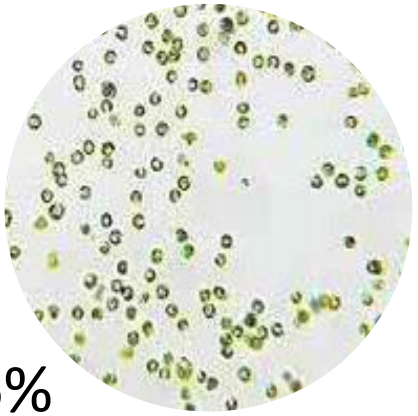
微拟球藻中的限速酶DGAT的探究
Bioinformatic Analysis of DGAT in *Nannochloropsis oceanica*

成员：G04 – 赵宇轩 何以琳 吴姝芳
G05 – 李可佳 张青 张昱
G06 – 何川 谭洁 张威 郑国旋
时间：2018.06.23

研究背景

研究背景

微拟球藻（Nannochloropsis）属于单细胞光合微藻中的眼点藻纲，细胞直径2–5 μm，生长快，可在盐碱、荒地等非耕地使用海水进行培养。

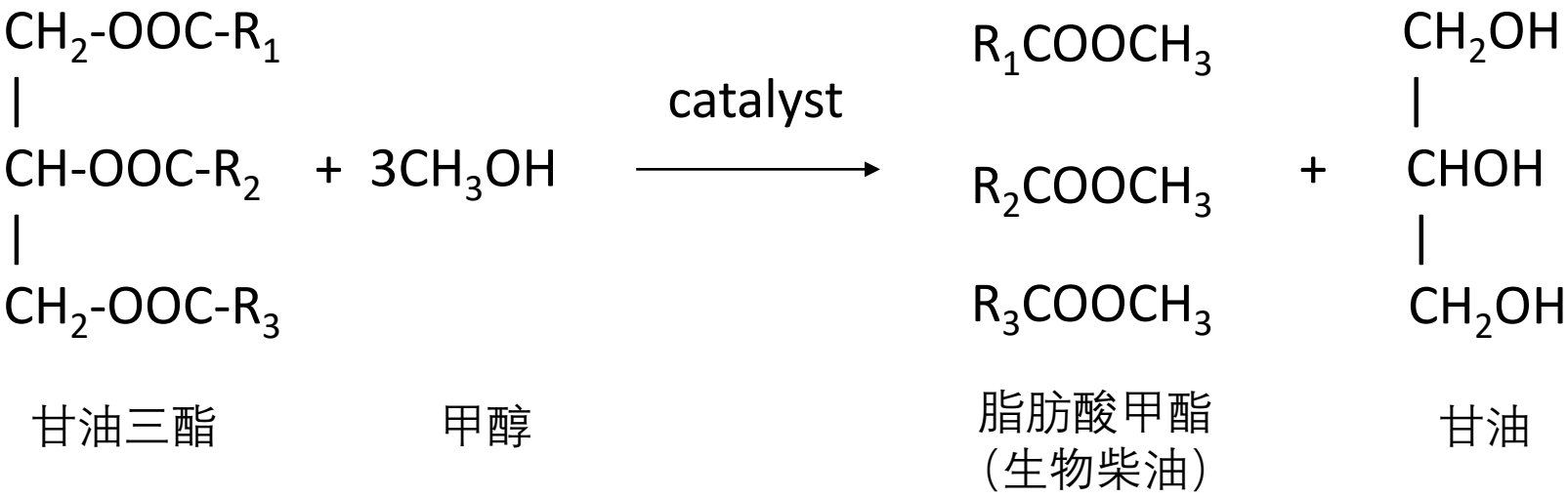


优势：

1.生长快，易培养
2.甘油三脂(TAG)产量在 10%-36%
3.细胞壁较薄易破碎
4. 20世纪80年代就已有较好的规模化生产方法和技术基础
5. 有较为成熟的分子工程平台

研究背景

生产生物柴油的转酯化反应



研究背景

DGAT

DGAT是Kennedy途径合成TAG的唯一限速酶，能催化DAG和acyl-CoA，将acyl-CoA的酰基基团转移到DAG的sn-3位置合成TAG。

Kennedy 途径

DGAT被认为在TAG合成及积累方面具有重要的作用，相关研究表明DAGT与植物种子发育、萌发，动物饮食性肥胖有关。

研究背景

DGAT分类及特点

	分布	类型	跨膜结构域	分子量 kDa	等电点	氨基酸数	基因定位	cDNA长度 bp	外显子/内含子数
DGAT1	动物、植物	跨膜蛋白	9-10	约59	8.71-9.63	约500	染色体Ⅱ（拟南芥）	1998	16/15（拟南芥）
DGAT2	动物、植物、酵母	跨膜蛋白	2-3	约39	8.86-9.70	约340	染色体Ⅲ（拟南芥）	1330	8/7（拟南芥）
DGAT3	植物（花生）	胞质溶解蛋白	0	约41	8.17-9.47	约350	*	1637	*

Cases,1998; Cao,2011; Saha,2006; Hobbs,1999; Bouvier,2000

DGAT基因及氨基酸序列

- 微拟球藻*Nannochloropsis oceanica*

  基因序列(CDs) (Genbank:KY073295.1)

  氨基酸序列(Genbank:ASL69957.1)

- 衣藻*Chlamydomonas reinhardtii*

  基因序列(CDs) NCBI Reference Sequence: XM_001693137.1
  Cre01.g045903.t1.1 CDS (phytozome数据库)

  氨基酸序列 NCBI Reference Sequence: XP_001693189.1
  Cre01.g045903.t1.1 (phytozome数据库)

- 小球藻*Chromochloris zofingiensis*

  基因序列(CDs) Cz03g14080.t1 CDS (phytozome数据库)
  Cz06g04190.t1 CDS (phytozome数据库)
  Cz09g23020.t1 CDS (phytozome数据库)

  氨基酸序列Cz03g14080.t1 (phytozome数据库)
  Cz06g04190.t1 (phytozome数据库)
  Cz09g23020.t1 (phytozome数据库)

DGAT基因及氨基酸序列

- 三角褐指藻*Phaeodactylum tricornutum*

  基因序列(CDs) (Genbank: >JF828904.1)

  氨基酸序列（Genbank: >AEG79729.1）

- 拟南芥*Arabidopsis thaliana*

  基因序列(CDs) AT1G48300.1 CDS

  AT3G51520.1 CDS

  氨基酸序列 AT1G48300.1

  AT3G51520.1

- 花生*Arachis hypogaea*

  基因序列(CDs) GenBank: KC736067.1

  GenBank: EU183333.1

  氨基酸序列GenBank: AGT57760.1

  GenBank: ABW34442.1

DGAT基因及氨基酸序列

- 小鼠*Mus musculus*

  基因序列(CDs) GenBank: AF078752.1
  氨基酸序列GenBank: AAC72917.1

- 大鼠*Rattus norvegicus*

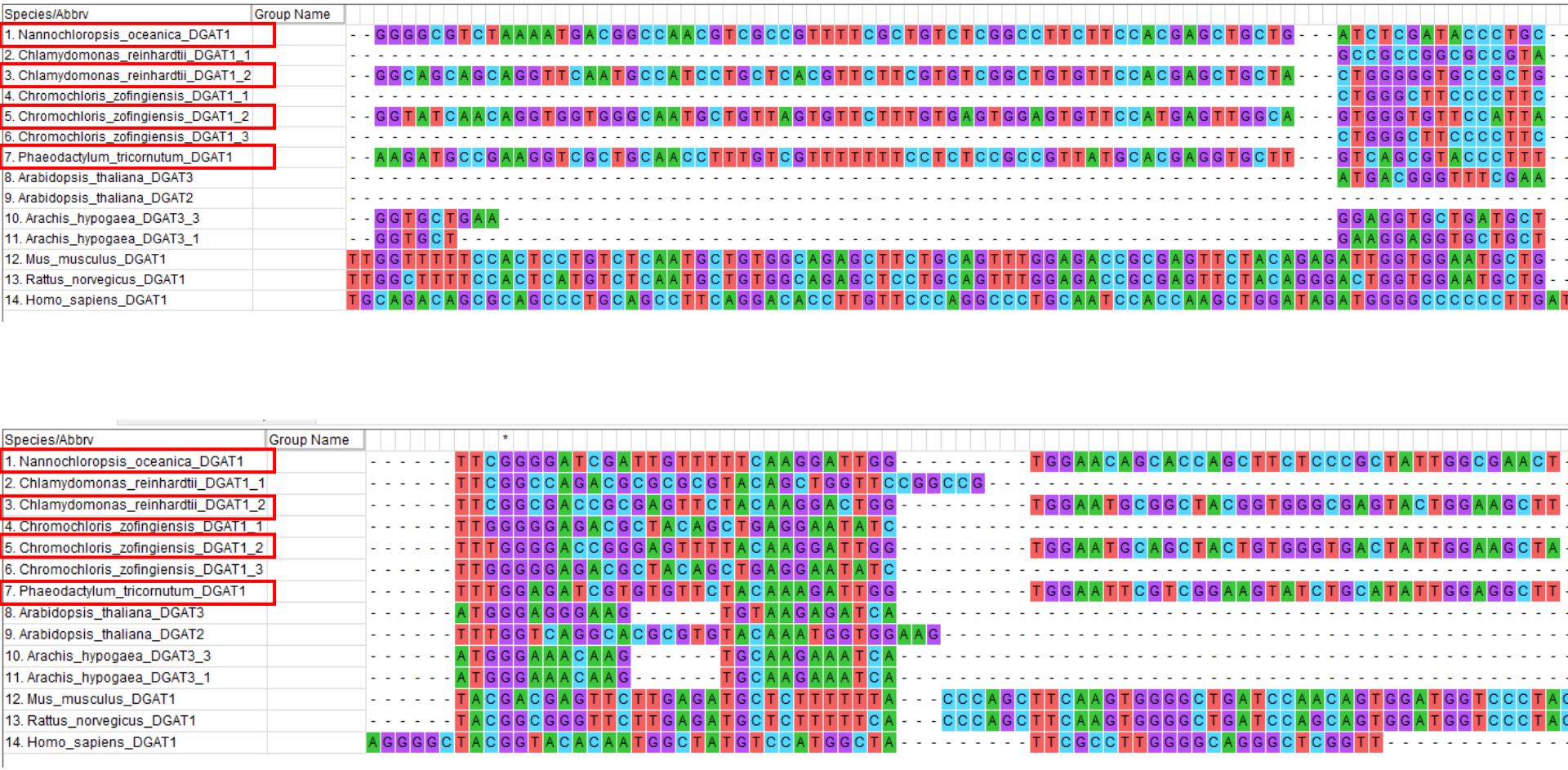
  基因序列(CDs) GenBank: AB062763.1
  氨基酸序列 GenBank: BAC43743.1

- 人*Homo sapiens*

  基因序列(CDs) NCBI Reference Sequence: NM_012079.5
  氨基酸序列NP_036211.2

多序列比对

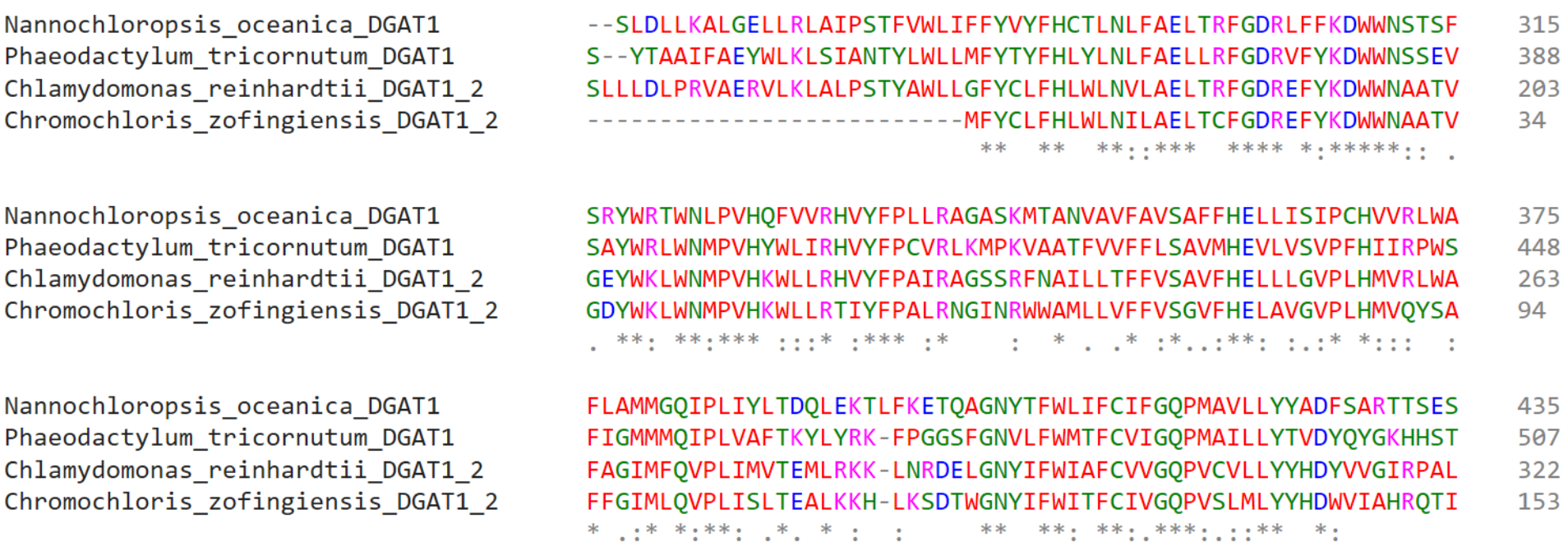
利用Muscle进行mRNA多序列比对



利用ClustalOmega进行蛋白多序列比对

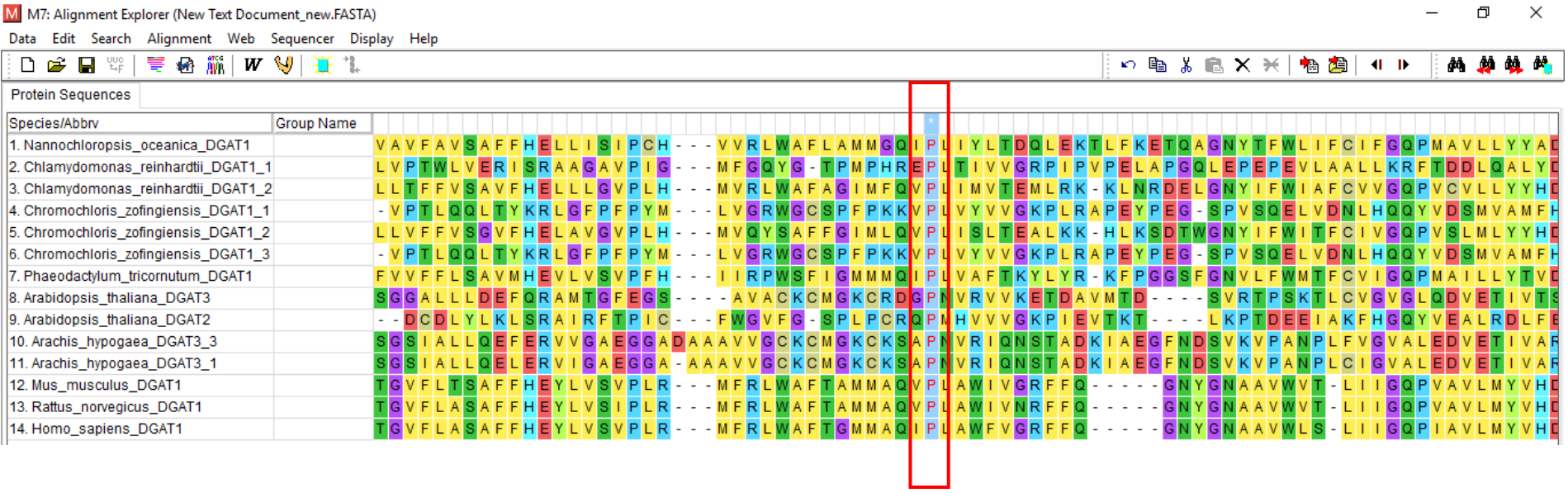

```
Arabidopsis_thaliana_DGAT3                ------------------------------------------------------------  360
Arachis_hypogaea_DGAT3_3                  ------------------------------------------------------------  340
Arachis_hypogaea_DGAT3_1                  ------------------------------------------------------------  340
Homo_sapiens_DGAT1                        GDREFYRDWWNSESVTYFWQNWNIPVHKWCIRHFYKPMLRRGSSKWMARTGVFLASAFFH  415
Mus_musculus_DGAT1                        GDREFYRDWWNAESVTYFWQNWNIPVHKWCIRHFYKPMLRHGSSKWVARTGVFLTSAFFH  426
Rattus_norvegicus_DGAT1                   GDREFYRDWWNAESVTYFWQNWNIPVHKWCIRHFYKPMLRLGSNKWMARTGVFLASAFFH  428
Nannochloropsis_oceanica_DGAT1            GDRLFFKDWWNSTSFSRYWRTWNLPVHQFVVRHVYFPLLRAGASKMTANVAVFAVSAFFH  360
Phaeodactylum_tricornutum_DGAT1           GDRVFYKDWWNSSEVSAYWRLWNMPVHYWLIRHVYFPCVRLKMPKVAATFVVFFLSAVMH  433
Chlamydomonas_reinhardtii_DGAT1_2         GDREFYKDWWNAATVGEYWKLWNMPVHKWLLRHVYFPAIRAGSSRFNAILLTFFVSAVFH  248
Chromochloris_zofingiensis_DGAT1_2        GDREFYKDWWNAATVGDYWKLWNMPVHKWLLRTIYFPALRNGINRWWAMLLVFFVSGVFH  79
Chromochloris_zofingiensis_DGAT1_1        GETLQLRNIFDVPTLQ-QL----------TYKRLGFPFPYMLVGRWG-C--SPFPKKVPL  206
Chromochloris_zofingiensis_DGAT1_3        GETLQLRNIFDVPTLQ-QL----------TYKRLGFPFPYMLVGRWG-C--SPFPKKVPL  206
Chlamydomonas_reinhardtii_DGAT1_1         GQTRAYSWFRPGPPLVPTW----------LVERISRAAGAVPIGMFG-QYGTPMPHREPL  271
Arabidopsis_thaliana_DGAT2                GQARVYKWWKPDCDLY--------------LKLSRAIRFTPICFWG-VFGSPLPCRQPM  262

Arabidopsis_thaliana_DGAT3                -------GSVSY*--------------------------------------------  360
Arachis_hypogaea_DGAT3_3                  ------------------------------------------------------------  340
Arachis_hypogaea_DGAT3_1                  ------------------------------------------------------------  340
Homo_sapiens_DGAT1                        ---------KDMDYSRIIERLLKLA----VPNHLIWLI---FFYWLFHSCLNAVAELMQF  355
Mus_musculus_DGAT1                        ---------KDMDYSRIIERLLKLA----VPNHLIWLI---FFYWFFHSCLNAVAELLQF  366
Rattus_norvegicus_DGAT1                   ---------KDMDYSRIIERLLKLA----VPNHLIWLI---FFYWLFHSCLNAVAELLQF  368
Nannochloropsis_oceanica_DGAT1            --V----D--SLDLLKALGELLRLA----IPSTFVWLI---FFYVYFHCTLNLFAELTRF  300
Phaeodactylum_tricornutum_DGAT1           SDLDETNGS--YTAAIFAEYWLKLS----IANTYLWLL---MFYTYFHLYLNLFAELLRF  373
Chlamydomonas_reinhardtii_DGAT1_2         FSVFDSGGSLLLDLPRVAERVLKLA----LPSTYAWLL---GFYCLFHLWLNVLAELTRF  188
Chromochloris_zofingiensis_DGAT1_2        ------------------------------------------MFYCLFHLWLNILAELTCF  19
Chromochloris_zofingiensis_DGAT1_1        VRALKEKGAVMLC-PGGQAELLHCYRAFKQQREFVLHIGHKGFCKLAIEHGAAVVPVLAL  160
Chromochloris_zofingiensis_DGAT1_3        VRALKEKGAVMLC-PGGQAELLHCYRAFKQQREFVLHIGHKGFCKLAIEHGAAVVPVLAL  160
Chlamydomonas_reinhardtii_DGAT1_1         SGLLRARKVAVLV-PGGVQEVLNME----HGKEVAYLSSRTGFVRLAVQHGAPLVPVWAF  222
Arabidopsis_thaliana_DGAT2                TSLLDSGYSCVLV-PGGVQETFHMQ----HDAENVFLSRRRGFVRIAMEQGSPLVPVFCF  218
```

藻类之间的蛋白多序列比对



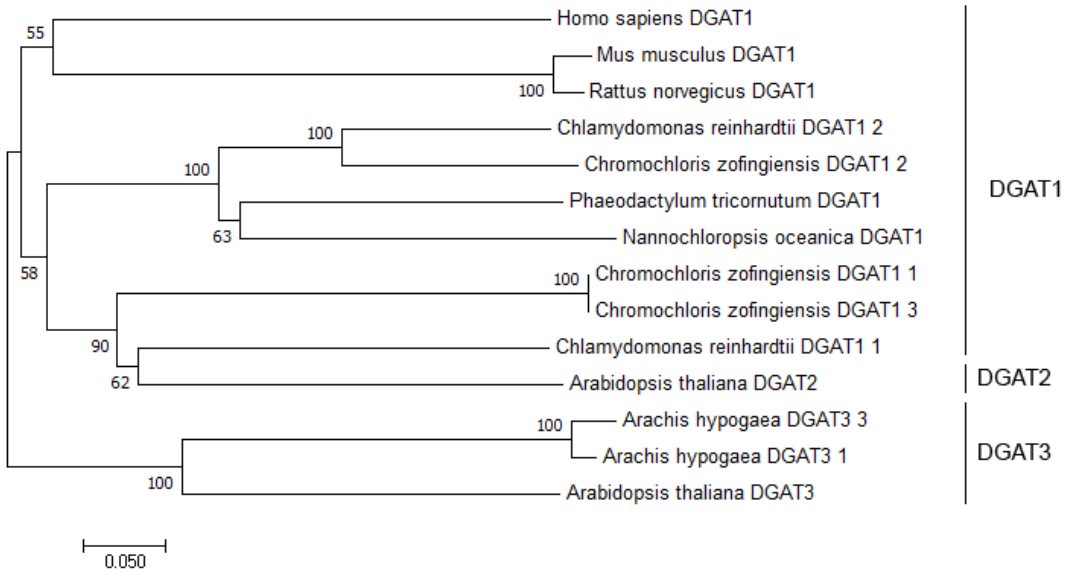
系统发育树构建

利用ClustalW进行多序列比对

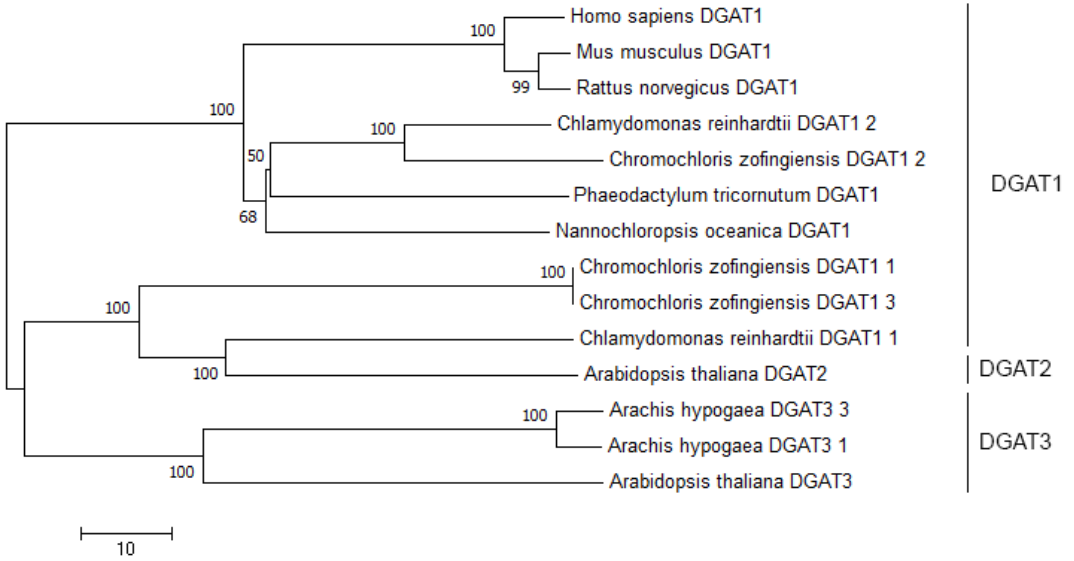


DGAT家族系统发育树构建

mRNA



Protein



结构分析

氨基酸组成

Amino acid composition:

Ala (A)	36	8.2%
Arg (R)	24	5.5%
Asn (N)	13	3.0%
Asp (D)	13	3.0%
Cys (C)	10	2.3%
Gln (Q)	11	2.5%
Glu (E)	11	2.5%
Gly (G)	13	3.0%
His (H)	19	4.3%
Ile (I)	23	5.3%
Leu (L)	60	13.7%
Lys (K)	10	2.3%
Met (M)	11	2.5%
Phe (F)	37	8.5%
Pro (P)	22	5.0%
Ser (S)	36	8.2%
Thr (T)	26	5.9%
Trp (W)	14	3.2%
Tyr (Y)	18	4.1%
Val (V)	30	6.9%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

密码子偏好性

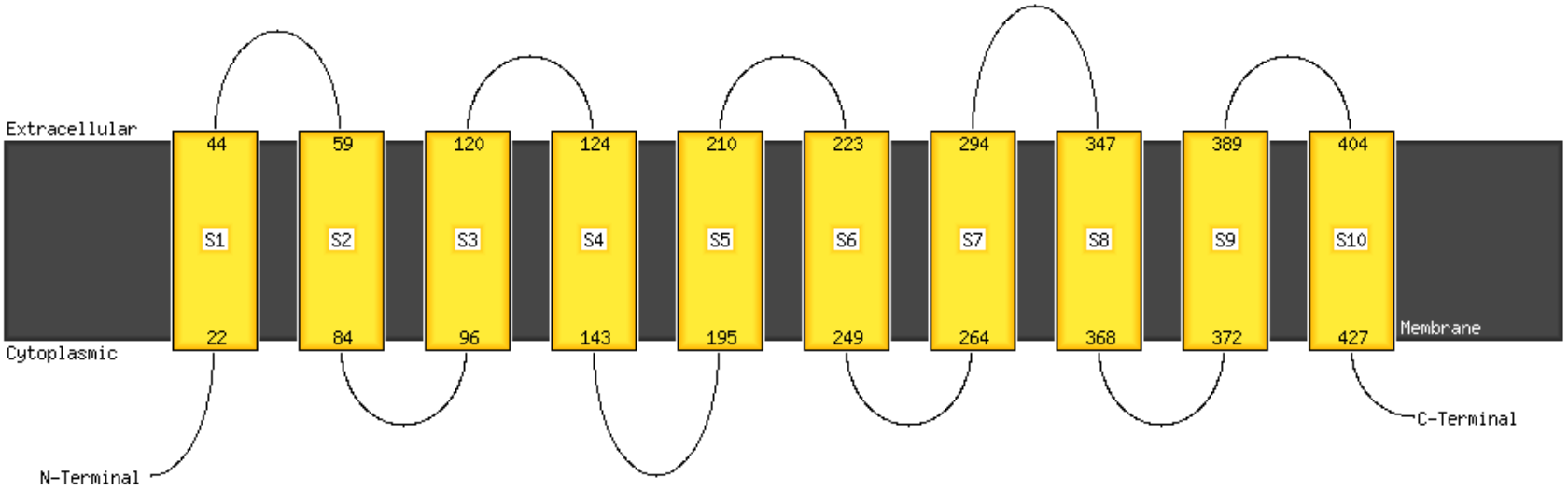
#CdsCount: 1

#Coding GC 56.47%
#1st letter GC 51.83%
#2nd letter GC 41.32%
#3rd letter GC 76.26%

#Codon	AA	Fraction	Frequency	Number
GCA	A	0.139	11.416	5
GCC	A	0.444	36.530	16
GCG	A	0.250	20.548	9
GCT	A	0.167	13.699	6
TGC	C	0.800	18.265	8
TGT	C	0.200	4.566	2
GAC	D	0.538	15.982	7
GAT	D	0.462	13.699	6
GAA	E	0.364	9.132	4
GAG	E	0.636	15.982	7
TTC	F	0.811	68.493	30
TTT	F	0.189	15.982	7
GGA	G	0.154	4.566	2
GGC	G	0.308	9.132	4
GGG	G	0.231	6.849	3
GGT	G	0.308	9.132	4
CAC	H	0.737	31.963	14
CAT	H	0.263	11.416	5
ATA	I	0.087	4.566	2
ATC	I	0.783	41.096	18
ATT	I	0.130	6.849	3
AAA	K	0.300	6.849	3
AAG	K	0.700	15.982	7
CTA	L	0.067	9.132	4
CTC	L	0.333	45.662	20
CTG	L	0.267	36.530	16
CTT	L	0.133	18.265	8
TTA	L	0.033	4.566	2
TTG	L	0.167	22.831	10
ATG	M	1.000	25.114	11
AAC	N	0.615	18.265	8
AAT	N	0.385	11.416	5
CCA	P	0.182	9.132	4
CCC	P	0.545	27.397	12
CCG	P	0.182	9.132	4
CCT	P	0.091	4.566	2
CAA	Q	0.000	0.000	0
CAG	Q	1.000	25.114	11
AGA	R	0.000	0.000	0
AGG	R	0.000	0.000	0
CGA	R	0.208	11.416	5
CGC	R	0.333	18.265	8
CGG	R	0.292	15.982	7
CGT	R	0.167	9.132	4
AGC	S	0.306	25.114	11
AGT	S	0.028	2.283	1
TCA	S	0.028	2.283	1
TCC	S	0.306	25.114	11
TCG	S	0.194	15.982	7
TCT	S	0.139	11.416	5
ACA	T	0.038	2.283	1
ACC	T	0.385	22.831	10
ACG	T	0.385	22.831	10
ACT	T	0.192	11.416	5
GTA	V	0.067	4.566	2
GTC	V	0.500	34.247	15
GTG	V	0.400	27.397	12
GTT	V	0.033	2.283	1
TGG	W	1.000	31.963	14
TAC	Y	0.722	29.680	13
TAT	Y	0.278	11.416	5
TAA	*	0.000	0.000	0
TAG	*	1.000	2.283	1
TGA	*	0.000	0.000	0

微拟球藻DGAT的二级结构预测

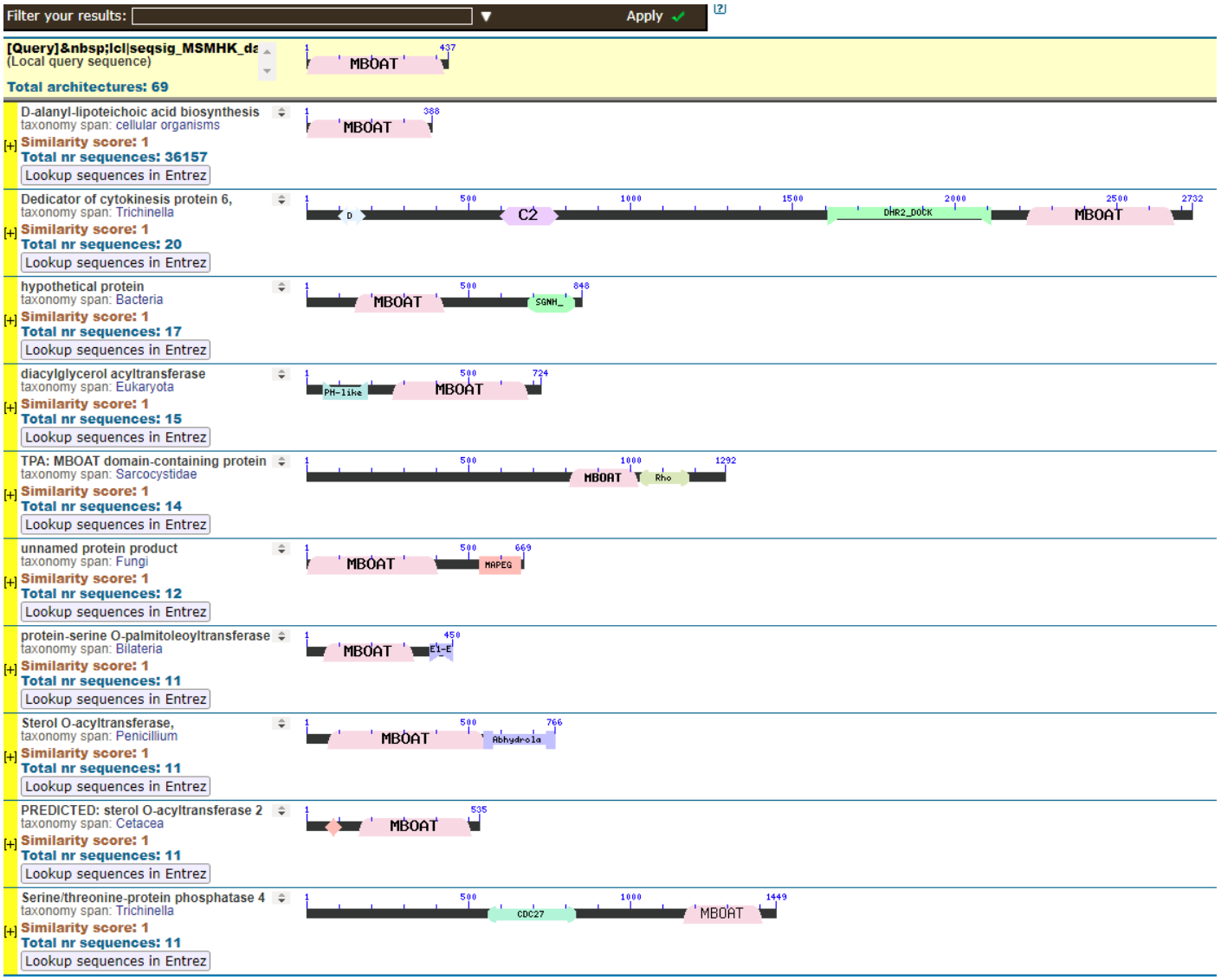
微拟球藻DGAT的跨膜螺旋预测



微拟球藻中的保守结构域

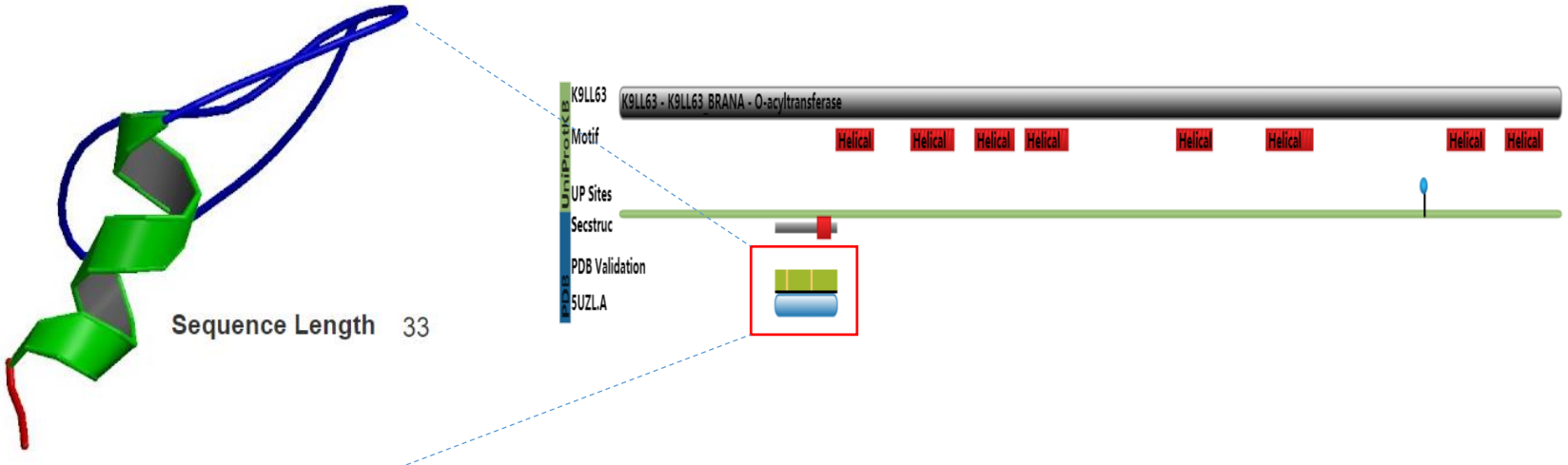
含有的结构域MBOAT, membrane-bound O-acyltransferase family

找到其他物种中的保守结构域



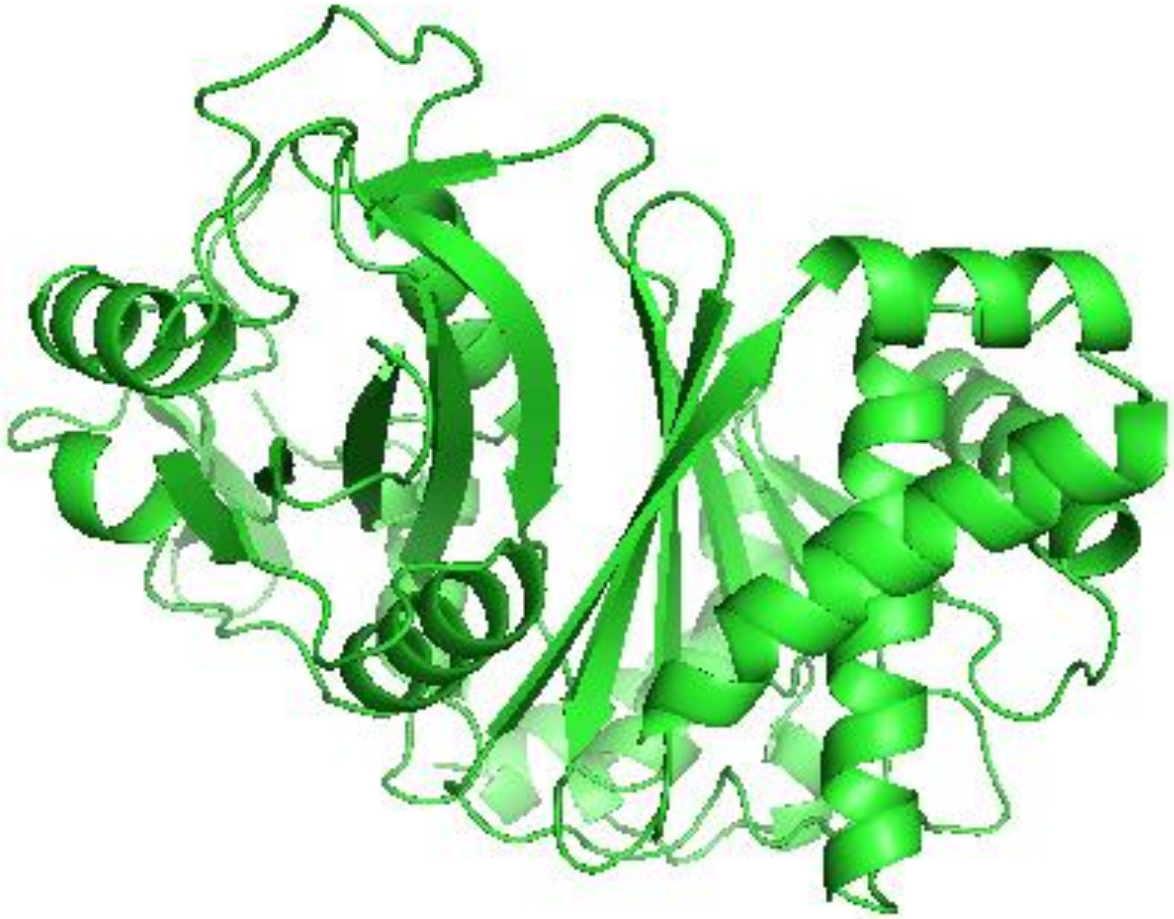
已经测定的结构：油菜O-乙酰基转移酶(NMR)

Brassica napus DGAT1 exosite

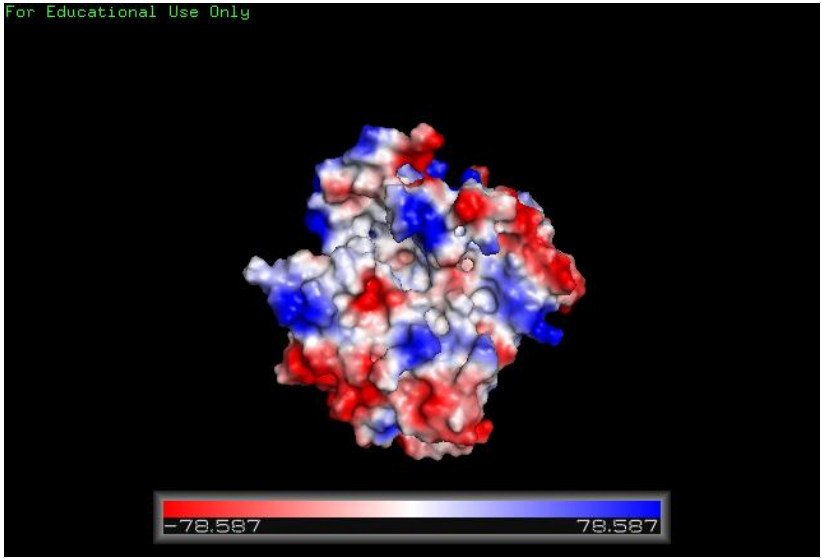


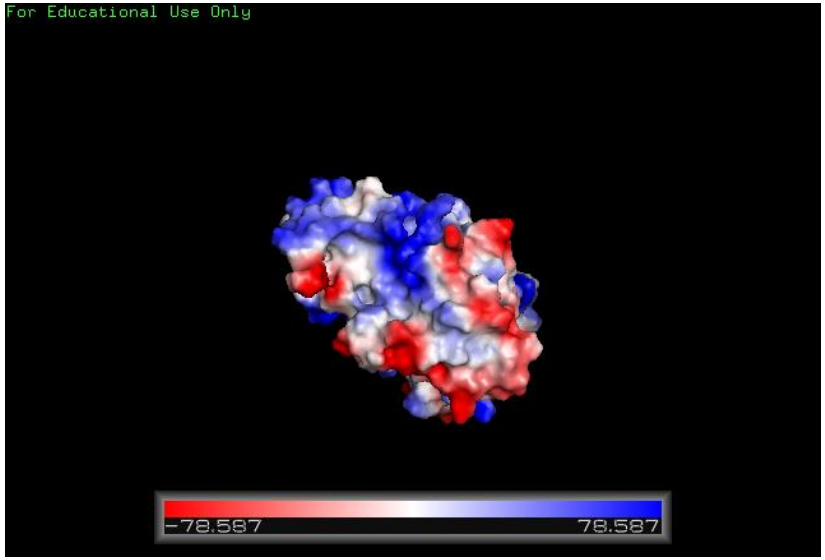
https://www.rcsb.org/structure/5UZL

来自海洋细菌的二酰基甘油酰基转移酶



蛋白活性口袋


For Educational Use Only
-78.587
78.587


For Educational Use Only
-78.587
78.587

Phyre2 预测：微拟球藻DGAT

三级结构预测 normal：

Top model

Image coloured by rainbow N → C terminus

Model dimensions (Å): **X**:26.997 **Y**:16.331 **Z**:18.285

Model (left) based on template c5uzlA_

Top template information

PDB header:transferase
Chain: A: **PDB Molecule:**o-acyltransferase;
PDBTitle: brassica napus dgat1 exosite

Confidence and coverage

Confidence:	62.2%	Coverage:	3%

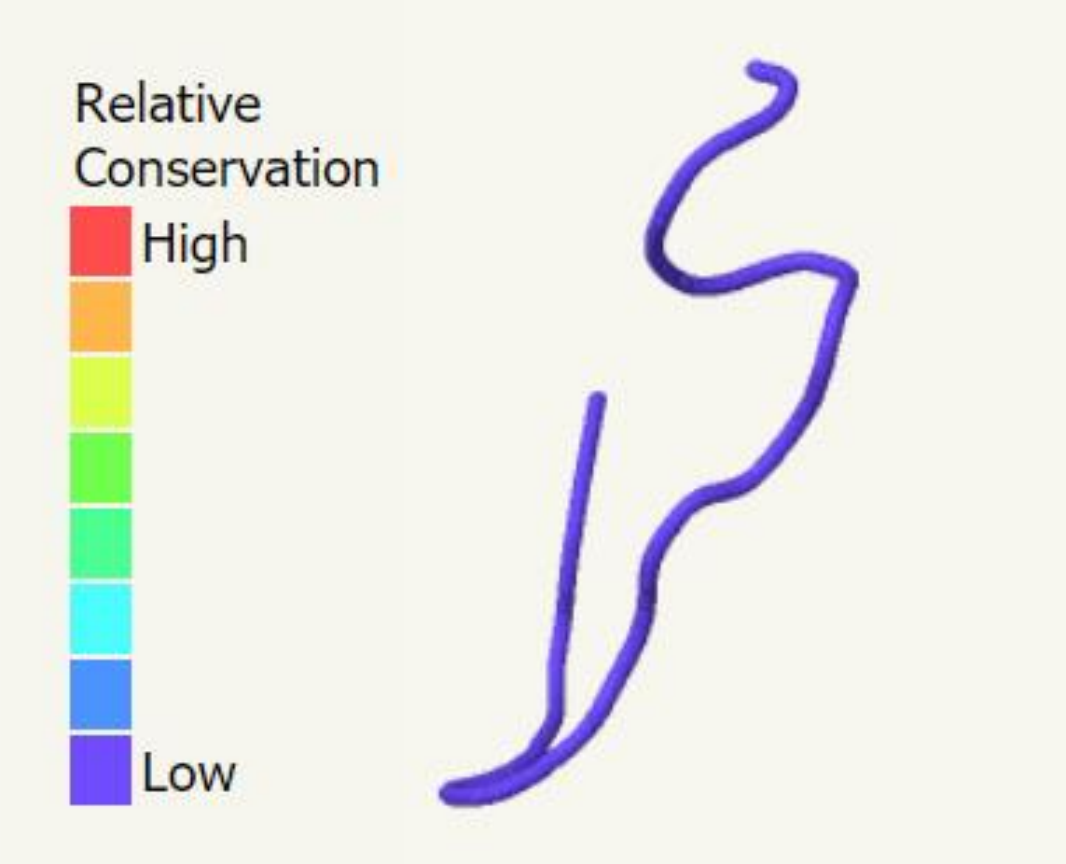
14 residues (3% of your sequence) have been modelled with 62.2% confidence by the single highest scoring template.

Confidence in the model: 62.2%

WARNING: Low confidence model. Please use with caution.

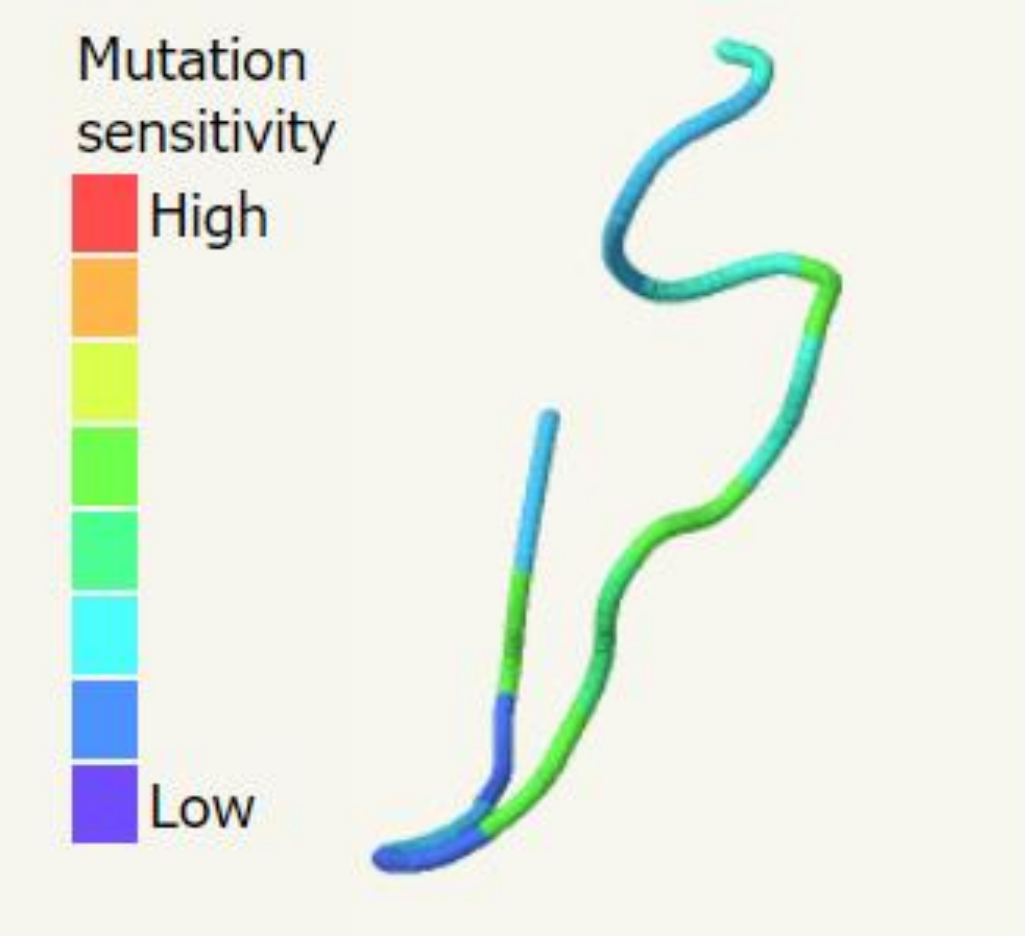
Phyre2 预测：微拟球藻DGAT

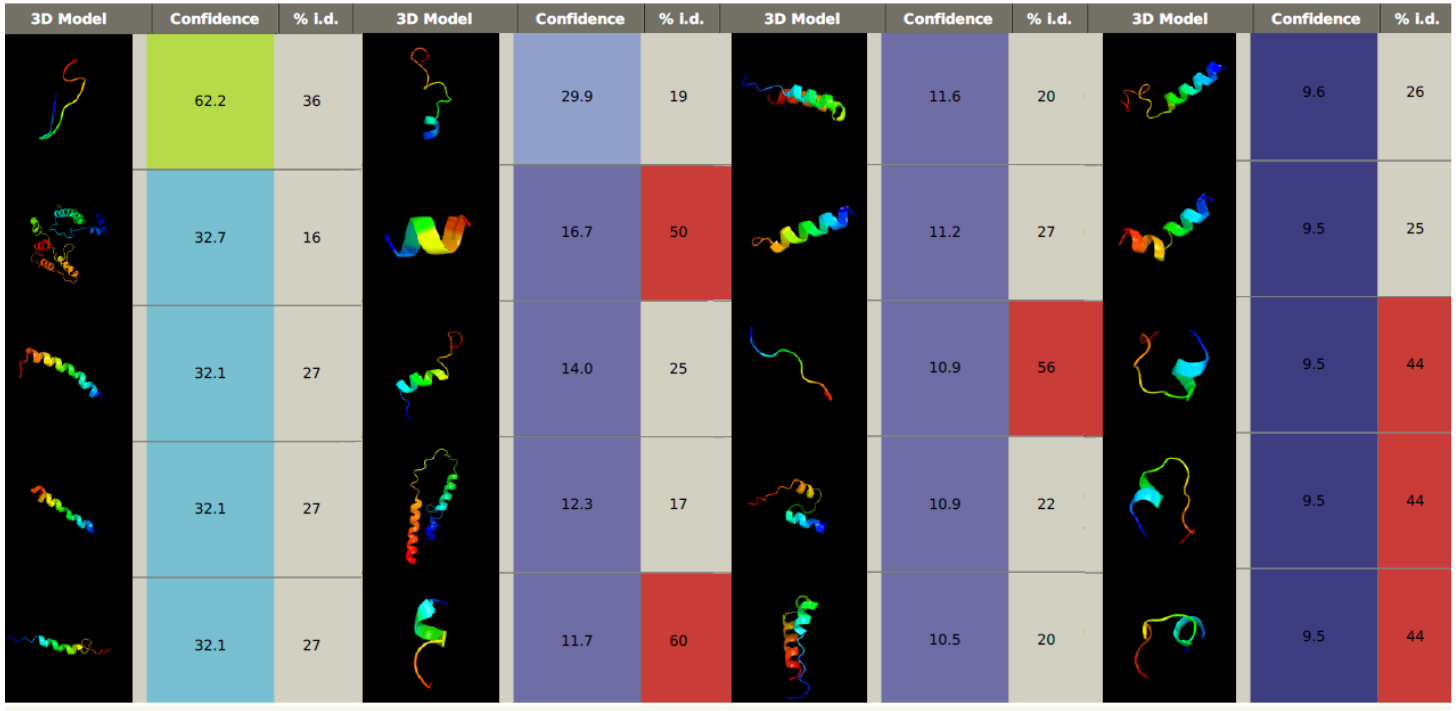
保守性



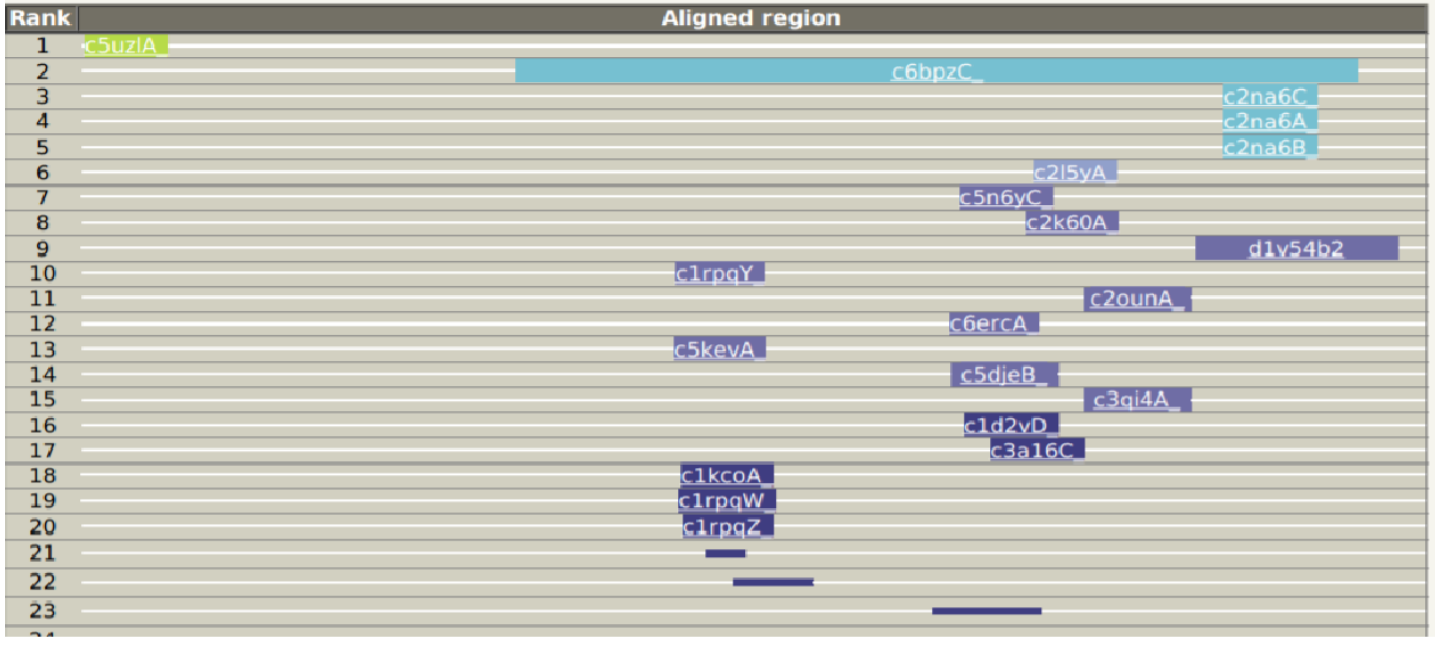
Phyre2 预测：微拟球藻DGAT

对突变的敏感性





Domain analysis

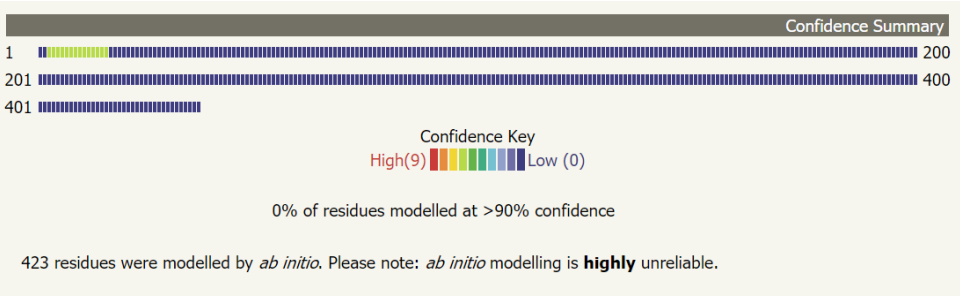


Phyre2 预测：微拟球藻DGAT

三级结构预测 intensive：

Image coloured by rainbow N → C terminus

Model dimensions (Å): **X**:64.261 **Y**:69.368 **Z**:78.838



Phyre2 预测：其他物种

衣藻 Chlamydomonas reinhardtii

Top model

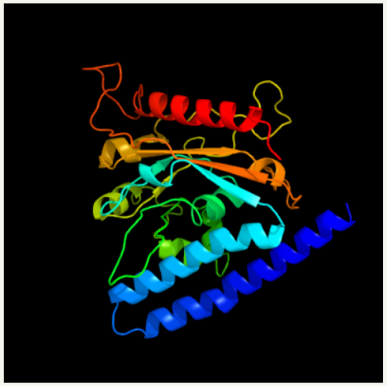


Image coloured by rainbow N → C terminus

Model dimensions (Å): **X**:50.079 **Y**:61.395 **Z**:52.060

Model (left) based on template c5kymA_

Top template information

PDB header:transferase
Chain: A: **PDB Molecule:**1-acyl-sn-glycerol-3-phosphate acyltransferase;
PDBTitle: crystal structure of the 1-acyl-sn-glycerophosphate (lpa)2 acyltransferase, plsc, from thermotoga maritima

Confidence and coverage

Confidence: **99.9%** Coverage: **72%**

234 residues (72% of your sequence) have been modelled with 99.9% confidence by the single highest scoring template.

小球藻Chromochloris zofingiensis

Top model

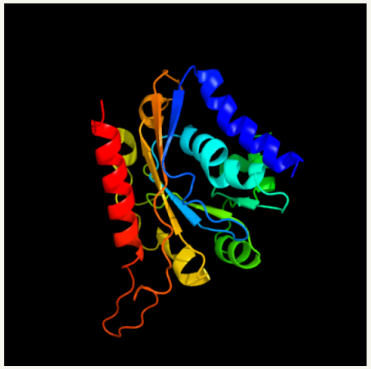


Image coloured by rainbow N → C terminus

Model dimensions (Å): **X**:44.902 **Y**:55.382 **Z**:44.483

Model (left) based on template c5kymA_

Top template information

PDB header:transferase
Chain: A: **PDB Molecule:**1-acyl-sn-glycerol-3-phosphate acyltransferase;
PDBTitle: crystal structure of the 1-acyl-sn-glycerophosphate (lpa)2 acyltransferase, plsc, from thermotoga maritima

Confidence and coverage

Confidence: **99.9%** Coverage: **73%**

193 residues (73% of your sequence) have been modelled with 99.9% confidence by the single highest scoring template.

拟南芥 Arabidopsis thaliana

Top model

Image coloured by rainbow N → C terminus

Model dimensions (Å): **X**:48.267 **Y**:57.978 **Z**:51.762

Model (left) based on template c5kymA_

Top template information

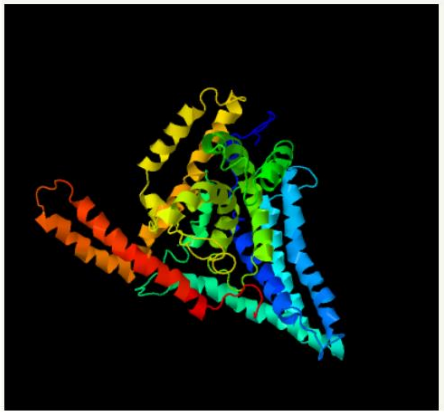
PDB header:transferase
Chain: A: **PDB Molecule:**1-acyl-sn-glycerol-3-phosphate acyltransferase;
PDBTitle: crystal structure of the 1-acyl-sn-glycerophosphate (lpa)2 acyltransferase, plsc, from thermotoga maritima

Confidence and coverage

Confidence: **100.0%** Coverage: **75%**

235 residues (75% of your sequence) have been modelled with 100.0% confidence by the single highest scoring template.

微拟球藻 Nannochloropsis.oceanica



总结

- 多序列比对表明微拟球藻的DGAT序列与藻类相似度更高，今后对微拟球藻的DGAT酶的研究中可以参考其他研究成熟的藻类
- 系统发育树表明DGAT1、DGAT2和DGAT3均具有跨物种的保守性，说明这三个蛋白在物种演化之前就已经存在
- 从结构上对微拟球藻进行了预测，可以为下一步对酶活性的优化提供一些建议

展望

随着环保挑战日益严峻，开发环保可再生的替代能源成为了当下的迫切任务。本次讨论报告通过与其他物种比对，并对DGAT结构进行预测，为后续研究提供两条思路：1.通过湿实验研究活性位点的突变是否会改变底物作用位点或增加酶活，从而提高TAG产量；2.基因共定位，解析DGAT作用机制，明确代谢途径，通过代谢工程手段提高油脂产量。

谢谢！