTGCGTTATCTACT
CACCTCCCGTCTTCCACATCGCCAGCCTGCACAATCTGGTGGTCCCGCTCTCGGA
CGGGCAGTGGCGTATCTGACGAGGGCGGCTTCGACGTGCAAGGTGCTCGGGT
TGCTCGAACCGGAGCGCTCACCAACTGGGGTGTGTCGCCACGATGGCATCGCGGT
ACTCGAACCGGCAACATCGCCAAGTACGACGTTCTTCGTTGACTGATTCTCGGTG
GCATCCGCGCGTCTGTCGCCCGCTTCAAGGAACGACTCAAGGAGCAGGGTCCCGTTC
GCACGCAATGCTCTGGTGCACAGCTACGGTCTCACCGAATGACAGCACCGCCATCGCAG
TCGCCGTGCGGCGCAACTCGAAGAGTTCAGGCACTCTCGGACGTCCGATCATCAC
CGTCTCGATGGAGATTTCGCAATCCCTCGCGGAGTGGCTGCCGACCGCATCGAAGGT
GAAGTGTGCGTACGACGTCGTTGCTGATCTCGGATATGGGAGAACGACGAGGCCA
CCGAGTCTGCAATCGACCCGACGCTGGCTCCGACCGCGGATTCGCGGTCATCGA
GAACGGTCTCTCCGCTCACGCGGCGTCTTCCGACCTGATCTGCGTGGCGCGGAG
AACGTCTACCCACCAGATCGAGCAGCGCTCGACGAACACCCCGACGTGATCGAG
TGCCTGTATCGGATGCGCGACCCGACCTCGGCCAGGAAAGTGTCCCGCTGGTGG
TCTGCGTCCGAGGAGGACACCAAGGAGGCGAGTTGCGGCACTACGCGCGGAGC
GCTCTGCTACTCAAGGTGCCGTGAAGTGGCGATACCAATGATCTGTTGCTCCG
1621 AACCGAACCGGCAAGATGGTTCGCGCGGAGATCACCGTATGA
    
```

注:下划线表示起始密码子 ATG 和终止密码子 TGA。

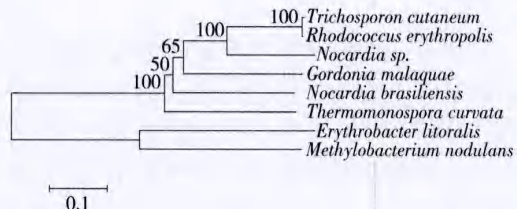
图 1 Facl 基因的全长 cDNA 序列



注:M: DL 2 000 分子质量标准; 1: cDNA 全长; 2: DNA 全长。

图 2 皮状丝孢酵母 Facl 基因的全长 cDNA 和 DNA

将 T. cutaneum Facl 氨基酸序列进行 Blast p 同源性比对后,发现与已知 Facl 氨基酸序列具有广泛的同源性。利用 Clustal W 和 MEGA 4 软件,构建了 Facl 的系统发育树 (图 3)。T. cutaneum Facl 与 Rhodococcus erythropolis Facl 表现出最高的相似性,这符合传统的分类系统。



注:使用 Clustal W 和 Mega 4 软件构建了系统进化树,节点上的数字表示分支可信度的百分比。

图 3 采用 Neighbor - Joining 方式

根据 Facl 氨基酸序列构建的系统进化树

### 2.4 Facl 全序列的分析

使用 NCBI CD - Search 进行分析,发现 T. cutaneum Facl 存在保守的功能域 FACL\_like\_2、LC - FACS\_euk 和 CaiC,这些功能域与合成脂酰 CoA 相关。在预测的 T. cutaneum Facl 蛋白序列中发现了保守的 ATP 结合 motif。ScanProsite Results Viewer 分析结果表明从 197 位到 208 位的序列 ILFTSGTSGrPK 是预测的 ATP 结合 motif,这是 Facl 家族的典型特征。T. cutaneum Facl 蛋白预测的等电点 pI 为 4.96,分子质量为 59.988 ku。

### 2.5 Facl 5' - 上游序列的分析

使用引物 Facl(qs) 和 Facl(qa),以 T. cutaneum 基因组 DNA 为模板,进行 PCR 扩增,获得了 482 bp 5' - 上游序列。该序列包含典型的启动子元件 CAAT box (图 4)。此外,还发现了与光照、激素和环境胁迫响应相关的调控元件,包括 CGTCA - motif (参与茉莉酸甲酯响应)、TCA - element (参与水杨酸响应)、MBS (参与干旱胁迫应答) 和光照响应元件 (C - box、CATT - motif、G - box、GTGGC - motif)。这些潜在的顺式作用元件为今后进一步研究环境条件对 Facl 表达的影响提供了线索。

```

-482 CGGATTGATGCTACGATCGTGGCTCCGGTCAAGGGCCGCGCCCGGAGCCA
MBS
GAACAGGAACTCGGGATCGTTCGAGAAGGACTCCGATGTAAACGGTCT
TCA-element
CAGGCGCGGAAGCGACTTCGCAAGCGAGGCCACGCGCTGCGCTCTCGCG
GTGGC-motif
ACCATCTCGTCCAGCTCCAATCTGCGTCCCGTCCGACAGCCAAAGTGC
TGATCACCGAGACGGTTCATCAACATCTGCGGACGTCCTCGCGCTGAGG
AGCTGCAAGAGTCCGATTCGAAAGTGTTCGGTGGCGGTTCTGCTATCGT
CATT-motif CGTCA-motif
GTCCTGTGTCCTAGTTCGATATCGGAATGTCGATTACTCCAATGCGATACGA
CTCGAACGTATCGACGCTCAAGCGGAAACCGTCTCAGTCTCGCTCAGTG
GTAAACCGCGCCACAGATATGATTTCTCGTGAGATTTGTGACGTGAAAA
G-box
-25 GCGCCACAATACGAAGGCCATCCCATG
CAAT-box +1
    
```

注:翻译起始位点 ATG,其位置定义为 +1。预测的顺式作用元件用下划线标记,名字写在对应的元件下面

图 4 在皮状丝孢酵母 Facl 基因的 5' - 上游序列区域预测的顺式作用元件

### 3 结论

对从 *T. cutaneum* 中克隆到的基因序列和其编码的氨基酸序列进行了分析,并以 GenBank 中部分细菌及真菌的脂肪酸 CoA 连接酶氨基酸序列为基础进行了系统进化分析。结果表明扩增所得的序列为编码脂肪酸 CoA 连接酶的基因,其大小为 1 662 bp。

#### 参考文献

- [1] Adamczak M, Bornscheuer U T, Bednarski W. The application of biotechnological methods for the synthesis of biodiesel [J]. *European Journal of Lipid Science and Technology*, 2009, 111(8): 808 - 813
- [2] Raetz K D. Challenges and advances in making microalgae biomass a cost efficient source of biodiesel [J]. *MMG 445 Basic Biotechnology*, 2009, 5: 37 - 43
- [3] Zhu L Y, Zong M H, Wu H. Efficient lipid production with *Trichosporon fermentans* and its use for biodiesel preparation [J]. *Bioresource Technology*, 2008, 99(16): 7881 - 7885
- [4] Papanikolaou S, Aggelis G. Lipids of oleaginous yeasts. Part II: Technology and potential applications [J]. *European Journal of Lipid Science and Technology*, 2011, 113(8): 1052 - 1073
- [5] Hu C, Wu S, Wang Q, et al. Simultaneous utilization of glucose and xylose for lipid production by *Trichosporon cutaneum* [J]. *Biotechnology for Biofuels*, 2011, 4(1): 25
- [6] Chen X F, Huang C, Xiong L, et al. Microbial oil production from corncob acid hydrolysate by *Trichosporon cutaneum* [J]. *Biotechnology Letters*, 2012, 34(6): 1025 - 1028
- [7] Huang X, Wang Y, Liu W, et al. Biological removal of inhibitors leads to the improved lipid production in the lipid fermentation of corn stover hydrolysate by *Trichosporon cutaneum* [J]. *Bioresource Technology*, 2011, 102(20): 9705 - 9709
- [8] Shen J J, Li F C, Yang Q L, et al. Fermentation of *Spartina anglica* acid hydrolysate by *Trichosporon cutaneum* for microbial lipid production [J]. *Marine Sciences*, 2007, 3(8): 38 - 41
- [9] Black P N, DiRusso C C. Transmembrane movement of exogenous long - chain fatty acids: proteins, enzymes, and vectorial esterification [J]. *Microbiology and Molecular Biology Reviews*, 2003, 67(3): 454 - 472
- [10] Faergeman N J, Knudsen J. Role of long - chain fatty acyl - CoA esters in the regulation of metabolism and in cell signaling [J]. *Biochemical Journal*, 1997, 323(1): 1 - 12
- [11] Faergeman N J, DiRusso C C, Elberger A, et al. Disruption of the *Saccharomyces cerevisiae* homologue to the murine fatty acid transport protein impairs uptake and growth on long - chain fatty acids [J]. *Journal of Biological Chemistry*, 1997, 272(13): 8531 - 8538
- [12] Fujino T, Man - Jong K, Minekura H, et al. Alternative translation initiation generates acyl - CoA synthetase 3 isoforms with heterogeneous amino termini [J]. *Journal of Biochemistry*, 1997, 122(1): 212 - 216
- [13] Saghai - Maroof M A, Soliman K M, Jorgensen R A, et al. Ribosomal DNA spacer - length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1984, 81(24): 8014 - 8018
- [14] 刘强, 刘玉珍, 刘东杰, 等. 利用大肠杆菌工程菌株生物转化脂肪酸的初步研究 [J]. *中国粮油学报*, 2013, 28(7): 41 - 47
- [15] 脂肪酸 *Bacillus pumilus* 羧基化调控基因 *fadD* - like 的克隆与序列分析 [J]. *中国粮油学报*, 2010, 25(2): 66 - 70.

## Cloning and Characterization of a Gene Encoding Fatty Acid CoA Ligase ( *Facl* ) from *Trichosporon Cutaneum*

Shang Changhua Wang Zhongming Yuan Zhenhong

Zhu Shunni Qin Lei Lü Pengmei

(CAS Key Laboratory of Renewable Energy, Guangzhou Institute of Energy Conversion,  
Chinese Academy of Sciences, Guangzhou 510640)

**Abstract** The *Facl* gene and its 5'-flanking sequence has cloned from *Trichosporon cutaneum* and analysis.



Based on the former result of genome library sequencing, 898 bp DNA sequence was obtained. Based on the 898 bp DNA sequence, the 5' genomic DNA was cloned by genome walking method; 854 bp and 1 236 bp DNA sequences were amplified respectively in two genome walking experiments. On the basis of the 5' and 3' termini sequences, primers were synthesized to amplify the full-length genomic DNA and cDNA sequence. The full-length *Facl* cDNA contained 1 662 bp open-reading frame (ORF), while the full-length DNA did't contain intron sequence. In addition, 482 bp 5'-flanking sequence was obtained. There were a number of predicted cis-acting elements in 5'-flanking sequence. The results of similarity analysis revealed that the highest identity of *Facl* amino acid sequence could be found between *T. cutaneum* and *Rhodococcus erythropolis*. The study has laid foundation for further research on the function and overexpression of *Facl* gene, genetic manipulation of lipid metabolism as a success.

**Key words** *Trichosporon cutaneum*, *Facl* gene, 5'-flanking sequence, sequences analysis

---

(上接第70页)

90% in tea seed oil. 28 varieties of tea seed oil utilized in the paper contain 17 kinds of fatty acids, 11 kinds of which can be measured in tea seed oil. Principal component analysis showed that stearic acid, iso-oleic acid, seventeen carbonic acid and palmitic acid are typical fatty acids. Different extraction processes have no significant differences on the relative contents of more than 1% fatty acids, such as palmitic acid, oleic acid, iso-oleic acid and linoleic acid. However, it mainly impacts on the other trace fatty acids, thus affects the oil quality.

**Key words** tea seed oil, fatty acid, oil extraction process, tea tree cultivar