

Evolutionary analysis of PEPCase family

磷酸烯醇式丙酮酸羧化酶家族的 演化分析

报告人：王子猛

小组成员：叶思达 王子猛 许聪 侯婕

2015. 1. 25

Outline

- Background
- Sequence collecting
- Phylogeny
- Protein structure
- Discussion

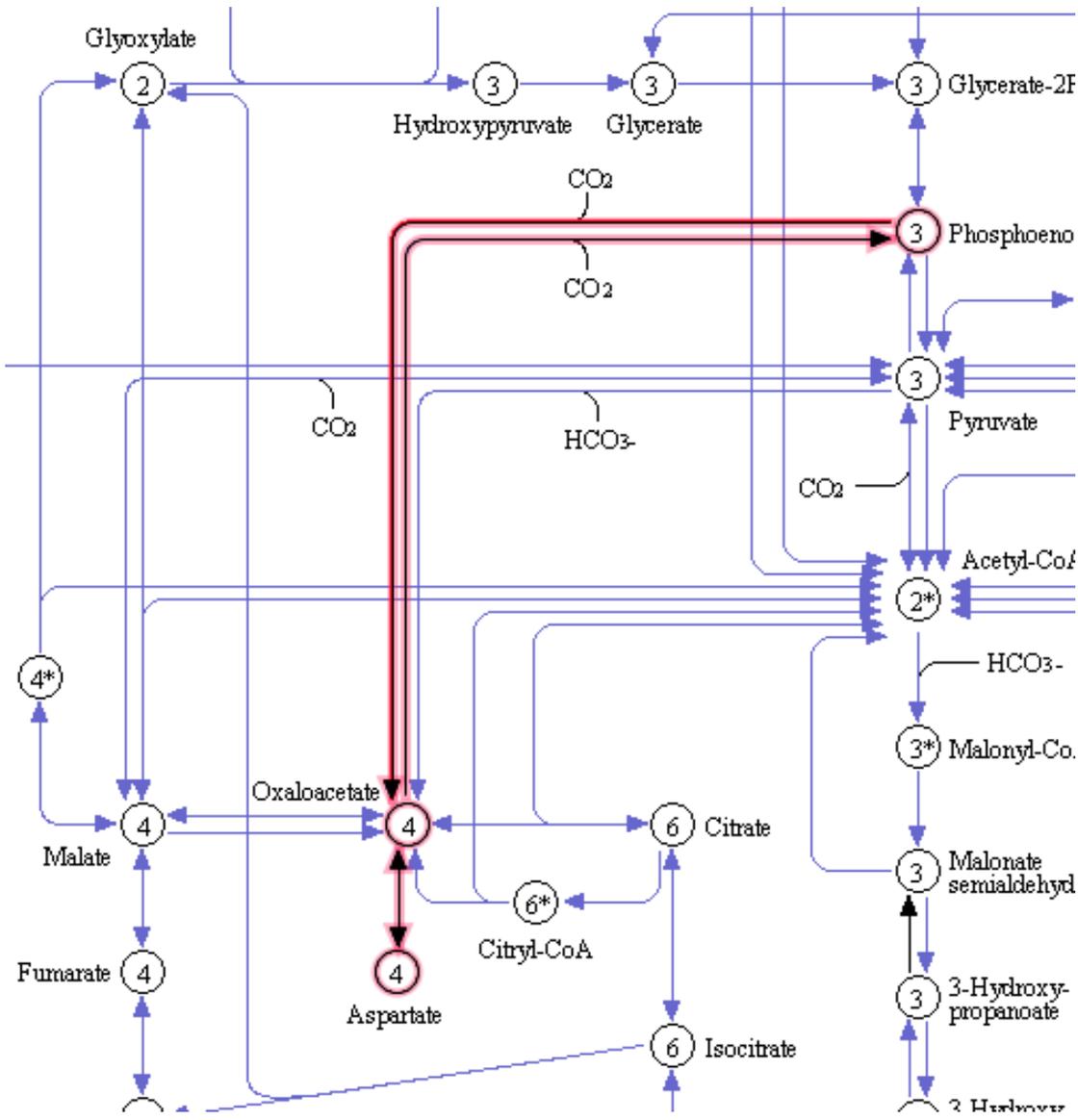
Background

- **PEPCase:**
 - phosphoenolpyruvate carboxylase(磷酸烯醇式丙酮酸羧化酶),
 - 简称**PEPCase**，也常被简称为ppc
- 起源非常古老，在现存的高等植物中，往往存在多个拷贝
- 在真菌中，尚未发现**PEPcase**的同源蛋白，原因不明

Function

- 对PEPCase的功能进行查询
 - 在KEGG数据库中，PEPCase(KO1595)被标注可参与5种通路：
 - Pyruvate metabolism: 丙酮酸代谢
 - Methane metabolism: 甲烷代谢
 - Carbon fixation in photosynthetic organisms
光合组织中碳的固定
 - Carbon fixation pathways in prokaryotes
原核生物中碳的固定
 - Carbon metabolism
含碳化合物的代谢

Background

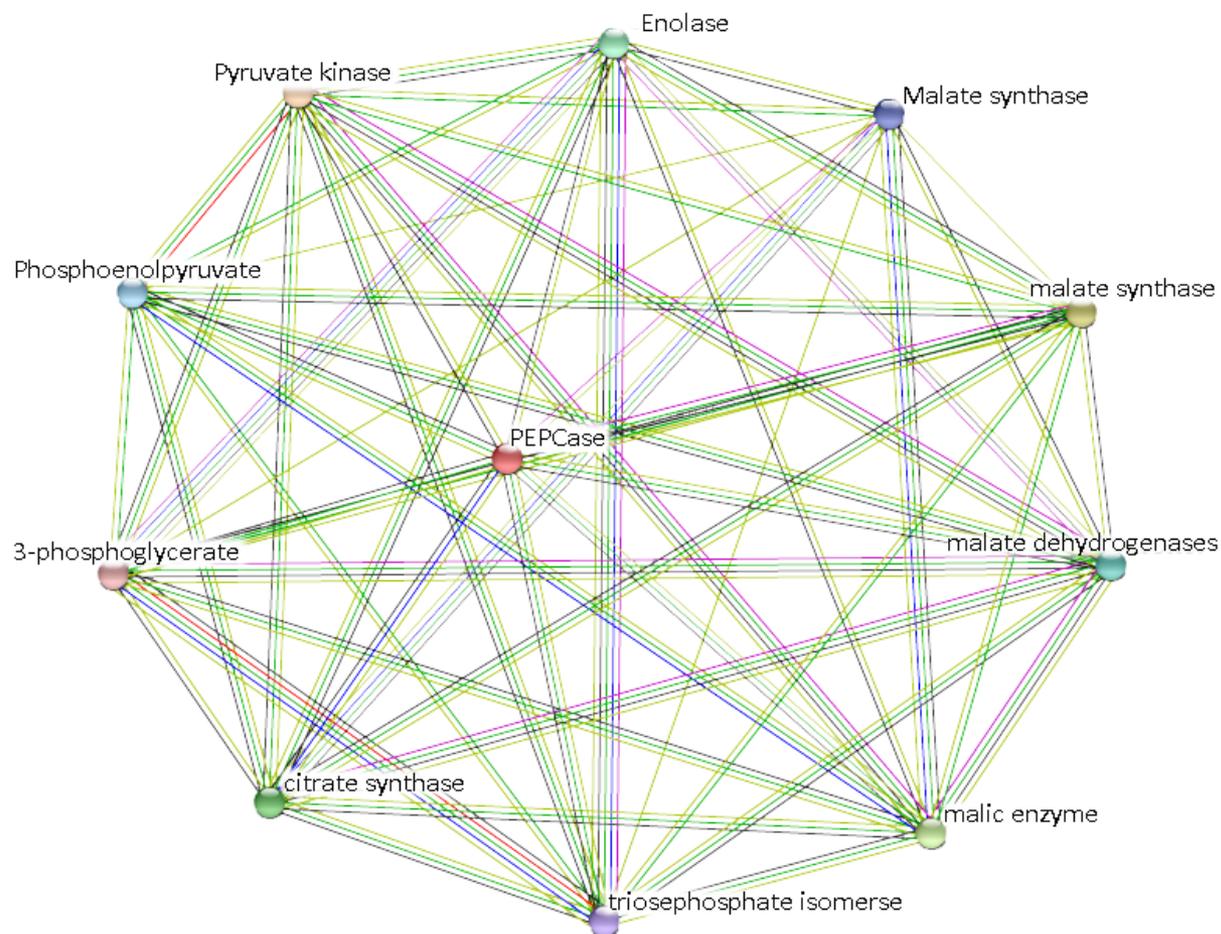


PEPCase参与的生化途径示意图。

PEPCase催化PEP生成草酰乙酸。在不同的生化途径中，PEPCase催化的反应是相同的。

(<http://www.genome.jp/kegg/>)

与PEPCase互作的基因



STRING数据库中与PEPCase有相互作用的因子示意图。

与PEPCase互作的蛋白，多数为含碳化合物代谢中的酶。

(<http://string.embl-heidelberg.de/>)

PEPCase与农业

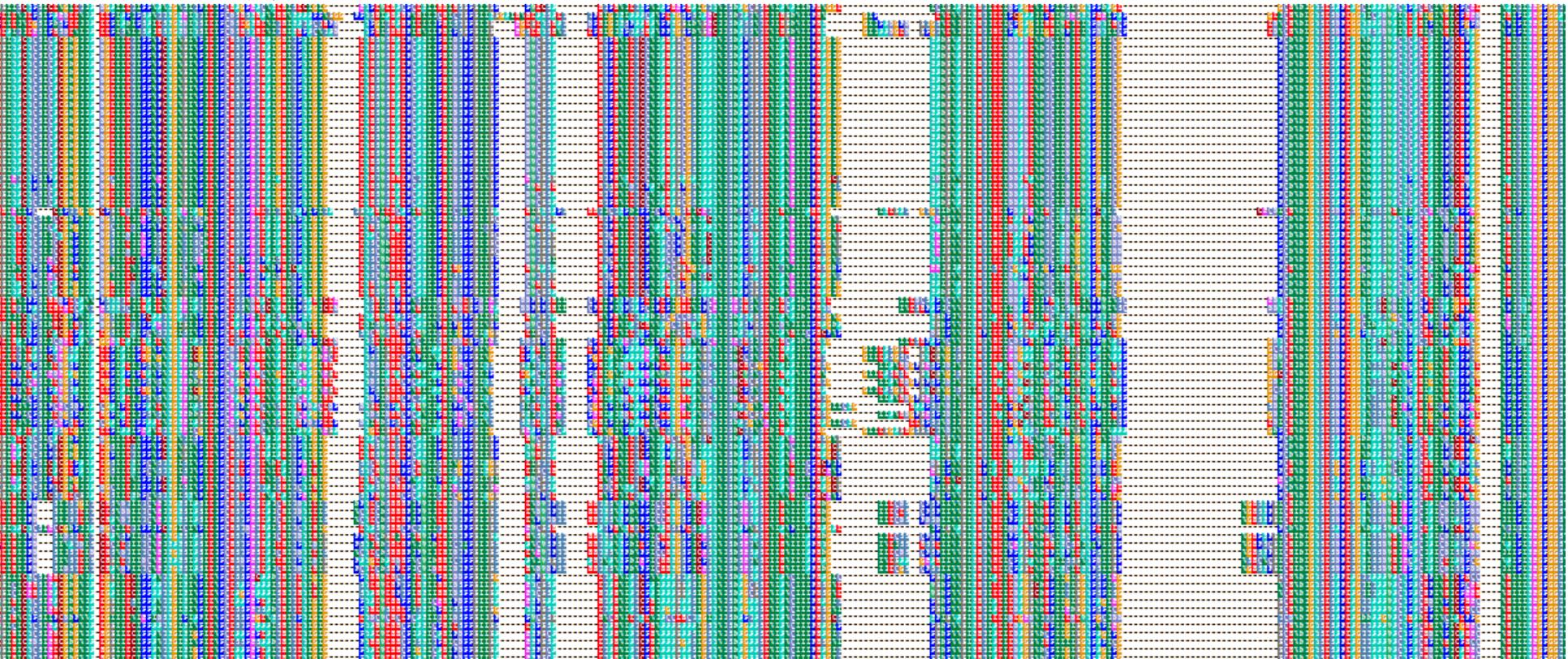
- **PEPCase**酶是**C4**光合途径或**CAM**(景天酸途径)光合途径的起始。
- 有研究组试图将**C3**农作物（如水稻）改造为**C4**作物，方法是转入**C4**光合途径所需的基因，包括**PEPCase**在内。

数据收集

- **AA序列数据来自PDB和NCBI**
- **数据分为两个部分：**
 - **PDB上：已审阅条目：264条**
 - **由NCBI补充：36条（已去除重复）**
 - **（PDB未审阅有11805条，未加入）**

序列比对

- 使用Mega6.0内置的clustalW进行序列比对，比对后全长1554AA



TIPs: clustal or Muscle?

(使用mega6.0中clustal进行比对，结果在BioEdit7.2.5中显示)

建树：算法的选择

- 对于亲缘关系较近者：不同算法区别不大。
- 对于亲缘关系较远者（跨门）：**ML(最大似然)法**，**Bayes法**
- 时间跨度：**Bayes≈ML>MP(最大简约法)**、**NJ(邻接法).....**
- 对于**PEPCase**的研究，跨越了不同的门，因此应选用**ML或Bayes法**。

AA替代模型检测

Table. Maximum Likelihood fits of 48 different amino acid substitution models

Model	Parameters	BIC	AICc	$\ln L$	(+I)	(+G)	$f(A)$	$f(R)$	$f(N)$	$f(D)$	$f(C)$	$f(Q)$	$f(E)$	$f(G)$	$f(H)$	$f(I)$
LG+G	624	143172.973	137109.138	-67927.407	n/a	1.04	0.079	0.056	0.042	0.053	0.013	0.041	0.072	0.057	0.022	0.062
LG+G+I	625	143184.701	137111.158	-67927.407	0.00	1.04	0.079	0.056	0.042	0.053	0.013	0.041	0.072	0.057	0.022	0.062
LG+G+F	643	143328.122	137079.849	-67893.567	n/a	1.03	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057
LG+G+I+F	644	143339.850	137081.870	-67893.567	0.00	1.03	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057
WAG+G	624	143608.562	137544.727	-68145.202	n/a	1.16	0.087	0.044	0.039	0.057	0.019	0.037	0.058	0.083	0.024	0.048
WAG+G+I	625	143620.290	137546.747	-68145.202	0.00	1.16	0.087	0.044	0.039	0.057	0.019	0.037	0.058	0.083	0.024	0.048
WAG+G+F	643	143770.706	137522.433	-68114.859	n/a	1.14	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057
WAG+G+I+F	644	143782.434	137524.454	-68114.859	0.00	1.14	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057
rtREV+G+F	643	144110.110	137861.838	-68284.561	n/a	1.03	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057
rtREV+G+I+F	644	144121.838	137863.859	-68284.561	0.00	1.03	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057
JTT+G	624	144339.416	138275.581	-68510.629	n/a	1.03	0.077	0.051	0.043	0.051	0.020	0.041	0.062	0.075	0.023	0.053
JTT+G+I	625	144351.144	138277.601	-68510.629	0.00	1.03	0.077	0.051	0.043	0.051	0.020	0.041	0.062	0.075	0.023	0.053
JTT+G+F	643	144420.032	138171.759	-68439.522	n/a	1.02	0.088	0.076	0.028	0.051	0.010	0.039	0.069	0.069	0.018	0.057

(使用mega6.0中进行替代模型的检测)

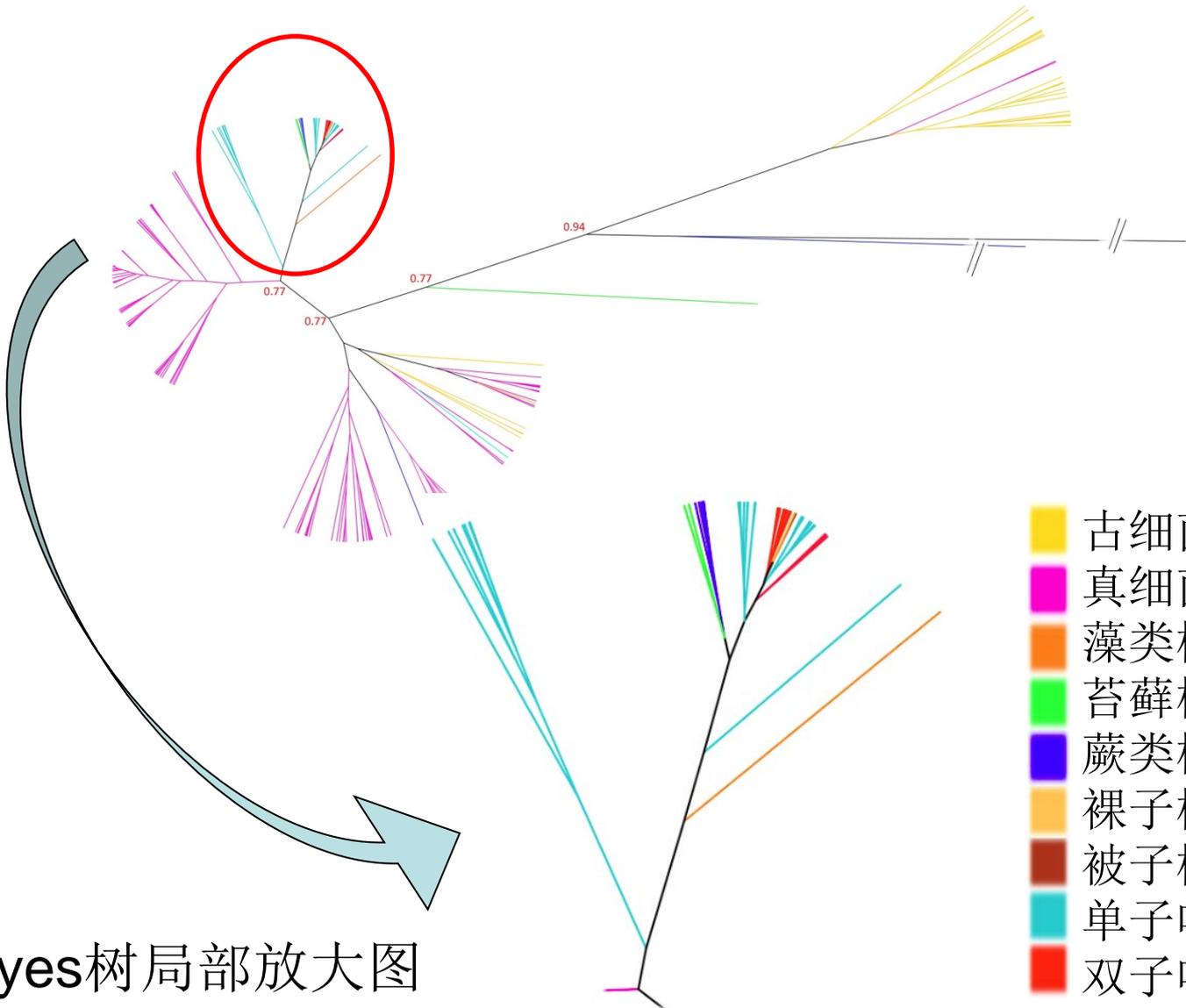
- 因为跨度大，不宜用MP树或NJ树，应用ML树或Payes树。而这两者均需要检测替代模型。
- **TIPs:** 在计算量较小时，推荐使用protest/jmodeltest进行替代模型的检测

Bayes树



- 使用Beast1.8.0建Bayes树，
使用FigureTree1.4.0进行调整并着色。
- 有一支伸出较远（文心兰+小立碗藓）

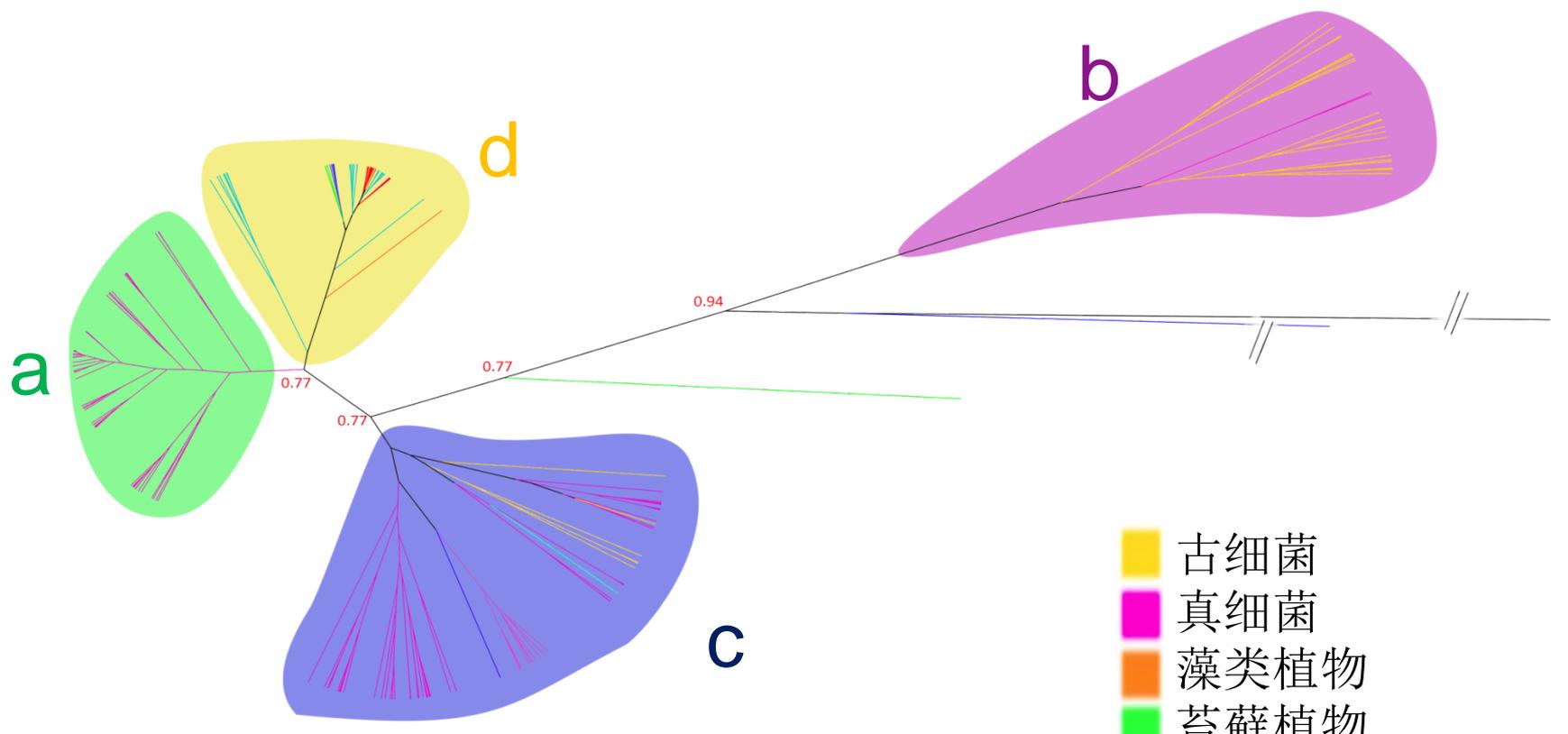
Bayes树分析



- 古细菌
- 真细菌
- 藻类植物
- 苔藓植物
- 蕨类植物
- 裸子植物
- 被子植物基部类群
- 单子叶植物
- 双子叶植物

Bayes树局部放大图

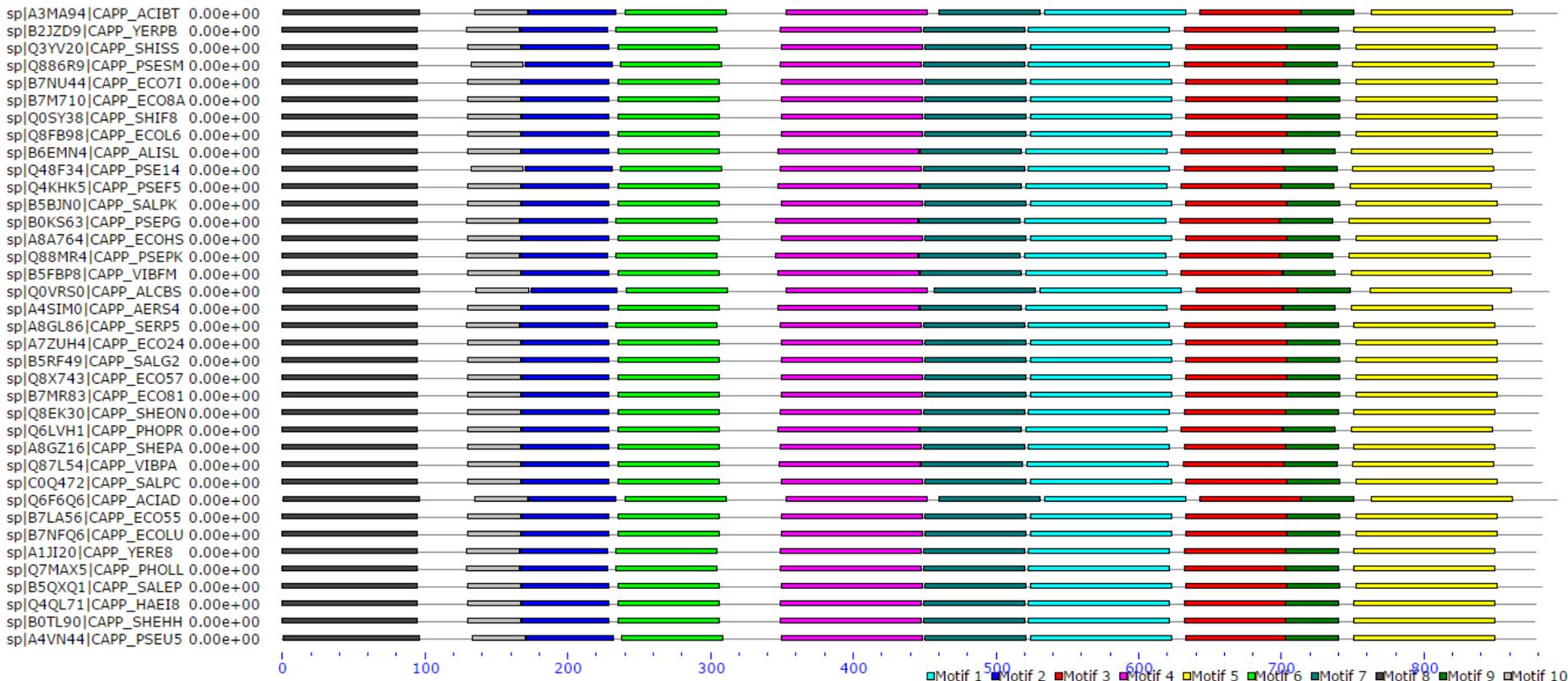
PEPCase家族分类



- 古细菌
- 真细菌
- 藻类植物
- 苔藓植物
- 蕨类植物
- 裸子植物
- 被子植物基部类群
- 单子叶植物
- 双子叶植物

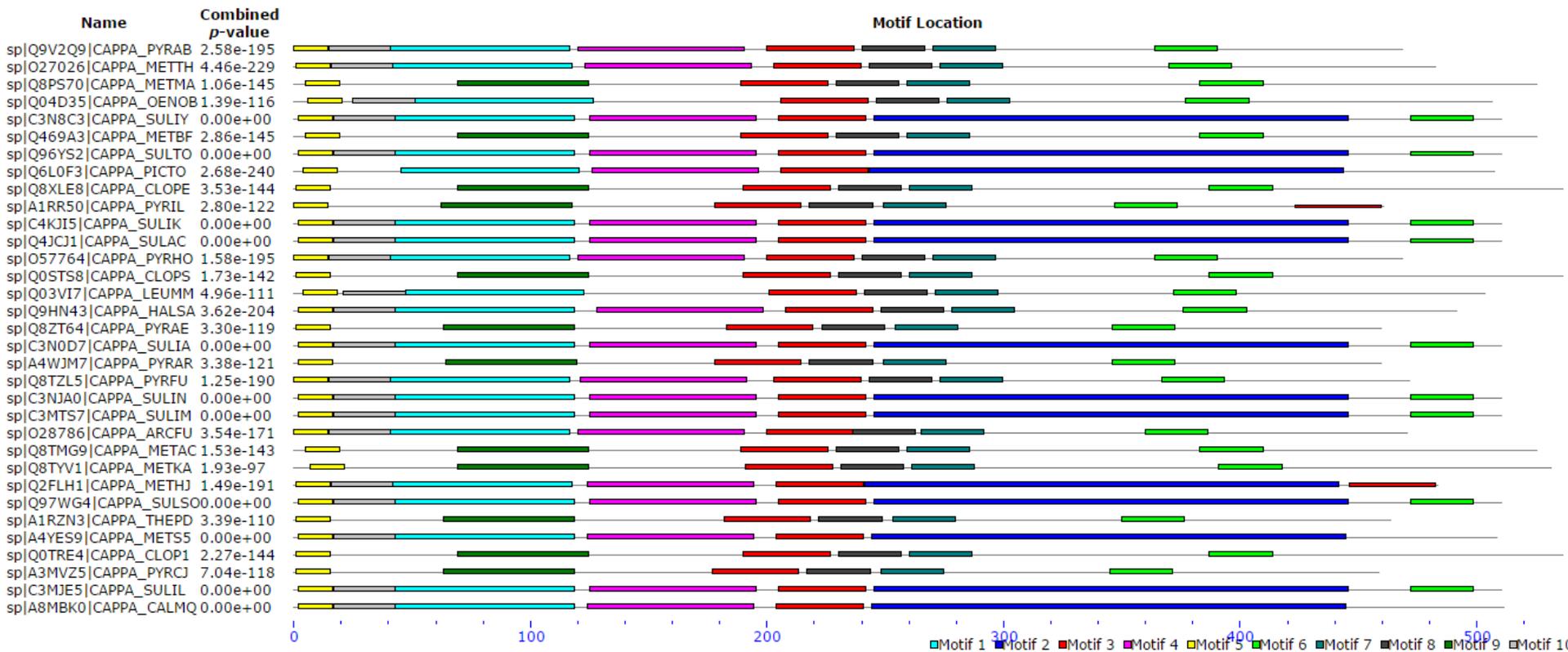
依据Bayes树的PEPCase家族成员分类图

MEME分析



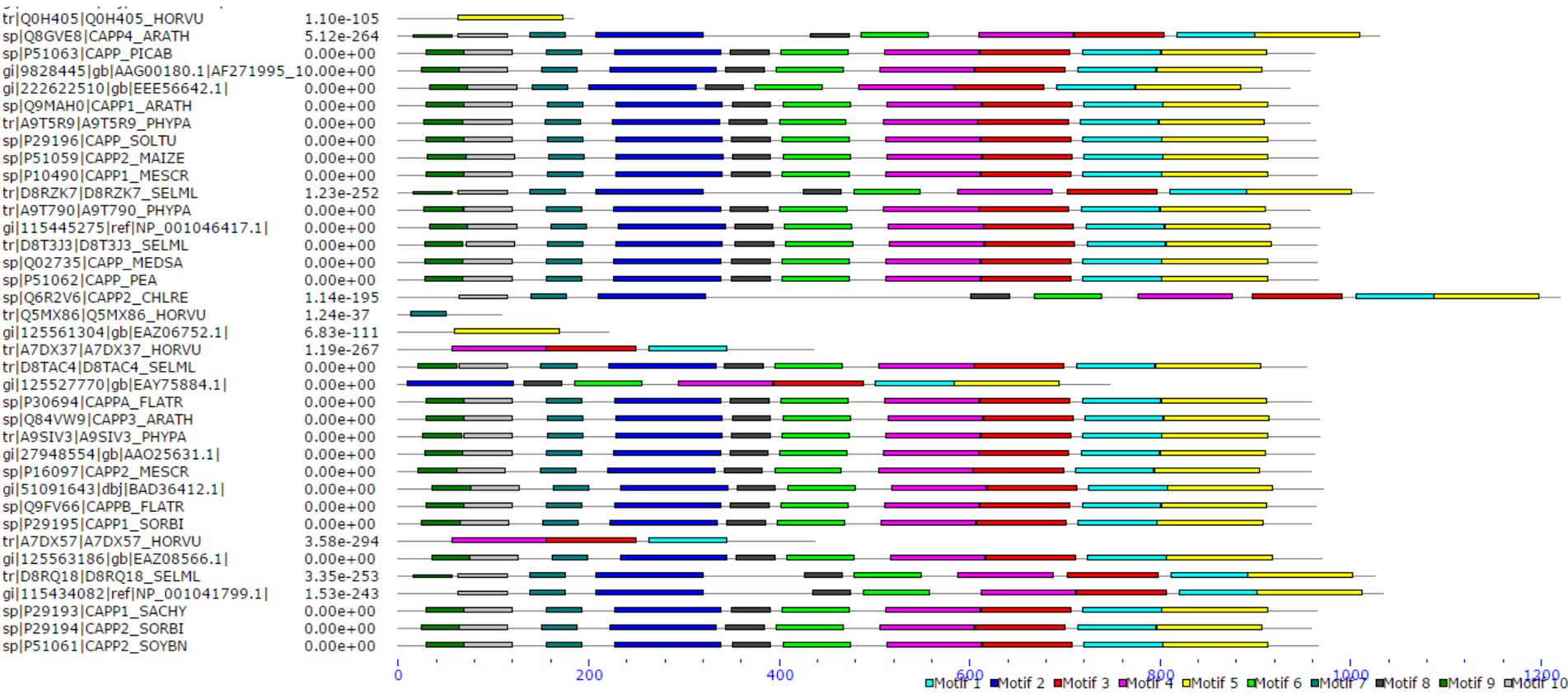
Beyes树a支(真细菌)MEME分析图(<http://meme.nbcr.net/meme/>)

MEME分析



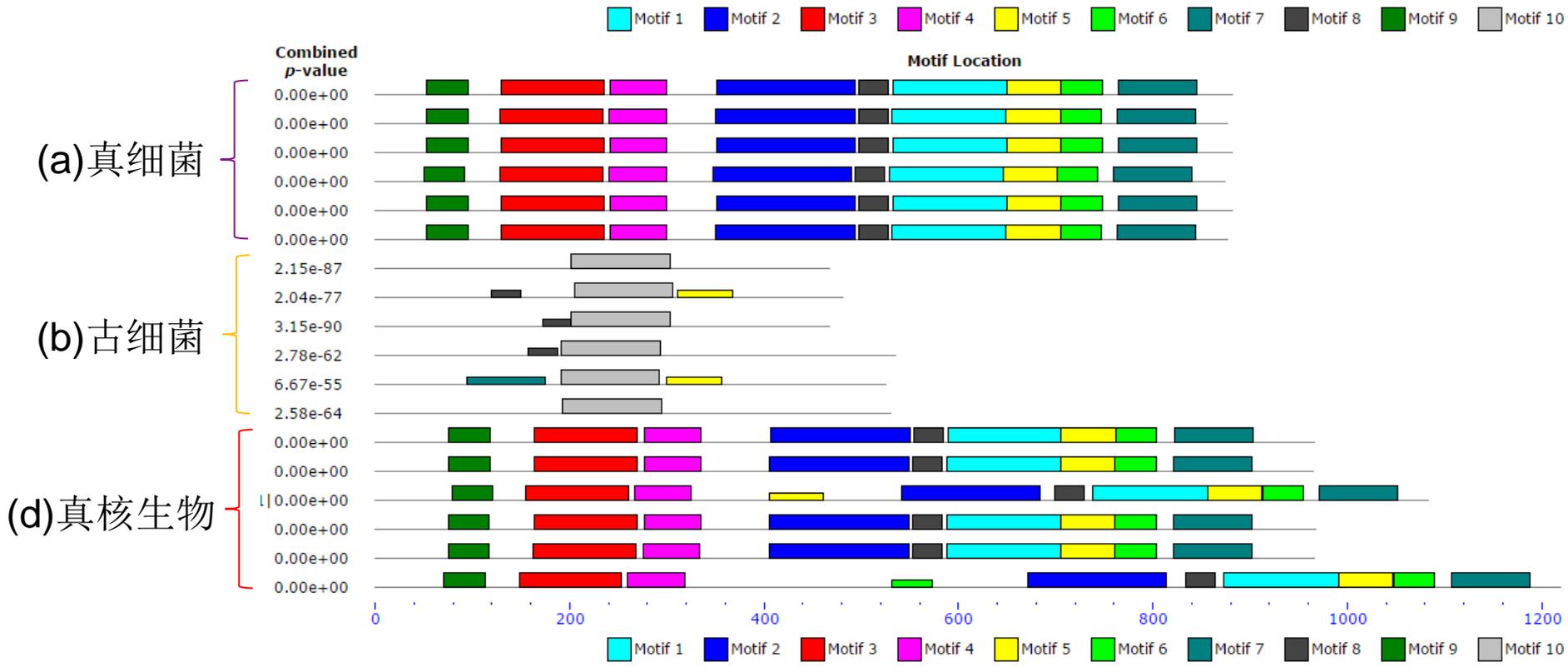
Beyes树b支(古细菌)MEME分析图(<http://meme.nbcr.net/meme/>)

MEME分析



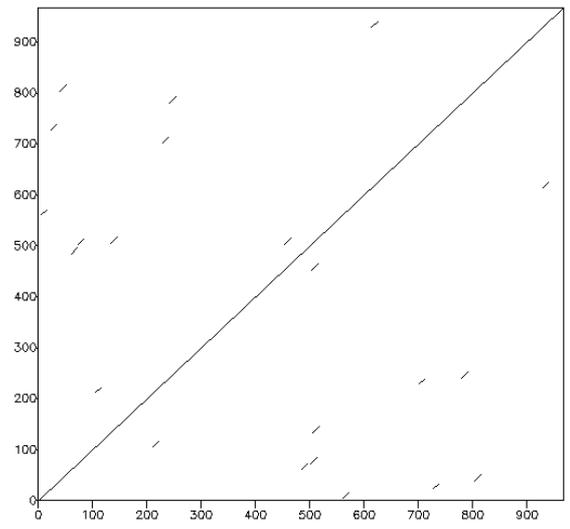
Beyes树d支(真核生物)MEME分析图(<http://meme.nbcr.net/meme/>)

MEME分析

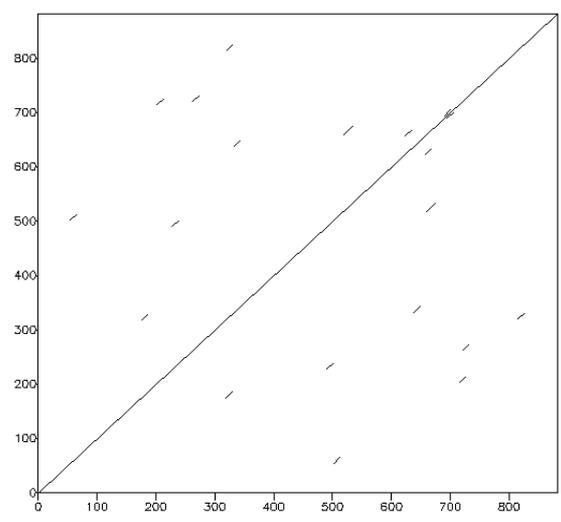


Beyes树abd支MEME分析图(<http://meme.nbcr.net/meme/>)

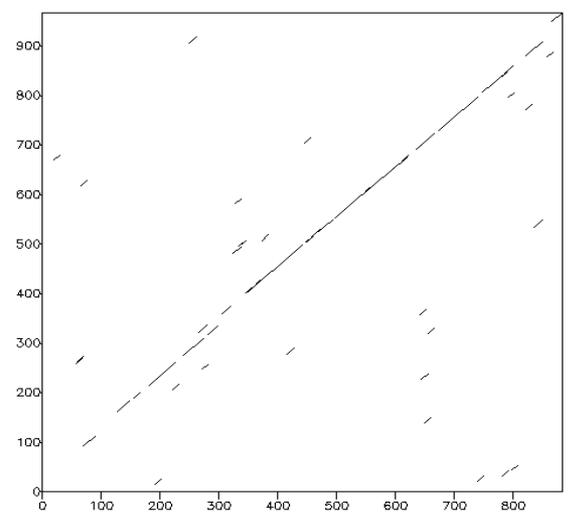
使用DotMatcher进行分析



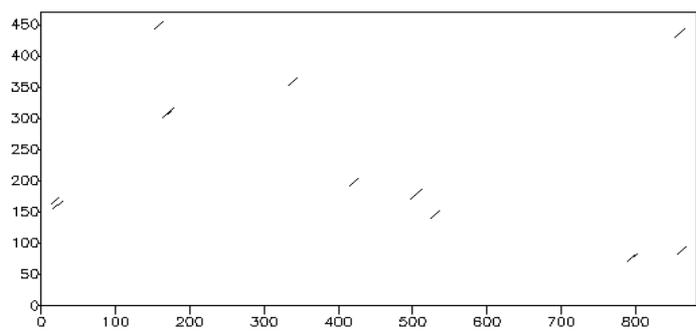
Arabidopsis(ppc3)



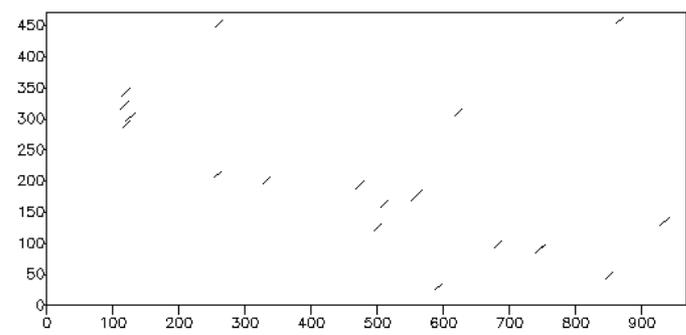
E. coli



At vs. E. coli

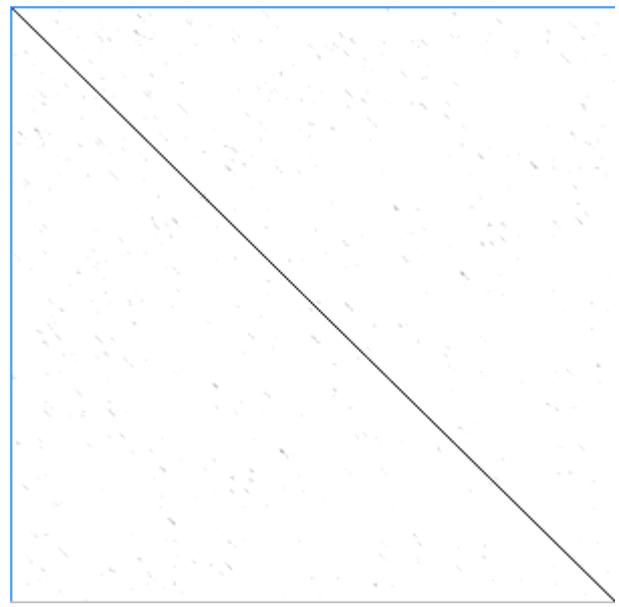


E. coli vs. 古细菌

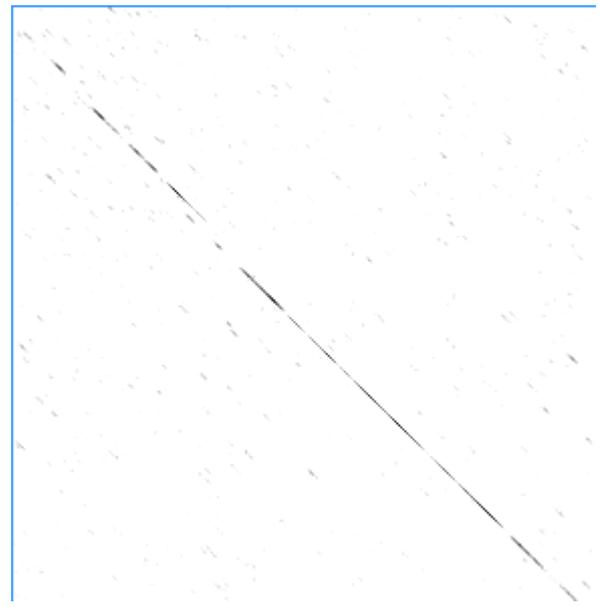


At vs. 古细菌

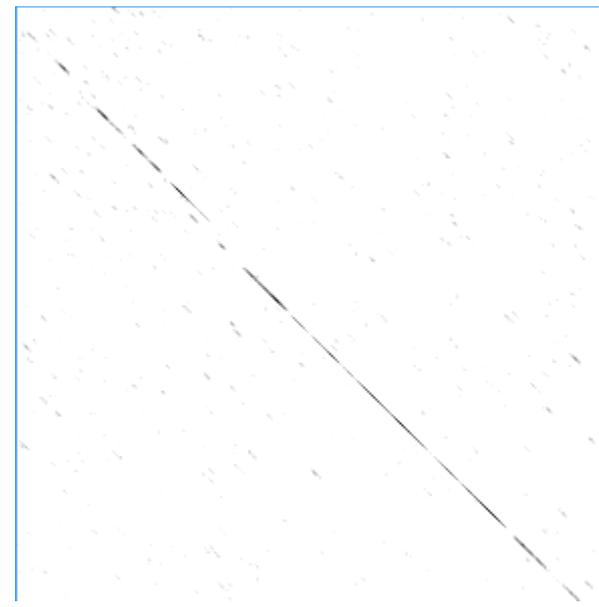
使用Dotlet进行分析



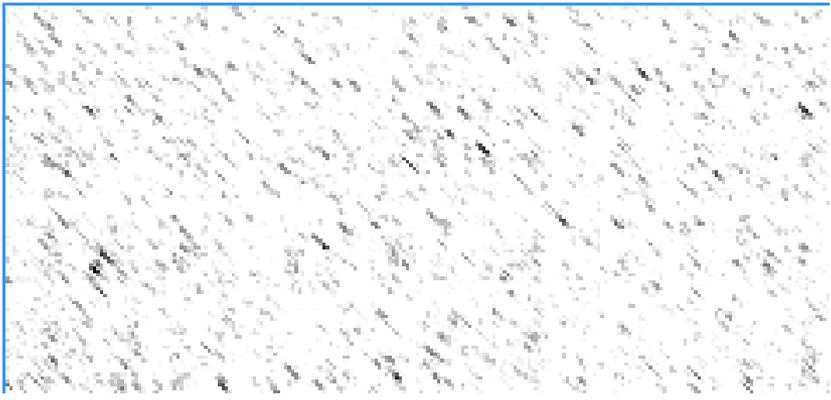
Arabidopsis(ppc3)



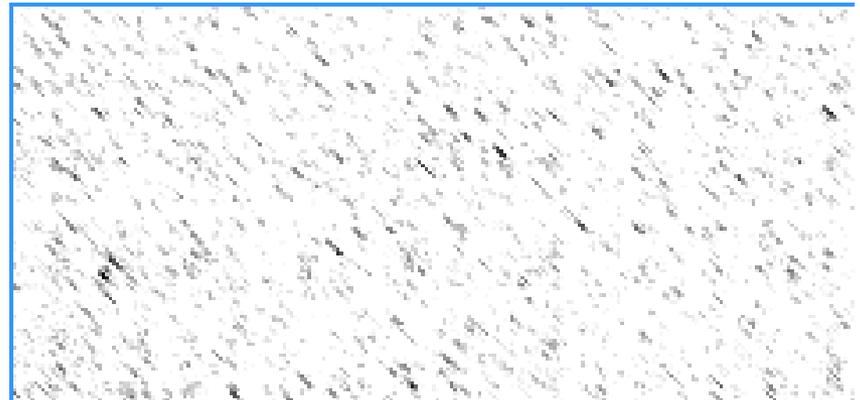
E. coli



At vs. E. coli



E. coli vs. 古细菌



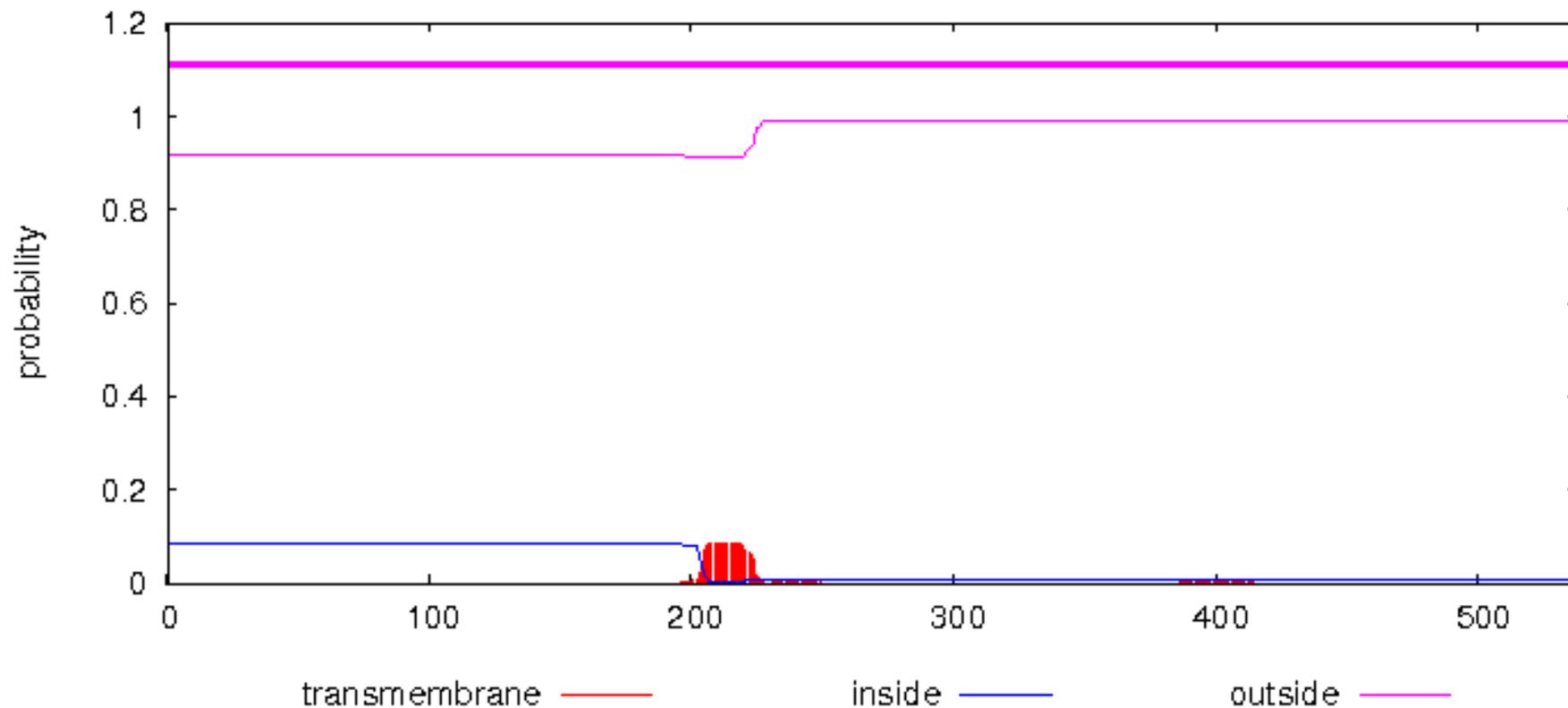
At vs. 古细菌

PEPCase在细胞中的定位

- PEPCase与光合作用有关。使用ChloroP 1.1进行检测，显示所选PEPCase并非定位于叶绿体。

Name	Length	Score	cTP	CS-score	cTP-length
sp_B1IVC2_CAPP_ECOLC	883	0.435	-	2.593	19
sp_B4F183_CAPP_PROMH	878	0.439	-	1.169	24
sp_B2VGA2_CAPP_ERWT9	883	0.436	-	1.067	24
sp_BOKS63_CAPP_PSEPG	875	0.427	-	1.706	68
sp_ABA764_CAPP_ECOHS	883	0.435	-	2.593	19
sp_A3Q9C4_CAPP_SHELP	878	0.432	-	4.688	5
sp_Q9V2Q9_CAPPA_PYRA	469	0.438	-	1.850	4
sp_Q27026_CAPPA_METT	483	0.436	-	8.052	5
sp_Q57764_CAPPA_PYRH	469	0.436	-	1.850	4
sp_Q0STS8_CAPPA_CLOP	537	0.427	-	5.475	5
sp_Q8TMG9_CAPPA_META	526	0.440	-	5.516	34
sp_Q8TYV1_CAPPA_METK	532	0.431	-	6.761	25
sp_Q02909_CAPP1_SOYB	967	0.428	-	1.865	9
gi_222619275_gb_EEE5	966	0.428	-	3.360	3
gi_586728416_ref_XP_	1085	0.426	-	9.642	54
tr_A9RUR8_A9RUR8_PHY	969	0.430	-	1.865	5
sp_P51062_CAPP_FEA	967	0.427	-	1.865	8
sp_Q6R2V6_CAPP2_CHLR	1221	0.428	-	0.689	34

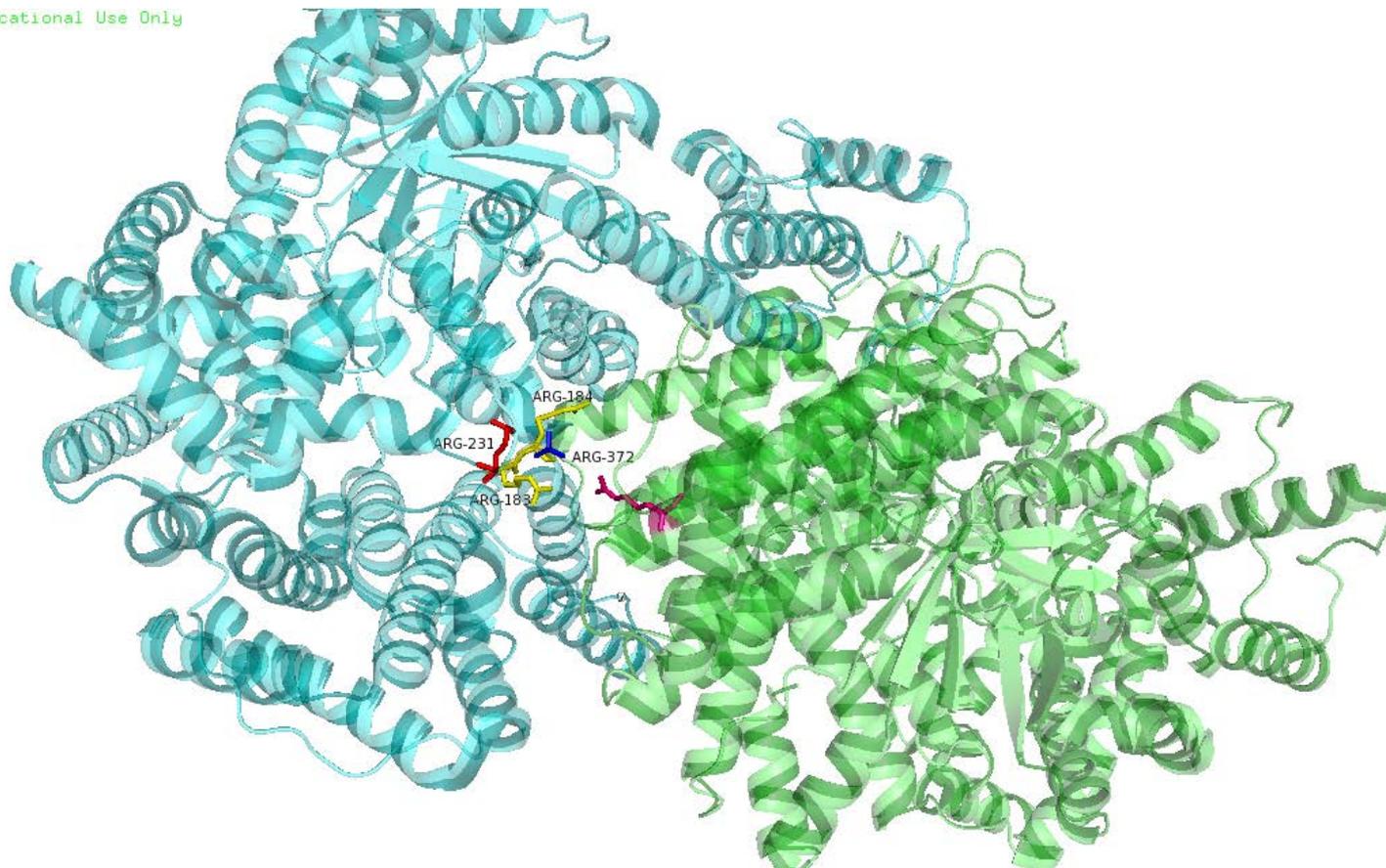
PEPCase在细胞中的定位



TMHMM v.2.0检测结果图。显示PEPCase中不存在跨膜结构域。
(<http://www.cbs.dtu.dk/services/TMHMM/>)

蛋白二聚体结构

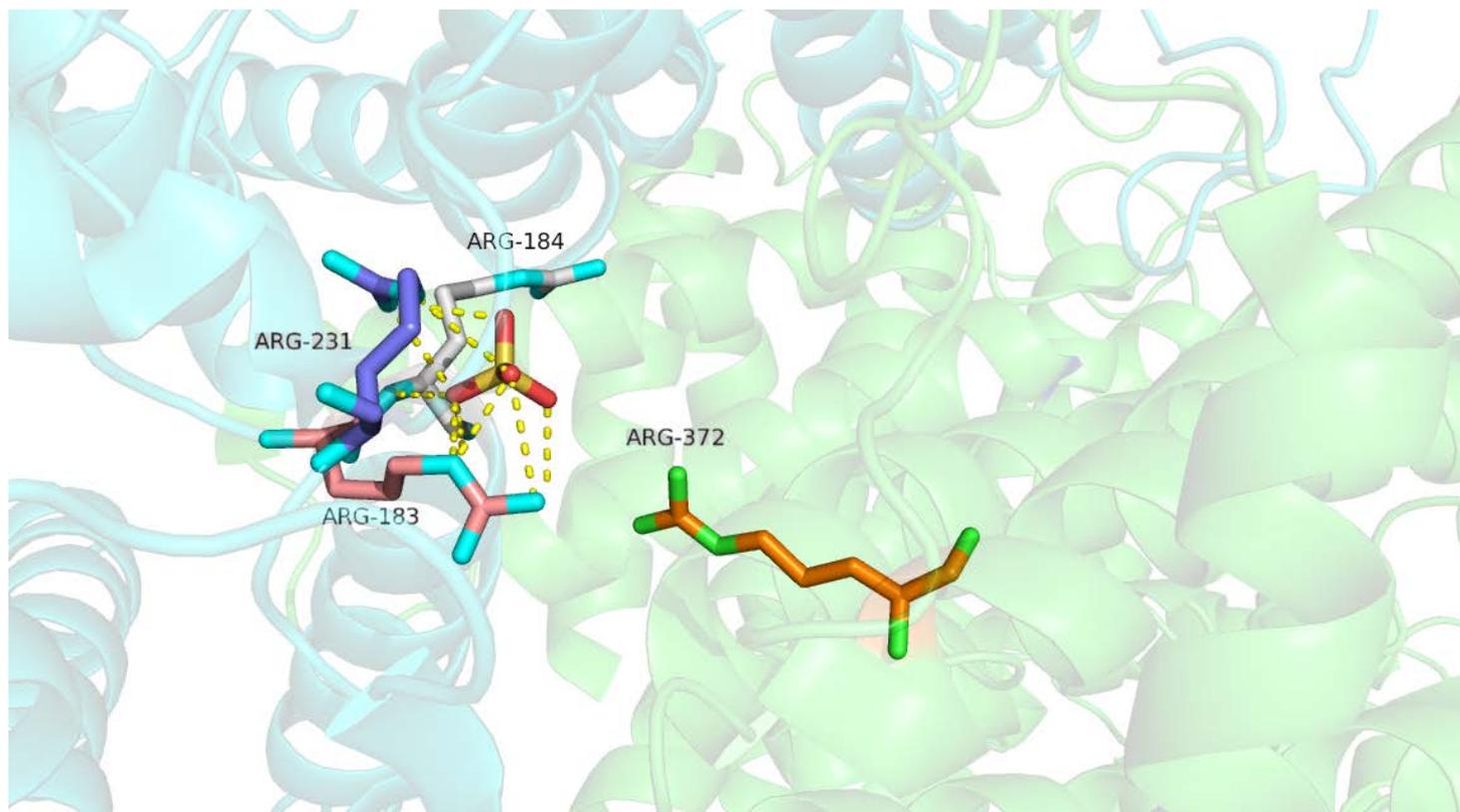
For Educational Use Only



(code: 1JQO, C4-form ppc from maize, 本图及以下蛋白结构图均使用pymol1.7.2.1完成)

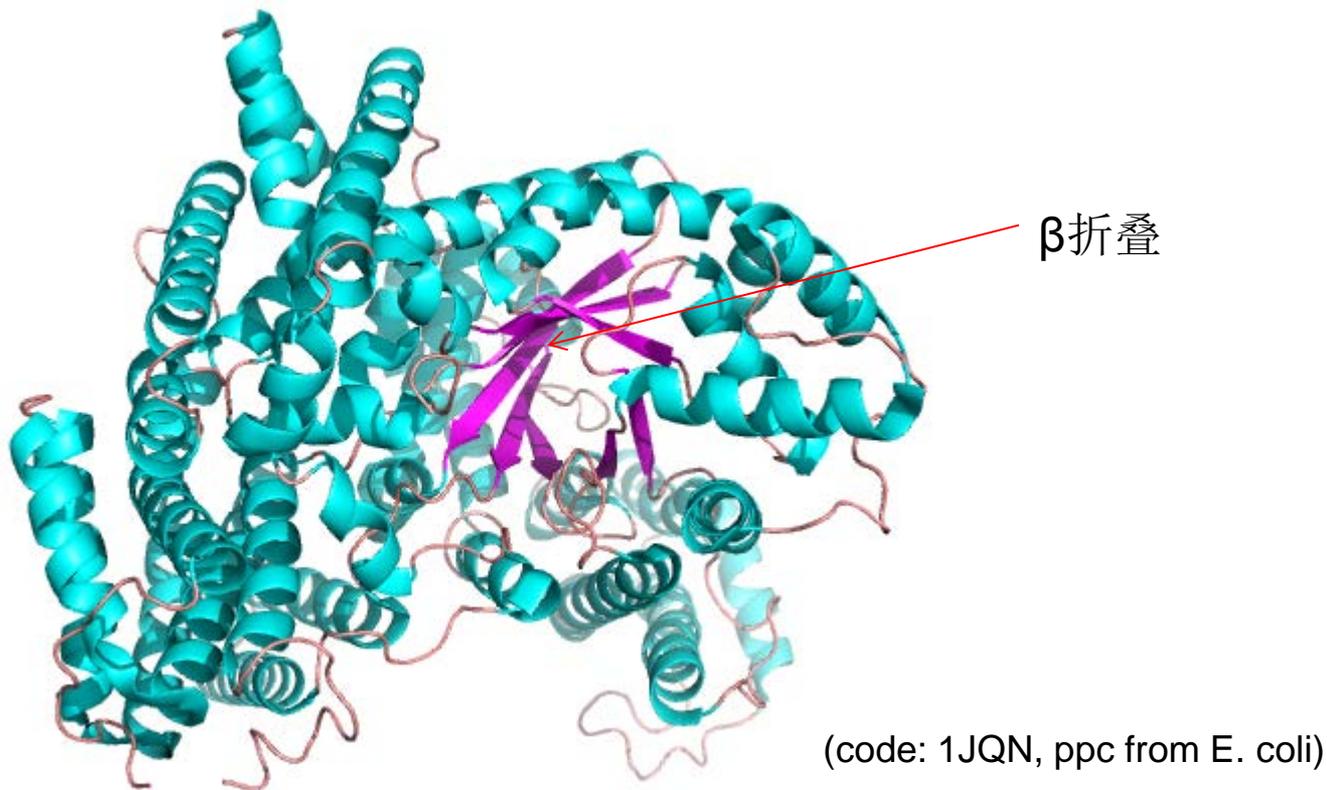
植物中PEPCase常以二聚体形式存在。富含 α 螺旋，其中单体往往含有约40个 α -helix(E.coli有38个)

α/β 亚基接触面



束缚硫酸盐(距活性中心 15\AA)被四个带有正电荷的氨基酸残基:R183、R184、R231 和R372 (R372 来自邻近亚基靠近结合处) 包围。与 α/β 亚基相互作用有关。

蛋白单体结构



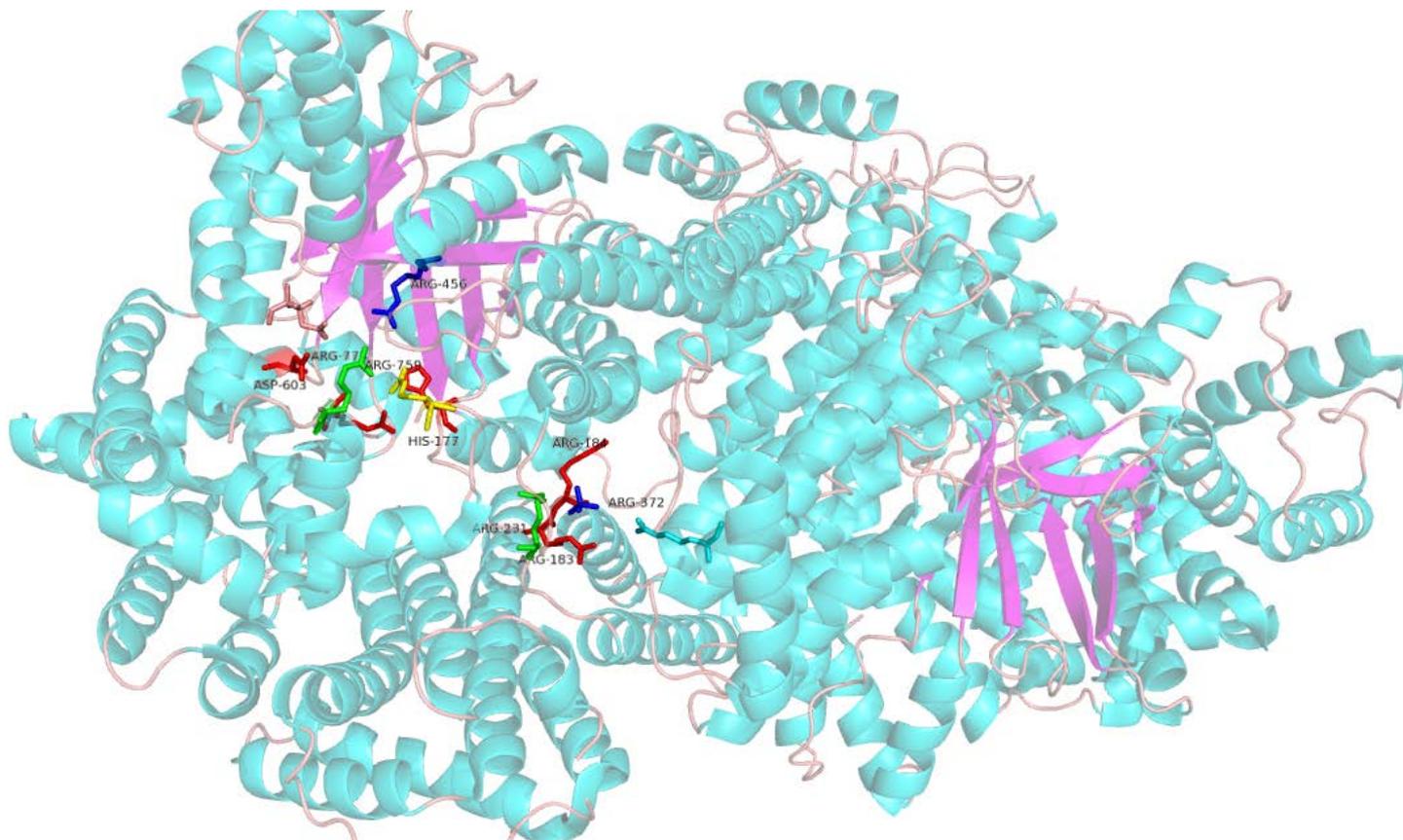
由 β 折叠形成桶状结构域，众多 α 螺旋环绕在周围。另外，周围有若干无规则蜷曲。没有明显的子域结构。

蛋白结构比对



使用pymol1.7.2.1对玉米 C4型 PEPCase 与大肠杆菌 PEPCase进行比对。结果显示二者高级结构基本相同。

催化位点



(code: 1JQO ppc maize)

催化位点位于 β 桶的一侧，由若干带正电荷的精氨酸与组氨酸，以及一个带正电的谷氨酸组成。催化位点距离结合面的精氨酸残基较近。

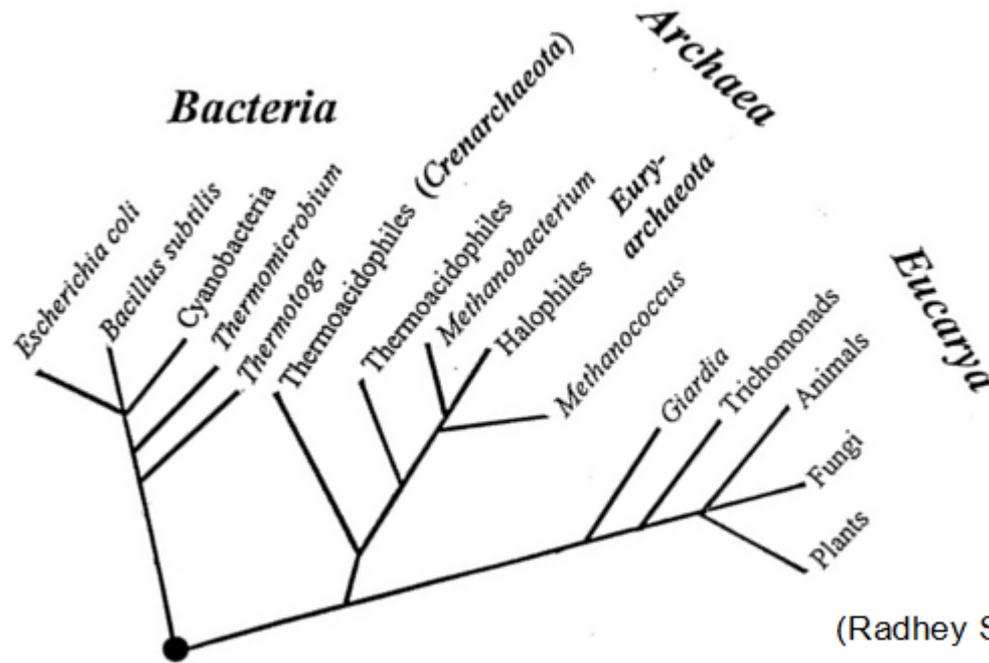
催化位点的保守性



催化位点170-190处的精氨酸（蓝）与组氨酸（红）是非常保守的。催化位点其他氨基酸也类似。(BioEdit 7.2.5)

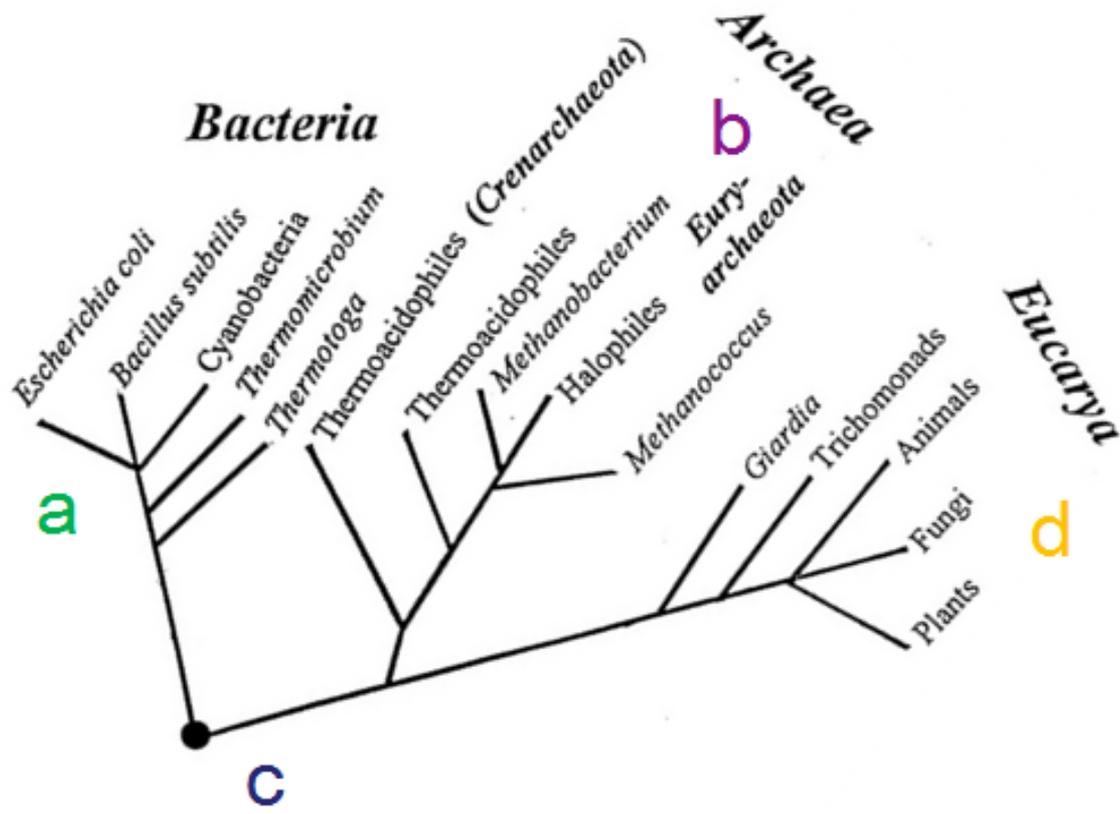
Discussion

- 一般认为，古细菌演化上处于真细菌与真核生物之间，较真细菌更为高等

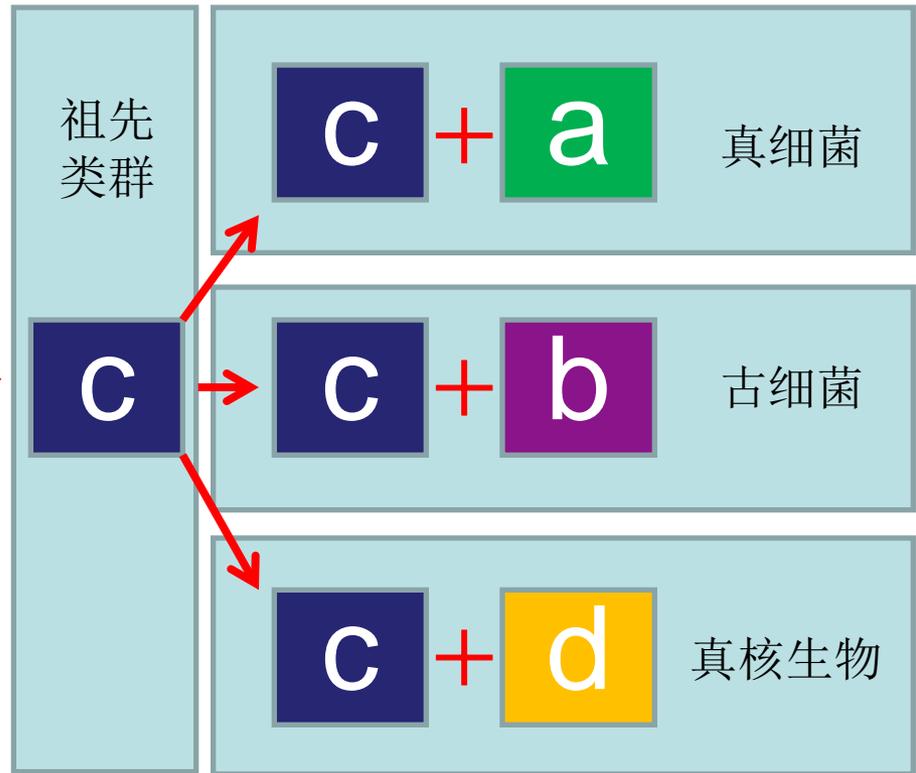
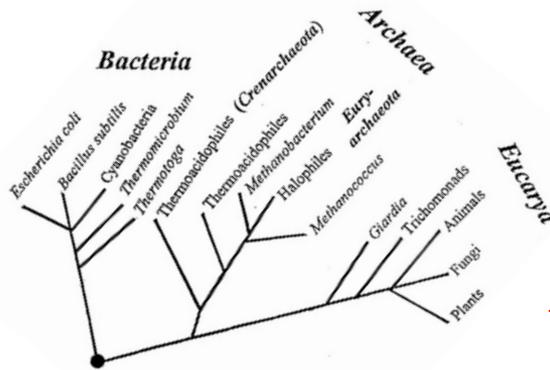
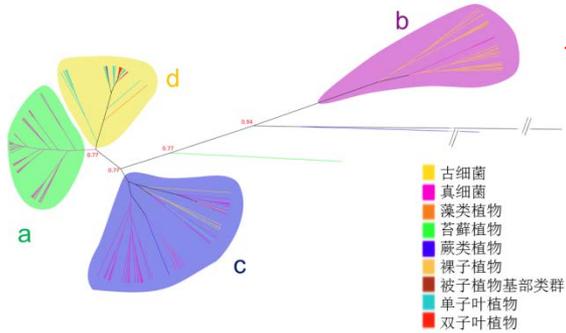


(Radhey S. Gupta, 1998)

PEPCase的演化



PEPCase演化过程猜想



PEPCase功能的演化

- **PEPCase**功能在不同类群中结构也很保守。
- 因此，**PEPCase**在各种代谢通路的演化中，可能仅仅是做为一个模块被招募。
- 因此，**C4**途径和**CAM**途径出现的关键，并不是**PEPCase**的创新性突变，而是调控其表达的方式发生了变化。

THANKS

